Literature review paper for the iterated prisoner's dilemma.

Nikoleta E. Glynatsi

2016

1 Introduction

The emergence of cooperation is a topic of continuing and public interest for social [23, 29], biological [30] and ecological sciences [31, 37, 46, 67]. Cooperation is essential for evolution but according to Darwins theory it is not always easy to achieve. The game called the prisoner's dilemma offers a theoretical framework for studying the emergence of altruist behaviour. Collecting data from 5 sources shows that more than 1170 papers related to the prisoner's dilemma have been published since it's origin.

In this work an extensive literature review will be presented. In addition, an introduction to the prisoner's dilemma is given in Section 2 and some major pieces of work will be discussed in Section 3. In Section 4 a comprehensive data set of literature regarding the prisoner's dilemma will be presented and analysed.

2 The Prisoner's Dilemma

The prisoner's dilemma a two player no-cooperative game where the decisions of the players are made simultaneously and independently. Both players can choose between cooperation (**C**) or defection (**D**).

The fitness of each player is influenced by its own behaviour, and the behaviour of the opponent. If both players choose to cooperate, both do better than if both defect. However, a player has the temptation to deviate. If a player was to defect while the other cooperates, the defector receives more than if both had cooperated. The reward for mutual cooperation is R units, for a mutual defection they receive P, and for cooperation-defection, the cooperator receives S where the defector receives T. Thus, the game's payoffs are given by,

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} \tag{1}$$

where $T > R > P > S$ and $2R > T + S$ are the conditions for a dilemma to exist. Due to rational behaviour and the knowledge that an individual is tempted to defect the game's equilibrium lies at a mutual defection and both players receive a payoff of P. Thus, the unbeatable strategy for the prisoner's dilemma is **D**.

Though both players end up with a worse payoff than if they had behaved otherwise and received a payoff of R. However, when the game is studied in a manner where prior outcomes matter, the defecting choice is no longer necessarily the unbeatable choice. The repeated form of the game is called the iterated prisoner's dilemma and now players interact more than just once.

In Section 3 it will be discussed how it was proven that the iterated prisoner's dilemma leaves room for cooperation. The prisoner's dilemma has attracted much attention and that's shown in Figure 1. Figure 1 illustrates the number of publications on the prisoner's dilemma per year from the following sources:

- arXiv;
- PLOS;
- IEEE;
- Nature;
- Springer.

The choice of sources is due to the fact that they have an open access API which will be discussed in more detail in Section 4. The data collection and the open source library used to generate the time plot will be described more comprehensively in Section 4. There are various specific timepoints that will be discussed in Section 3.

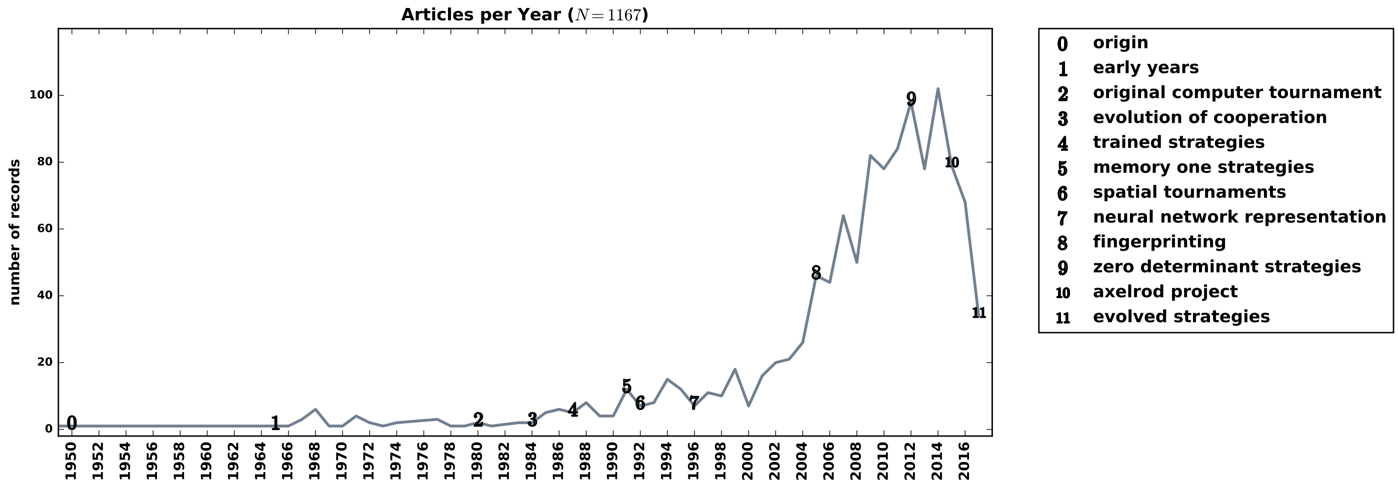

Figure 1: A timeline highlighting the milestones of the prisoner's dilemma.

3 Timeline

3.1 Origin and the 60's

The origin of the prisoner's dilemma go back to 1950s in early experiments conducted in RAND [25] to test the applicability of games described in [66]. In [25] the two player game was introduced but the name behind the game was given later the same year. A. W. Tucker, the PhD advisor of John Nash, in an attempt to delivery the game with a story during a talk used prisoners as players and the game is known as the prisoner's dilemma ever since [62].

The study of the prisoner's dilemma has attracted people from various fields across the years. An early figure within the field is Professor Anatol Rapoport, a mathematical psychologist, whose work focused on peacekeeping. In his early work [54] Rapoport conducted experiments using humans to simulate a play of the prisoner's dilemma. Experimental groups were not been used only by Rapoport but it was a common mean of studying the game [24, 27, 43, 44, 58] and are still being use to date.

Several experiments explored the conditions under which altruist behaviour emerges in human societies. Conditions that were studied include, the gender [24, 43, 44] of individuals, the representation of the game [24], the distance between players [58], the start effect [61] and whether the experimenter has been biaised [27].

Though the aspects of representation, distance and starting effect were proven in the 60's to have a significant effect on the cooperating ratio of the individuals, the aspect of gender was found to be insignificant. An interesting result has been the effect of the representation, or the explanation of the game itself. It shows that when people are playing the prisoner' dilemma any minor change in the presentation of the game can have an effect on their interpretation of the game.

Other experiments focused on analysing the play of a test subject. Researchers believed that they could identify an unbeatable strategy to play the game. Inspired by the work of Rapoport and the idea that AI was now being trained to play a game of chess the political scientist Robert Axelrod performed the first ever computer tournament, known to the author, of the iterated prisoner's dilemma [16, 18].

3.2 Axelrod's Tournaments

In 1980 [13] a computer tournament of the iterated prisoner's dilemma took place. R. Axelrod invited 14 participants to submit a strategy written in the programming languages Fortran or Basic, the names of the attendants were the following,

1. T Nicolaus Tideman and Paula Chieruzz;
2. Rudy Nydegger;
3. Bernard Grofman;
4. Martin Shubik;
5. Stein and Anatol Rapoport;
6. James W Friedman;
7. Morton Davis;
8. Jim Graaskamp;
9. Leslie Downing;
10. Scott Feld;
11. Johann Joss;
12. Gordon Tullock;
13. Name not given;

Each strategy played against all the 14 opponents, itself and a player that played randomly a match of 200 turns. This topology is called round robin and is the equivalent of a complete graph. The tournament was repeated 5 times to reduce variation in the results. Each participant knew the exact length of the matches and had access to the full history of each match. Furthermore, Axelrod performed an preliminary tournament and the results were known to the participants. The payoff values used where $R = 3, P = 1, T = 5$ and $S = 0$. These values are commonly used in the literature and unless specified will be the values used in the rest of the work 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 submitted by Rapoport and was called **Tit For Tat**. Tit for Tat, is a strategy that always cooperates on the first round and then mimics the opponent's previous move. Tables 1 2 show two examples of Tit for Tat interacting with two very deterministic opponents. An opponent refereed here as **Cooperator**, that always cooperates and and opponent that always defects, **Defector**.

Turns	Tit for Tat	Cooperator
1	C	C
2	C	C
3	C	C
⋮	⋮	⋮
200	C	C

Table 1: Tit for Tat example match against Cooperator

Turns	Tit for Tat	Defector
1	C	D
2	D	D
3	D	D
⋮	⋮	⋮
200	D	D

Table 2: Tit for Tat example match against Defector

The results of the first tournament was filled with surprises. Tit for Tat won but was the simplest strategy of all. Tit for Tat managed to defeat even entrants that tried to improve on Tit for Tat after the preliminary tournament.

Axelrod explained the reasons behind Tit for Tat's success. The top eight ranked strategies have been strategies that never defected on the first round, thus they were 'nice'. Amongst those top strategies the results were determined by just two of the other seventh rules. Those strategies were described as kingsmakers and they were the strategies submitted by Grofman and Graaskamp. The final property that was described to have affect Tit for Tat's performance compared to the rest strategies was 'forgiveness'. Tit for Tat punished it's opponent for a defection but just once.

Axelrod was the first to speak about 'niceness' and 'forgiveness' as properties of a winning strategy for the iterated prisoner's dilemma. To further test the robustness of the results Axelrod performed a second tournament [14]. A total of 63 participants submitted strategies for the second tournament, their names were the following,

1. Gail Grisell;
2. Harold Rabbie;
3. James W Friedman;
4. Abraham Getzler;
5. Roger Hotz;
6. George Lefevre;
7. Nelson Weiderman;
8. Tom Almy;
9. Robert Adams;
10. Herb Weiner;
11. Otto Borufsen;
12. R D Anderson;
13. William Adams;
14. Michael F McGurrin;
15. Graham J Eatherley;
16. Richard Hufford;
17. George Hufford;
18. Rob Cave;
19. Rik Smoody;
20. John Willaim Colbert;
21. David A Smith;
22. Henry Nussbacher;
23. William H Robertson;
24. Steve Newman;
25. Stanley F Quayle;
26. Rudy Nydegger;
27. Glen Rowsam;
28. Leslie Downing;
29. Jim Graaskamp and Ken Katzen;
30. Danny C Champion;
31. Howard R Hollander;
32. George Duisman;
33. Brian Yamachi;
34. Mark F Batell;
35. Ray Mikkelson;
36. Craig Feathers;
37. Fransois Leyvraz;
38. Johann Joss;
39. Robert Pebly;
40. James E Hall;
41. Edward C White Jr;
42. George Zimmerman;
43. Edward Friedland;
44. X Edward Friedland;
45. Paul D Harrington;
46. David Gladstein;
47. Scott Feld;
48. Fred Mauk;
49. Dennis Ambuehl and Kevin Hickey;
50. Robyn M Dawes and Mark Batell;
51. Martyn Jones;
52. Robert A Leyland;
53. Paul E Black;
54. T Nicolaus Tideman and Paula Chieruzz;
55. Robert B Falk and James M Langsted;
56. Bernard Grofman;
57. E E H Schurmann;
58. Scott Appold;
59. Gene Snodgrass;
60. John Maynard Smith;
61. Jonathan Pinkley;
62. Anatol Rapoport.

All the participants knew the results of the previous tournament. The rules were similar to those of the first tournament with only one exception; the number of turns was not specified. A probabilistic ending tournament was meant to be used. In a probabilistic ending tournament each match has probability of ending after each turn. This is also refereed as 'shadow of the future' [17].

However, the tournament was not a probabilistic ending one. A fixed probability of 0.0036 was chosen as a chance of ending a match with each given move. The value was chosen so that the expected median length of a match would be 200 turns. The topology was of a round robin and each pair of players was matched 5 times. The length of the matches was determined once by drawing a random sample. Each of the five matches had a length of 63, 77, 151 and 308.

The results of the tournament once again came as a surprise. Tit for Tat was the simplest submission in the second tournament and won the second tournament as well. Tit for Tat provided proof that reciprocity behaviour can allow cooperation to emerge in the iterated prisoner's dilemma game. In [18] the main conclusions indicating strong performance was:

- that it start of by cooperating
- it would forgive it's opponent after a defection
- after opponents identified that they were playing Tit for Tat choose to cooperate for the rest of the game.

Though a full explanation of all 13 strategies in given in [18] the same does not hold for all 63 strategies of the second tournament [18]. The author mainly focuses on the high ranked participants. However, the source code for of the strategies be found on Axelrod's personal website [1]. The code has been written by Axelrod and several other contributors in the programming languages Fortran or Basic. The strategies written in Basic were translated to Fortran before the tournament. The source code includes the code only for the strategies and not for creating and performing the tournament.

The code for the winning strategy Tit for Tat is illustrated in Figure 2. A few strategies were submitted in Basic but where translated into Fortran by Axelrod's team.

```
    FUNCTION K92R(J,M,K,L,R, JA)
C BY ANATOL RAPOPORT
C TYPED BY AX 3/27/79 (SAME AS ROUND ONE TIT FOR TAT)
c replaced by actual code, Ax 7/27/93
c  T=0
c   K92R=ITFTR(J,M,K,L,T,R)
      k92r=0
      k92r = j
c test 7/30
c   write(6,77) j, k92r
c77   format(' test k92r. j,k92r: ', 2i3)
      RETURN
      END
```

Figure 2: Source code for Tit for Tat in Fortran. Provided by [1].

Unfortunately, the source code of the first tournament is not available as stated in Axelrod's personal website [1].

Other successful strategies from Axelrod's tournament that can been seen in literature to date are,

- **Grudger** is a strategy that will cooperate as long as the opponent does not defect. The name Grudger was give to the strategy in [41]. Though the strategy goes by many names in the literature such as, Spite [20], Grim Trigger [19] and Grim [65].
- etc

3.2.1 Mis implementation

Success often comes with criticism. Axelrod's tournaments assumed that each player has perfect information of the opponent's actions. In real life situations this is not always the case. Colleagues' interactions often suffer from measures of uncertainty. In the original tournaments there was no possibility of mis implementation or misunderstanding. These stochastic variations are refereed to as **noise** and **mis perception**. Noise is the concept of flipping one's move based on a given probability. On the contrary, mis perception is the probability that the opponent's current move is flipped before being recorded. Noise will flip a player's action and it will be recorded correctly in the history where mis perception will not have an effect on the player's move but it will be recorded wrong [35].

The performance of Tit for Tat was proven to suffer from such stochasticity in the tournament environment, especially against itself [5, 31, 48, 49, 59]. If two strategies playing Tit for Tat were to compete against each other in a noisy environment the strategies will get a series of unwanted defections. In a non noisy environment the two strategies would have been cooperating for the entire match. An interesting result was introduced by [48]. Molander stated that if two strategies playing Tit for Tat meet in a noisy match the average payoff that a strategy will receive will be the same as that of a Random player (with probability 0.5 of cooperating).

In [5] a similar tournament to that of Axelrod was performed but this time noise was used. Bendor invited academics to submit strategies to participate in his tournament. A total of thirteen strategies were used including already existed strategies such as Tit for Tat and **Tit for 2 Tats**. The results showed that Tit for Tat performed purely placing eight in the tournament. Bendor stated that a more forgiving strategy was needed, in his tournament a strategy called **Nice and Forgiving** ranked first.

The work of [51] following a similar approach agreed with this result. In [51], the space of re-active strategies was explored in a noisy environment. The strategy that was performing the best in that environment was the re-active strategy known as **Generous Tit for Tat**. A reactive strategy is a strategy that consider only the past move of the opponent, but they will be discussed later on in more detail. Generous Tit for Tat, attracted attention as it was a generous variant of the famous strategy Tit for Tat able to withstand noisy environments.

The author published yet another paper three years later introducing another interesting player. The new strategy had the tolerance of Generous Tit for Tat but also the capability of resisting and invading an all-out cooperators population was. The strategy is called **Pavlov**, and is based on the fundamental behavioural mechanism win-stay, lose-shift. The strategy starts off with a **C**, then Pavlov will repeat it's last move it was awarder with by R or T but will shift if punished by P or S.

3.3 Strategies Stability

So far it has been discussed that a strategy's dominance is tested through the performance of the strategy in a tournament against other strategies. But is the overall success of a strategy based only on it's performance in a round robin tournament or should it be checked through other ways as well?

Following his initial tournaments Axelrod performed an 'ecological' tournament in 1981 [18]. In [18], the set of strategies from Axelrod's second tournament was used to perform the ecological tournament. The 63 strategies interacted generation after generation to a round robin competition where their frequencies were proportional to their payoff in the previous round. The ecological approach is based on the payoff matrix of the tournament. The highest performing strategies are adapted by lower scoring individuals within a fixed population. Over time a strategy takes over the population. Figure 3 demonstrates an example of the natural selection proceeder.

The ability of strategies to be favoured under natural selection and their ability to withstand invasion from other strategies soon became an new measure of performance; refereed to as the stability of a strategy.

In [18], the results showed that in a homogeneous population of Tit for Tat invasion by mutant strategies was not successful.

The results of [22] argued that no pure strategy is evolutionary stable in the iterated prisoner's dilemma. This was not proven analytically, instead a series of examples using strategies such as Tit for Tat, Suspicious Tit for Tat and Defector where explored; a very constrained set of strategies.

The results were questioned by [45], stating that much was still no fully explored and more research had to be put into the results. Another attempt to explore stability of strategies in the prisoner's dilemma was done in [21]. This time exploring the results in a noisy environment, but similarly a analytical proof was not achieved.

3.4 Spatial tournament

An extension to the natural selection was introduced in the 1992 [52], recommending a different type of topology. A population of two deterministic strategies, Defector and Cooperator, were placed on a a two dimensional square array

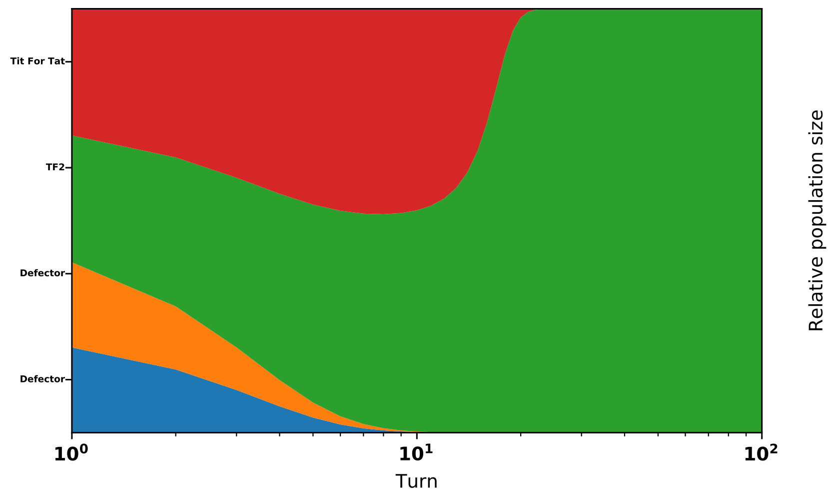

Figure 3: System evolving over time based on natural selection using [6].

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 4. The authors claimed that the essential results remain true of all topologies; the results also hold whether self interactions are taken into account.

Thus each cell of the lattice is occupied by a **C** or a **D** and in each generation step each cell owner interacts with its immediate neighbours and play the game. The score of each player is the sum of the overall games the player competed in. At the start of the next generation, each lattice cell is occupied by the player with the highest score among the previous owner and the immediate neighbours. Nowak and all created this model where the model parameter has been the temptation payoff denoted as b. Thus $T = b$. For different values of the parameter b it was shown that cooperators and defectors can persist together indefinitely. This topology is refereed to as spatial topology.

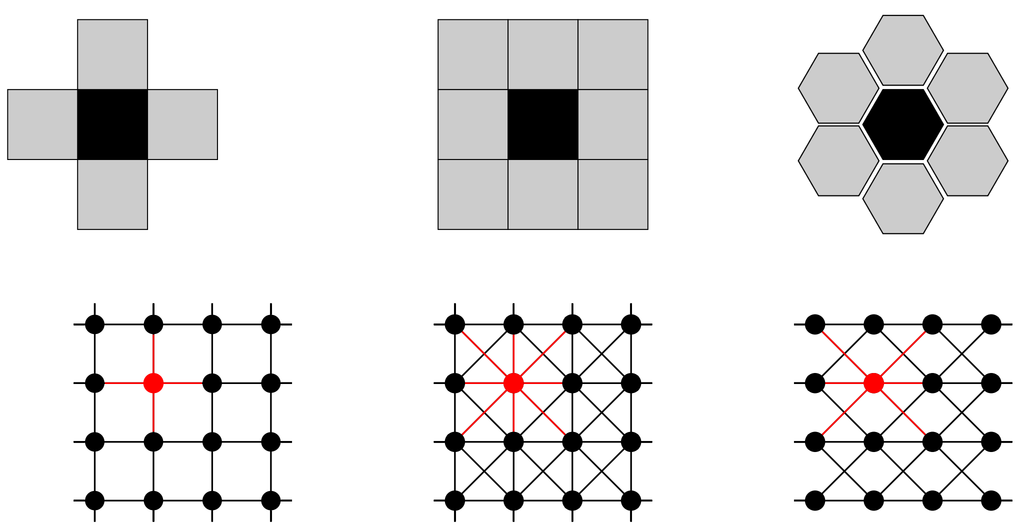

Figure 4: Spatial neighbourhoods

3.5 Tit for Tat variants

Several other strategies were introduced as more generous versions and described as more dominant than Tit for Tat. These include, **Contrite Tit for Tat** [68] and **Adaptive Tit for Tat** [64]. **Tit for Two Tats** [17], which defects only if the other player defected on the two preceding moves. On the other hand, defector variants have also been studied [34]. **Anti Tit for Tat**, is a strategy that plays the opposite of the opponents previous move. Another limitation of the strategy was discussed in [59]. Tit for Tat was proven to hit a deadlock. Deadlock meaning a loop between cooperation and defection. **Omega Tit For Tat** was introduced and was a strategy capable of avoiding the deadlock [59].

3.6 Gradual and Handshake

Other strategies that made an impact have been **Gradual** [20] and **Handshake** [56] presented in 1997 and 1989 respectively. Gradual starts off by cooperating, then after the first defection of the other player, it defects one time and cooperates twice. After the second defection of the opponent, it defects two times and cooperates twice. After the n^{th} defection it reacts with n consecutive defections and then two cooperations. Handshake is a strategy that starts with cooperation, defection. If the opponent plays in a similar way then it will cooperate forever, otherwise it will defect forever.

3.7 Li, Hingston and Kendall Tournament

In 2011 the authors of [40] performed their own tournament where several interesting strategies made an appearance.

- **Periodic player CCD**, plays **C**, **C**, **D** periodically. Note that variations of a period player also make appearance in the article but will not be listed here.
- **Prober**, starts with the pattern **D**, **C**, **C** and then defects if the opponent has cooperated in the second and third move; otherwise, it play as Tit for Tat.
- **Reverse Pavlov**, a strategy that does the reverse of Pavlov.

In earlier work the same author introduced a strategy called **APavlov**, which stands for adaptive Pavlov [39]. The strategy attempts to classify the opponent as one of the following strategies, All Cooperator, All Defector, Pavlov, Random or **PavlovD**. PavlovD, is just Pavlov but it starts the game with a **D**. Once Adaptive Pavlov has classified the opponent plays to maximize it's payoff.

3.8 Memory One Strategies

Reactive strategies are a subset of memory one strategies introduced in 1989 [50]. Reactive strategies are denoted by the probabilities to cooperate after a **C** and a **D** of the opponent. Thus, a reactive strategy only considers the previous turn of the opponent. Strategies such as, Tit for Tat and Generous Tit for Tat are reactive.

Memory one strategies, are a set of strategies that consider only the last turn of the game to decide on the next action [51]. They are represented by the four conditional probabilities p_1, p_2, p_3 and p_4 to cooperate after CC, CD, DC and DD respectively (the four possible states a player can be in if only the last turn of the game was to be considered). Reactive strategies are just a constrained version where $p_1 = p_3$ and $p_2 = p_1$. The first action of the strategy (when the history does not exist yet) is assumed to be **C** unless is stated otherwise. For example, a reactive strategy called **Suspicious Tit for Tat**, studied in [49], has the same representation as Tit for Tat but plays **D** in the first round.

In [53], a new set of memory one strategies were introduced, called **zero determinant (ZD)** strategies. The ZD strategies, manage to force a linear relationship between the score of the strategy and the opponent. Press and Dyson, prove their concept of the ZD strategies and claim that a ZD strategy can outperform any given opponent.

The ZD strategies have attracted a lot of attention. It was stated that "Press and Dyson have fundamentally changed the viewpoint on the Prisoner's Dilemma" [60]. In [60], a new tournament was performed including ZD strategies and a new set of ZD strategies the **Generous ZD**. Even so, ZD and memory one strategies have also received criticism. In [38], the 'memory of a strategy does not matter' 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.

3.9 Ashlock's work

Another worth mentioning figure is Daniel Ashlock. In [12] the author presented two new strategies. These strategies have been trained using a finite state machine representation. They are called, **Fortress3** and **Fortress4**. Figure 5 illustrates their diagrammatic representation where the transition arrows are labelled O/P where O is the opponent's last action and P is the player's response.

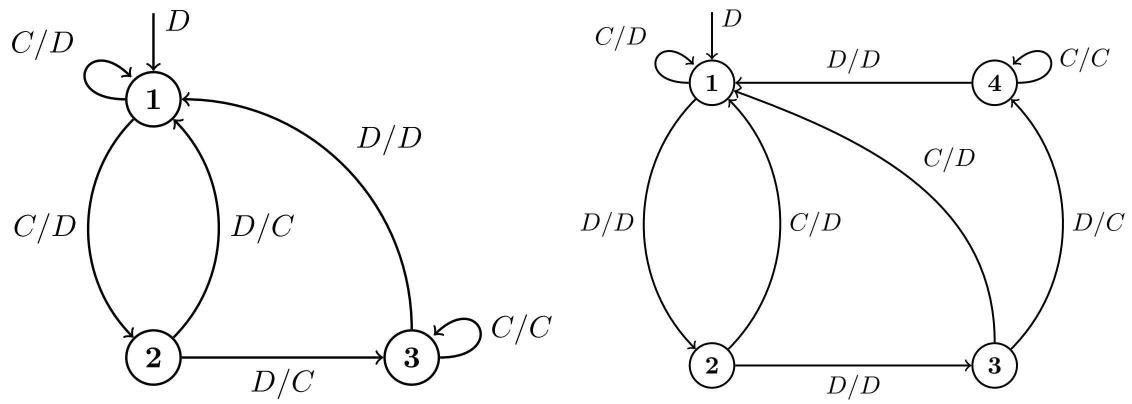

Figure 5: Representations of Fortress 3 and Fortress 4. Note that the strategy's first move, enters state 1, is defection for both strategies.

Finite state machines are commonly used to represent iterated prisoner's dilemma strategies [47, 57]. Strategies based on finite state machines are described by the number of states. The strategy selects the next action in each round based on the current state and the opponent's last move, transitioning to a new state each time. Figure 6, illustrates the finite state representation of Tit For Tat.

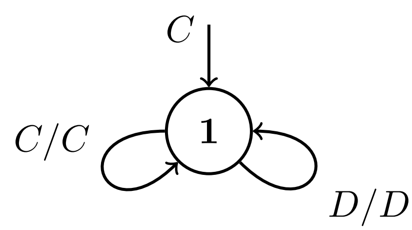

Figure 6: Finite state machine representation of Tit for Tat.

3.9.1 Fingerprinting

With a large number of strategies and different representations in the literature a question was soon risen. How can we be sure that all these strategies are indeed different to each other? In [7] a method called fingerprinting was given as an answer to the problem. The method of fingerprinting is a technique for generating a functional signature for a strategy [8]. 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 [7] Tit for Tat is used as the probe strategy. Fingerprint functions can then be compared to allow for easier identification of similar strategies. In Figure 7 an example of Pavlov's fingerprint is given. Fingerprinting has been studied in depth in [8, 9, 10, 11].

3.10 Evolved strategies

Complex strategies are defined as a set of strategies tha can use a variety of features computed from the history of play. The term complex can also be refereed to strategies that have been trained with evolutionary methods to be dominant. In [15], Axelrod used an evolutionary algorithm to identify a strategy that was equal to or better than Tit for Tat.

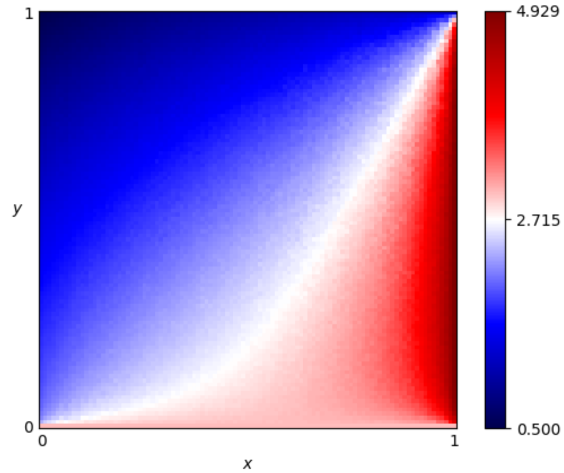

Figure 7: Pavlov fingerprinting with Tit for Tat used as the probe strategy. Figure was generated using [6].

Other representation methods include lookup tables [15, 42] and artificial neural networks [33, 38]. In [15], lookup tables are introduced as a mean of representing a strategy in a gene format. A lookup table is a set of deterministic responses based on the opponents m last moves; [15] considered $m = 3$. Figures 8 shows a look up representation of Tit for Tat where $m = 1$.

Opponent's last move	**Next action**
D	***D***
C	***C***

Figure 8: Lookup table representation of Tit for Tat.

Similarly, artificial neural networks provide a mapping function to an action based on a selection of features computed from the history of play. A number of strategies based on artificial neural networks are introduced by [32]. These strategies are refereed to as **EvovlvedANN** strategies and are based on a pre-trained neural network with the following features,

- Opponent's first move is C
- Opponent's first move is D
- Opponent's second move is C
- Opponent's second move is D
- Player's previous move is C
- Player's previous move is D
- Player's second previous move is C
- Player's second previous move is D
- Opponent's previous move is C
- Opponent's previous move is D
- Opponent's second previous move is C
- Opponent's second previous move is D
- Total opponent cooperations
- Total opponent defections
- Total player cooperations
- Total player defections
- Round number

A representation of **EvovlvedANN 5** is given in Figure 9. The inputs of the neural network are the 17 features as listed above. Number 5 reefers to the size of the hidden layer.

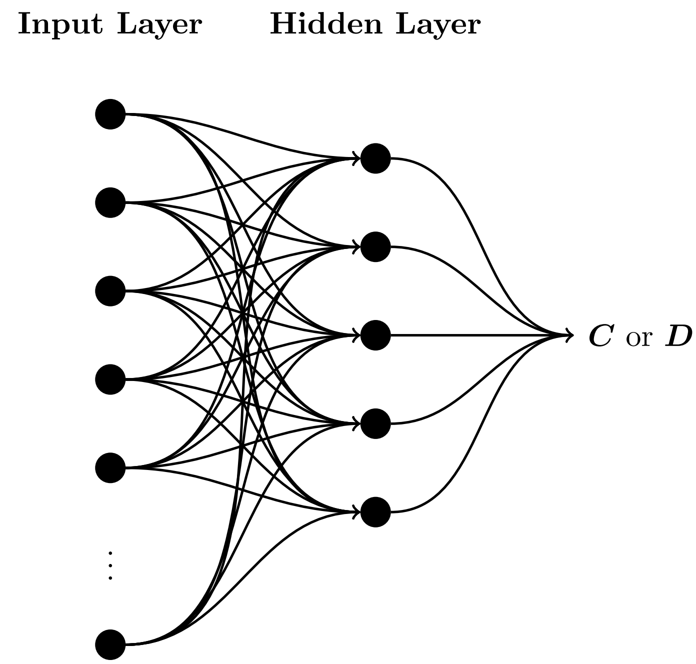

Figure 9: Neural network representation of EvovlvedANN 5.

In [32], these representing methods are refereed to as archetypes. Finite state machines and artificial neural networks are included in the work but also new archetypes are introduced, such as hidden Markov models. A variant of a finite state machine that use probabilistic transitions based on the prior round of play to other states and cooperate or defect with various probabilities at each state. Finite state machines and hidden Markov models based strategies are characterized by the number of states. Similarly, artificial neural networks based players are characterized by the size of the hidden layer and number of input features.

Additionally a variant of a look up table is also presented called the lookerup archetype. The lookerup archetype responses based on the opponent's first n_1 moves, the opponent's last m_1 moves, and the players last m_2 moves. Taking into account the initial move of the opponent can give many insights. For it is the only move a strategy is truly itself without being affected by the other player. As a reminder, Axelrod in his work highlighted the importance of the initial move and believed that it was one of the secrets of success of the strategy Tit for Tat.

Finally, a new archetype called the Gambler is also introduced, which is a stochastic variant of the lookerup archetype.

Archetypes are used with evolutionary algorithms to train set of new strategies. The evolutionary algorithm used in both [15, 28] is called genetic algorithm. Other algorithms including particle swarm optimization have been used in research of the most dominant strategy [26].

In [32] the approach in used to introduce as stated by the authors the best performing strategies for the iterated prisoner's dilemma. These strategies will be refereed as **Evolved** strategies. Several successful new strategies are,

- **EvolvedLookerUp2_2_2** a looker up strategy trained with a genetic algorithm; EvolvedLookerUp2_2_2 responses based on the opponent's 2 first and last moves and the player's 2 last moves. Thus $n_1 = 2, m_1 = 2$ and $m_2 = 2$.
- **Evolved HMM 5** a 5 states hidden markov model trained with a genetic algorithm;
- **Evolved FSM 16** a 16 state machine trained with a genetic algorithm;
- Finally **PSO Gambler 2 2 2** a looker up strategy trained with a particle swarm algorithm, where $n_1 = 2, m_1 = 2$ and $m_2 = 2$.

Though several papers have claimed before to have discovered the dominant strategies for the game the work of [32] is promising. This is due the fact that the introduced strategies have been trained using different types of evolutionary algorithms in a pool of 176 well known strategies for the literature. Including all the strategies that have been discussed in this section.

This was made possible due an open source library, called the Axelrod project [6]. The project is written in the programming language Python, it is accessible and open source. To date the list of strategies implemented within the library exceed the 200. The project has been used in several publications including [32] and a paper describing it and it's capabilities was published in 2016 [36]. The source code for Tit for Tat as implement within the library is shown in Figure 10. Furthermore, performing a tournament with a selection of strategies is possible in five lines of code, shown in Figure 11.

```python
def strategy(self, opponent: Player) -> Action:
    """This is the actual strategy"""
    # First move
    if not self.history:
        return C
    # React to the opponent's last move
    if opponent.history[-1] == D:
        return D
    return C
```

Figure 10: Source code for Tit for Tat in Python as implemented in Axelrod Python library [6]

.

```python
>>> import axelrod as axl
>>> players = (axl.Cooperator(), axl.Defector(), axl.TitForTat(), axl.Grudger())
>>> tournament = axl.Tournament(players)
>>> results = tournament.play()
>>> results.ranked_names
['Defector', 'Tit For Tat', 'Grudger', 'Cooperator']
```

Figure 11: Performing a computer tournament using [6].

Software has a crucial role in research. Well written and maintained software allows the reproducibility of prior work and can accelerate findings within the field. The field of the iterated prisoner's dilemma has suffered the consequences of poor research software. As stated above the source code of the initial computer tournament is not retrievable. Several of the strategies that competed in the tournament are not given a full explanation of how the decided on their next move. In terms of best practice and reproducibility the Axelrod library is the lead software in the field.

3.11 Other Software

Due the nature of the research regarding the iterated prisoner's dilemma several software packages have been created in order to simulate computer tournaments.

Another piece of software includes a library called PRISON [4]. PRISON is written in the programming language Java and it has been used by it's authors in several publications. The project includes a good number of strategies from the literature but unfortunately the last update of the project dates back in 2004.

More recent projects include [2, 3], both are education platforms and not research tools. In [2], several concepts such as the iterated game, computer tournaments and evolutionary dynamics are introduced through a user interface game. Project [3] offers a big collection of strategies and allows the user to try several match and tournaments configurations. Such as noise.

3.12 Applications

3.12.1 Social Applications

3.12.2 Ecological Applications

The reciprocal period of the prisoner's dilemma spread the knowledge of the game not only worldwide but also across different scientific principles. The study of cooperation was once again a critical issue. The applications of the game soon found their way to ecological studies, for example [46] conducted an experiment using sticklebacks to test the robustness of the strategy Tit for Tat in the interactions of fish. Fish usually travel in pairs and monitor their hunters to gain information on the enemy. Other works that include applications to ecological settings have been those of [31, 67]. There the reciprocal food sharing between vampire bats was studied.

3.12.3 Biological Applications

- [63] uses evolutionary game theory to study the spread of virus.
- [30] a shout for his work, using tit for tat to study cells.

3.12.4 not sure

In [55], the authors claim that they have managed to re-run the first tournament that Axelrod performed. They tried to push his work further by altering aspects such as, the format of the tournament, the objective and the population. One of the authors claimed to have been a contributor to the first tournaments, which would explain how it was managed to reproduce the tournament.

4 Analysis

References

[1] Complexity of cooperation web site. `http://www-personal.umich.edu/~axe/research/Software/CC/CC2.html`. Accessed: 2017-10-23.

[2] The evolution of trust. `http://ncase.me/trust/`. Accessed: 2017-10-23.

[3] The iterated prisoner's dilemma game. `http://selborne.nl/ipd/`. Accessed: 2017-10-23.

[4] Lifl (1998) prison. `http://www.lifl.fr/IPD/ipd.frame.html`. Accessed: 2017-10-23.

[5] When in doubt... cooperation in a noisy prisoner's dilemma. *The Journal of Conflict Resolution*, 35(4):691–719, 1991.

[6] The Axelrod project developers . Axelrod: ¡release title¿, April 2016.

[7] Daniel Ashlock and Eun-Youn Kim. Techniques for analysis of evolved prisoner's dilemma strategies with fingerprints. 3:2613–2620 Vol. 3, Sept 2005.

[8] Daniel Ashlock and Eun-Youn Kim. Fingerprinting: Visualization and automatic analysis of prisoner's dilemma strategies. *IEEE Transactions on Evolutionary Computation*, 12(5):647–659, Oct 2008.

[9] Daniel Ashlock, Eun-Youn Kim, and Wendy Ashlock. Fingerprint analysis of the noisy prisoner's dilemma using a finite-state representation. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(2):154–167, June 2009.

[10] Daniel Ashlock, Eun-Youn Kim, and Wendy Ashlock. A fingerprint comparison of different prisoner's dilemma payoff matrices. pages 219–226, Aug 2010.

[11] Daniel Ashlock, Eun-Youn Kim, and N. Leahy. Understanding representational sensitivity in the iterated prisoner's dilemma with fingerprints. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 36(4):464–475, July 2006.

[12] Wendy Ashlock and Daniel Ashlock. Changes in prisoners dilemma strategies over evolutionary time with different population sizes. pages 297–304, 2006.

[13] Robert Axelrod. Effective choice in the prisoner's dilemma. *The Journal of Conflict Resolution*, 24(1):3–25, 1980.

[14] Robert Axelrod. More effective choice in the prisoner's dilemma. *The Journal of Conflict Resolution*, 24(3):379–403, 1980.

[15] Robert Axelrod. The evolution of strategies in the iterated prisoner's dilemma. *Genetic Algorithms and Simulated Annealing*, pages 32–41, 1987.

[16] Robert Axelrod. Launching the evolution of cooperation. *Journal of Theoretical Biology*, 299(Supplement C):21 – 24, 2012. Evolution of Cooperation.

[17] Robert Axelrod and Douglas Dion. The further evolution of cooperation. *Science*, 242(4884):1385–1390, 1988.

[18] Robert Axelrod and William D. Hamilton. The evolution of cooperation, 1981.

[19] Jeffrey S Banks and Rangarajan K Sundaram. Repeated games, finite automata, and complexity. *Games and Economic Behavior*, 2(2):97–117, 1990.

[20] Bruno Beaufils, Jean paul Delahaye, and Philippe Mathieu. Our meeting with gradual: A good strategy for the iterated prisoner's dilemma. 1997.

[21] R Boyd. Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *Journal of theoretical biology*, 136 1:47–56, 1989.

[22] R. Boyd and J. P. Lorberbaum. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, 327:58–59, 1987.

[23] Valerio Capraro, Jillian J Jordan, and David G Rand. Heuristics guide the implementation of social preferences in one-shot prisoner's dilemma experiments. *Scientific reports*, 4, 2014.

[24] Gary W. Evans and Charles M. Crumbaugh. Payment schedule, sequence of choice, and cooperation in the prisoner's dilemma game. *Psychonomic Science*, 5(2):87–88, Feb 1966.

[25] Merrill M. Flood. Some experimental games. *Management Science*, 5(1):5–26, 1958.

[26] Nelis Franken and Andries Petrus Engelbrecht. Particle swarm optimization approaches to coevolve strategies for the iterated prisoner's dilemma. *IEEE Transactions on Evolutionary Computation*, 9(6):562–579, 2005.

[27] Philip S. Gallo and Irina Avery Dale. Experimenter bias in the prisoner's dilemma game. *Psychonomic Science*, 13(6):340–340, Jun 1968.

[28] Marco Gaudesi, Elio Piccolo, Giovanni Squillero, and Alberto Tonda. Exploiting evolutionary modeling to prevail in iterated prisoners dilemma tournaments. *IEEE Transactions on Computational Intelligence and AI in Games*, 8(3):288–300, 2016.

[29] Carlos Gracia-Lázaro, José A Cuesta, Angel Sánchez, and Yamir Moreno. Human behavior in prisoner's dilemma experiments suppresses network reciprocity. *Scientific reports*, 2, 2012.

[30] Douglas R. Green. 'tit-for-tat' in cell biology. *Nature Reviews Molecular Cell Biology*, 12:73, 2011.

[31] H. C. J. Godfray. The evolution of forgiveness. *Nature*, 355:206–207, 1992.

[32] Marc Harper, Vincent Knight, Martin Jones, Georgios Koutsovoulos, Nikoleta E. Glynatsi, and Owen Campbell. Reinforcement learning produces dominant strategies for the iterated prisoner's dilemma. *CoRR*, abs/1707.06307, 2017.

[33] PG Harrald and DB Fogel. Evolving continuous behaviors in the iterated prisoner's dilemma. *Bio Systems*, 37(1-2):135145, 1996.

[34] Christian Hilbe, Martin A. Nowak, and Arne Traulsen. Adaptive dynamics of extortion and compliance. *PLOS ONE*, 8(11):1–9, 11 2013.

[35] R. Hoffmann and N. C. Waring. *Complexity Cost and Two Types of Noise in the Repeated Prisoner's Dilemma*, pages 619–623. Springer Vienna, Vienna, 1998.

[36] Vincent Knight, Owen Campbell, Marc Harper, Karol Langner, James Campbell, Thomas Campbell, Alex Carney, Martin J. Chorley, Cameron Davidson-Pilon, Kristian Glass, Tomás Ehrlich, Martin Jones, Georgios Koutsovoulos, Holly Tibble, Müller Jochen, Geraint Palmer, Paul Slavin, Timothy Standen, Luis Visintini, and Karl Molden. An open reproducible framework for the study of the iterated prisoner's dilemma. *CoRR*, abs/1604.00896, 2016.

[37] Tatjana Krama, Jolanta Vrublevska, Todd M. Freeberg, Cecilia Kullberg, Markus J. Rantala, and Indrikis Krams. You mob my owl, ill mob yours: birds play tit-for-tat game. *Scientific Reports*, 2(800):73, 2012.

[38] Christopher Lee, Marc Harper, and Dashiell Fryer. The art of war: Beyond memory-one strategies in population games. *PLOS ONE*, 10(3):1–16, 03 2015.

[39] Jiawei Li. How to design a strategy to win an ipd tournament. In *The Iterated Prisoners' Dilemma 20 Years On*, World Scientific Book Chapters, chapter 4, pages 89–104. World Scientific Publishing Co. Pte. Ltd., 04 2007.

[40] Jiawei Li, Philip Hingston, and Graham Kendall. Engineering design of strategies for winning iterated prisoner ' s dilemma competitions. 3(4):348–360, 2011.

[41] SIWEI LI. Strategies in the stochastic iterated prisoners dilemma. *REU Papers*, 2014.

[42] Kristian Lindgren and Mats G. Nordahl. Evolutionary dynamics of spatial games. *Phys. D*, 75(1-3):292–309, 1994.

[43] Daniel R. Lutzker. Sex role, cooperation and competition in a two-person, non-zero sum game. *Journal of Conflict Resolution*, 5(4):366–368, 1961.

[44] David Mack, Paula N. Auburn, and George P. Knight. Sex role identification and behavior in a reiterated prisoner's dilemma game. *Psychonomic Science*, 24(6):280–282, Jun 1971.

[45] R. M. May. More evolution of cooperation. *Nature*, 327:15–17, 1987.

[46] M. Milinski. Tit for tat in sticklebacks and the evolution of cooperation. *Nature*, 325:433–435, January 1987.

[47] John H. Miller. The coevolution of automata in the repeated prisoner's dilemma. *Journal of Economic Behavior and Organization*, 29(1):87 – 112, 1996.

[48] Per Molander. The optimal level of generosity in a selfish, uncertain environment. *The Journal of Conflict Resolution*, 29(4):611–618, 1985.

[49] M. A. Nowak and K. Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355:250–253, January 1992.

[50] Martin Nowak and Karl Sigmund. Game-dynamical aspects of the prisoner's dilemma. *Applied Mathematics and Computation*, 30(3):191 – 213, 1989.

[51] Martin Nowak and Karl Sigmund. The evolution of stochastic strategies in the prisoner's dilemma. *Acta Applicandae Mathematica*, 20(3):247–265, Sep 1990.

[52] May R. M. Nowak M. A. Evolutionary games and spatial chaos. *Letters to nature*, 359:826–829, 1992.

[53] W H Press and F J Dyson. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26):10409–10413, 2012.

[54] A. Rapoport and A.M. Chammah. *Prisoner's Dilemma: A Study in Conflict and Cooperation, by Anatol Rapoport and Albert M. Chammah, with the Collaboration of Carol J. Orwant.* University of Michigan Press, 1965.

[55] Amnon Rapoport, Darryl A. Seale, and Andrew M. Colman. Is tit-for-tat the answer? on the conclusions drawn from axelrod's tournaments. *PLOS ONE*, 10(7):1–11, 07 2015.

[56] A.J. Robson. Efficiency in evolutionary games: Darwin, nash and secret handshake. 1989.

[57] Ariel Rubinstein. Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory*, 39(1):83 – 96, 1986.

[58] John Sensenig, Thomas E. Reed, and Jerome S. Miller. Cooperation in the prisoner's dilemma as a function of interpersonal distance. *Psychonomic Science*, 26(2):105–106, Feb 1972.

[59] Wolfgang Slany and Wolfgang Kienreich. On some winning strategies for the iterated prisoner's dilemma or mr. nice guy and the cosa nostra. *CoRR*, abs/cs/0609017, 2006.

[60] Alexander J. Stewart and Joshua B. Plotkin. Extortion and cooperation in the prisoners dilemma. *Proceedings of the National Academy of Sciences*, 109(26):10134–10135, 2012.

[61] James T. Tedeschi, Douglas S. Hiester, Stuart Lesnick, and James P. Gahagan. Start effect and response bias in the prisoner's dilemma game. *Psychonomic Science*, 11(4):149–150, 1968.

[62] A. W. Tucker. The mathematics of tucker: A sampler. *The Two-Year College Mathematics Journal*, 14(3):228–232, 1983.

[63] Paul E. Turner and Lin Chao. Prisoner's dilemma in an rna virus. *Nature*, 398:441–443, 1999.

[64] E. Tzafestas. Toward adaptive cooperative behavior. 2:334–340, Sep 2000.

[65] Pieter van den Berg and Franz J Weissing. The importance of mechanisms for the evolution of cooperation. In *Proc. R. Soc. B*, volume 282, page 20151382. The Royal Society, 2015.

[66] J Von Neumann and O Morgenstern. Theory of games and economic behavior. *Princeton University Press*, page 625, 1944.

[67] G. S. Wilkinson. Reciprocal food sharing in the vampire bat. *Nature*, 308:181–184, 1984.

[68] Jianzhong Wu and Robert Axelrod. How to cope with noise in the iterated prisoner's dilemma. *Journal of Conflict Resolution*, 39(1):183–189, 1995.