A systematic literature review of the Prisoner's Dilemma; collaboration and influence.

Nikoleta E. Glynatsi, Vincent A. Knight

2016

Abstract

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 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. This paper aims to provide a detailed literature review of the field. This is achieved by partitioning the timeline into six different sections. Furthermore, a comprehensive data set of literature is analysed using network theoretic approaches in order to explore the influence and the collaborative behaviour of the field itself.

1 Introduction

To add a sentence about selfness and selfishness. There is a simple way of representing these behaviours/concepts. This is to use a particular two player non-cooperative game called the prisoner's dilemma, originally described in [34].

Each player has two choices, to either be selfness and cooperate or to act in a selfish manner and choose to defect. Each decision is made simultaneously and independently. The fitness 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 a mutual cooperation.

A player's payoffs are generally represented by (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 constrains (2) and (3).

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

$$T > R > P > S \tag{2}$$

$$2R > T + S \tag{3}$$

Another common representation of the payoff matrix is given by (4), where b is the benefit of the altruistic behaviour and c it's its cost (constrains (2) - (3) still hold).

$$\begin{pmatrix} b - c & c \\ b & 0 \end{pmatrix} \tag{4}$$

Constrains (2) - (3) and rational behaviour guarantee that it never benefits a player to cooperate. 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 and theoretical works have shown that cooperation can emerge once players interact repeatedly. Arguably, the most important of these works has been R. Axelrod's "The Evolution of Cooperation" [23]. In his book Axelrod reports on a series of computer tournaments he organised of a finite turns games of the iterated prisoner's dilemma. Participants had to choose between cooperation and defection again and again while having memory of their previous encounters. Academics from several fields were invited to design computer strategies to compete. The pioneering work of Axelrod 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 prisoner's dilemma has attracted much attention ever since the game's origins. In Section 3 a comprehensive data set of literature regarding the prisoner's dilemma, and collected from the following sources, is presented and analysed.

- arXiv [60]; 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 [5]; 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) [46]; 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 [38]; a British 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 [61]; a leading global scientific publisher of books and journals. It publishes close to 500 academic and professional society journals.

Reviewing the amount of published academic articles as well as measuring and exploring the collaborations within the field became able tasks using this data set of literature.

2 Timeline

In this section a large amount of literature regarding the prisoner's dilemma is reviewed. The review starts from the year the game was formulated and covers publications all the way to today.

2.1 Origins of the prisoner's dilemma

The origin of the prisoner's dilemma goes back to the 1950s in early experiments conducted at RAND [34] to test the applicability of games described in [93]. The game received it's name later the same year. According to [90], A. W. Tucker (the PhD supervisor of J. Nash [66]), in an attempt to delivery the game with a story during a talk he described the players as prisoners and the game has been known as the prisoner's dilemma ever since.

The early research on the prisoner's dilemma had been very constrained. The only source of experimental results was through groups of humans that simulated rounds of the games and human groups came with disadvantages. Human can behave very randomly and in several experiments both the size and the background of the individuals were different. Thus comparing results of two or more studies is tricky.

The main aim of these experiments had been to understand the conditions behind the emergence of cooperation. Conditions such as the gender [33, 56, 57] of individuals, the representation of the game [33], the distance between players [81], the initial effects [89] and whether the experimenter was biased [35] were being explored and still are today.

An early figure that sought out to understand the conditions under which altruist behaviour emerged was Prof A. Rapoport. A mathematical psychologist, whose work focused on how to promote international and national cooperation. Rapoport tried to conceptualize strategies that could promote international cooperation. In his teaching and research he used the prisoner's dilemma. Rapoport offered the field many insights, he is the creator of strategies such as Tit for Tat and Pavlov, which are going to be discussed in later parts of this paper.

Decades later the political scientist R. Axelrod introduced the pioneer computer tournaments that have largely replaced human subjects in the study of the iterated prisoner's dilemma ever since. In the next section these tournaments and several strategies that were design by researchers, such as Rapoport, are introduced.

2.2 Axelrod's tournaments and intelligent design of strategies

As discussed in Section 2.1, before 1980 a great deal of research was done in the field, however, as described in [21], the political scientist R. 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 sciences Axelrod created a framework for exploring these questions using computer tournaments.

Axelrod's tournaments made the study of cooperation of critical interest once again, academic articles were being published reproducing Axelord's work, accessing and further developing his results. Mainly performing new tournaments and designing new strategies. As described in [78], "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 [22] summarized a number of works that were immediately inspired from the "Evolution of Cooperation" and in 2012, [48] wrote a review on iterated prisoner's dilemma strategies and big competitions that had occurred since the originals.

In essence, Axelrod in the tournaments had asked researchers to design a strategy, set a number of rules, with the purpose of wining an iterated prisoner's dilemma tournament. These strategies were constructed by an intelligent cause and not an undirected process, and here there are refereed to as strategies of intelligent design. This section covers Axelrod's original tournaments as well as research that introduced new strategies of intelligent design.

The first reported computer tournament took place in 1980 [17]. Several scientists were invited to submit their strategies, written in the programming languages Fortran or Basic. There was a total of 13 submissions made by the following researchers,

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;

and a 13^{th} who remained anonymous.

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 run only once, each participant knew the exact length of the matches and had access to the full history of each match. Furthermore, Axelrod performed a preliminary tournament and the results were known to the participants. The payoff values used for equation (1) 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. The success of Tit for Tat came as a surprise. It was not only the simplest submitted strategy but it had also won the tournament even though it could never do better than any player it was interacting with.

In order to further test the results Axelrod performed a second tournament later in 1980 [18]. The results of the first tournament had been publicised and the second tournament received much more attention, with 62 entries made by the following people,

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.

The new participants knew the results of the previous tournament. The rules were similar with only a few alternations. 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' [22]) of the game ending on the next move. However, 5 different length matches were selected for each match 63, 77, 151, 308 and 401, such that the average length would be around 200 turns.

Several entries of the second tournament tended to be variants of Tit for Tat, such as **Tit for Two Tats** submitted by John Maynard Smith. Tit for Two Tats defects only when the opponent has defected twice in a row. Another well known entry was **Grudger**. Grudger was originally submitted by James W. Friedman. Grudger is a strategy that will cooperate as long as the opponent does not defect. The name Grudger was give to the strategy in [54]. Though the strategy goes by many names in the literature such as, Spite [26], Grim Trigger [24] and Grim [92].

Despite the larger size of the second tournament, none of the variants and new strategies managed to outperform the simple designed strategy. The winner was once again Tit for Tat. The conclusions made from the first two tournaments were that the strong performance of the strategy was due to:

- 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.
- As soon as the opponents identified that they were playing Tit for Tat, they would choose to cooperate for the rest of the game.

However, the success of Tit for Tat was not unquestionable. Several papers showed that stochastic uncertainty severely undercut the effectiveness of reciprocating strategies. Though such stochastic uncertainty are unlikely to occur in a computer tournament, but have to be expected in real life situations [62].

In [64] it was proved that in an environment where **noise** is introduced two strategies playing Tit for Tat receive the same average payoff as two Random players. Noise is a probability that a player's move will be flipped. In 1986, [30] ran a computer tournament with a 10 percent chance of noise and Tit for Tat finished sixth out of 21 strategies. Bendor in [47] performed tournament similar to Axelrod's with noise and a probability of 0.0067 of ending in the next turn. His results demonstrated the poor performance of Tit for Tat once again and showed that the highest ranked strategies were more generous ones. His top ranked strategy was called **Nice and Forgiving**. Nice and Forgiving, differs in significant ways from Tit for Tat. Initially, Nice's generosity takes the form of a benign indifference. It will continue to play cooperation as long as its rival's cooperation level exceeded 80%. Secondly, although it will retaliate if its rival's observed cooperation fell below 80, it is willing to revert to full cooperation before its partner does, so long as the partner satisfies a certain thresholds of acceptable behaviour.

Hammerstein [80], pointed out another weakness of Tit for Tat in noisy environments. 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. To overcome this error [82] introduced another more generous variant of Tit for Tat, **OmegaTFT**. They also altered their strategy so that it had the ability to to recognize and exploit the Random strategy in a way that after an opponent strategy crosses a certain randomness threshold they conclude that the opponent is a Random strategy and change the behaviour to act as a **Defector**, a strategy that always choose to defect.

A second type of stochastic uncertainty is misperception, where a player's action is made correctly but it's recorder incorrectly by the opponent. In 1986, [87] introduced a strategy called **Contrite Tit for Tat** 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 a defection. This hints, alongside more generous versions of Tit for Tat as discussed above, that the counter attack to stochastic uncertainties is a strategy's readiness to defect after a defection.

Another protagonist in the literature and better perform strategy than Tit for Tat came along in 1993. The strategy was **Pavlov** and though the name was formally given by Nowak [71] the strategy had been around since 1965 known as Simpleton [77], introduced by Rapoport himself. The strategy is based on the fundamental behavioural mechanism win-stay, lose-shift. It starts off with a cooperation and then repeats it's previous move only if it was awarder with a payoff of R or T. Pavlov is heavily studied in the literature and Similarly to Tit for Tat it's used in tournaments perform until today and has had many variants trying to build upon it's success. **PavlovD**, just a Pavlov which starts the game with a defection and **Adaptive Pavlov**. Adaptive Pavlov tried to classify it's opponents as as one of the following strategies, **Cooperator**, Defector, Pavlov, Random or Pavlov and chooses a strategy that maximise it's payoff against the now 'known' opponent. Cooperator is a deterministic strategy that conditionally cooperates.

Several researchers, and this will discussed in later sections as well, argued with Axelrod's result on simplicity. The advantages of complexity were shown by [26] in 1997 where they introduced another well known strategy **Gradual**. 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. In a tournament of 12 strategies [26], Gradual had managed to outperform strategies such as Tit for Tat and Pavlov. Gradual was surpassed by yet another complex and intelligent designed strategy **Adaptive Tit for Tat**. The authors of [91] conducted the exact same tournament as [26] with now 13 strategies and their strategy ranked first.

Another interesting research on teams of strategies [28, 29, 79]. These strategies are of intelligent design and they have been programmed with a recognition mechanism by default. Once the strategies recognise one another, one would act as leader and the other as a follower. The follower then plays as a Cooperator, cooperates unconditionally and the leader would play as a Defector gaining the hghest achievable score. Followers would behave as Defector towards other strategies to defect their score and help the leader. In [79], a team for the University of Southampton used teams and recognition patterns and managed to win the 2004 Anniversary IPD Tournament.

2.3 Evolutionary Dynamics

Following Axelrod's tournaments it was proven that direct reciprocity can make cooperation successful is an round robin setting. A long standing question had been understand the conditions required for the emergence and maintenance of cooperation in evolving populations. The complex nature of the iterated prisoner's dilemma strategies makes their evolutionary stability more complex to study and though the question "when does cooperation emergences in an evolutionary dynamic setting" still remains open several insights have been published over the years. In this section several remarks that have been made on the evolutionary dynamics are discussed.

In the later sections of [18], Axelrod discusses about an ecological tournament he performed using the 62 strategy of the second tournament An ecological approach is a simulation of theoretical future rounds of the game where strategies that do better are more likely to be included in future rounds than others. The simulation of the process, as described in [18], is straightforward. Let us consider an example. Let the four strategies Tit for Tat, Tit for Two Tat, Cooperator and Defector compete in an ecological tournament. The expected payoff matrix, when these four strategies interact, is give by,

$$\begin{bmatrix} 3.0, & 3.0, & 3.0, & 0.99 \\ 3.0, & 3.0, & 3.0, & 0.99 \\ 3.0, & 3.0, & 3.0, & 0.0 \\ 1.02, & 1.039, & 5.0, & 1.0 \end{bmatrix}$$

Starting with proportions of each type in a given generation, their proportions for the next generation needs to be calculated. This is achieved by calculating the weighted average of the scores of a given strategy with all other players.

- The weights are the numbers of the other strategies which exist in the current generation.
- The numbers of a given strategy in the next generation is then taken to be proportional to the product of its numbers in the current generation and its score in the current generation.

The process is then repeated for a given number of future tournaments. Figure 1 illustrates a simulation of our ecological tournament, as shown strategies that cooperate quickly kill off the Defector.

A similar result was presented by Axelrod. In his ecological tournament cooperative strategies managed to take over the population over time. On the other hand exploitative strategies started to die off as weaker strategies were becoming extinct. In other words they were dying because there was fewer and fewer prey for them to exploit. In 1981, Axelrod also studied the prisoner's dilemma in an evolutionary context based on the evolutionary approaches of John Maynard Smith [59, 83, 84]. Smith is a well know evolutionary biologist as well as an attendant of Axelord's second tournament. John Maynard Smith alongside George Price are considered fundamental figures of evolutionary game theory. In [59] they introduced the definition of an evolutionary stable strategy (ESS).

Imagine a population made up of individuals where everyone follows the same strategy B and a single individual adopts a mutant strategy A. Strategy A is said to invade strategy B if the payoff of A against B is greater than the expected payoff received by B against itself. Since strategy B is in a population that interacts only with itself, the concept of invasion is equivalent to a single mutant being able to outperform the average population. Thus for a strategy to be ESS it must be able to resists any invasion.

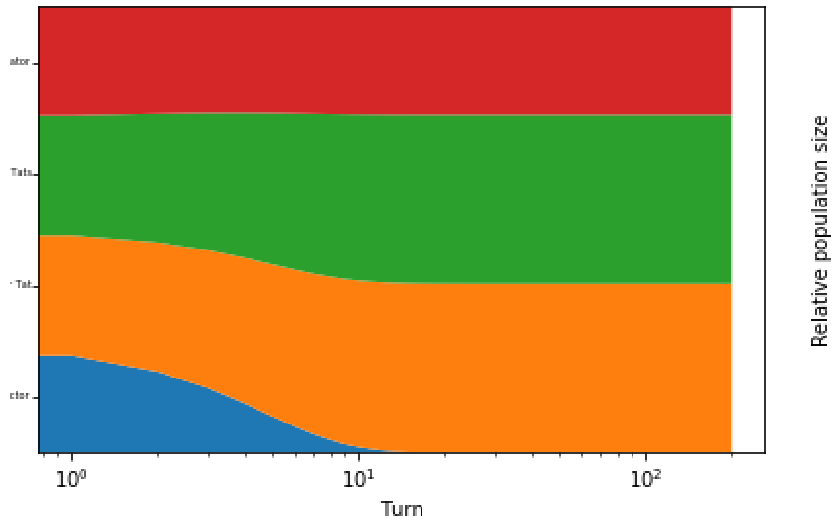


Figure 1: Results on an ecological tournament with Tit for Tat, Tit for Two Tats, Cooperator and Defector.

The work described in [19], studied the evolutionary stability of Tit for Tat and although the strategy was likely to take over the population, its stability was conditional on the importance of the future of the game. This is represented by a discounting factor denoted as w. Axelrod showed that if w was sufficiently large, Tit for Tat could resist invasion by any other strategy. Moreover, he showed how a small cluster of Tit for Tat players could invade a extortionate environment. Alongside the biologist William Donald Hamilton they wrote about the biological applications of the evolutionary dynamics of the iterated prisoner's dilemma [23] and won the Newcomb-Cleveland prize of the American Association for the Advancement of Science. Arguing with Axelrod's results. In [27] Boyd and Lorderbaum show that if w, the importance of the future of the game, is large enough then no deterministic strategy is ESS because it can always be invaded by any pair of other strategies. This was also independently proven by [76]. Furthermore, Boyd Boyd showed that introducing noise into the evolutionary IPD allows such strategies to be evolutionary stable and that for given combination of strategies cooperation can start without a cluster of cooperative strategies. Arguing with this remark with Axelrod's results. In [69], Nowak and Sigmund studied the dynamics of the evolutionary iterated prisoner's dilemma with a spectrum of stochastic strategies that only remember their opponent's last move, not their own. They found that there can be multiple fixed points that there can be an evolutionary stable coexistence among multiple such strategies.

In 1992, [72] explored how local interaction alone can facilitate population wide cooperation in a oneshot Prisoner's Dilemma. 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, where each node represents a player and the edges denote whether two players will interact. Thus each cell of the lattice is occupied by a Cooperator or a Defector.

- 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 the immediate neighbours.

This topology is refereed to as spatial topology. The population dynamics of these experiments were studied as a function of the temptation (T) payoff. More specifically the following payoff matrix was used, which is equivalent to equation (1):

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} \tag{5}$$

where ($b > 1$).

It was shown that as long as small clusters of cooperators form and they can benefit from interactions with their own kind while avoiding interactions with defectors, global cooperation will continue, Also, cooperation cannot survive under a more mundane condition that the benefit-to-cost ratio is lower relative to the number of neighbors in such a population purely composed of unconditional cooperators and defectors In [58] showed that cooperation is more likely to emerge in a small-world topology. Another topic worth mentioning is that of coevolution on graphs. In [97] studied graphs were a probability of rewiring ones connections was in place, however, the rewire could be with any given node in the graphs and not just with imitate neighbours. Perc etc all showed that "making of new friends" may be an important activity for the successful evolution of cooperation, but also that partners must be selected carefully and one should keep their number limited.

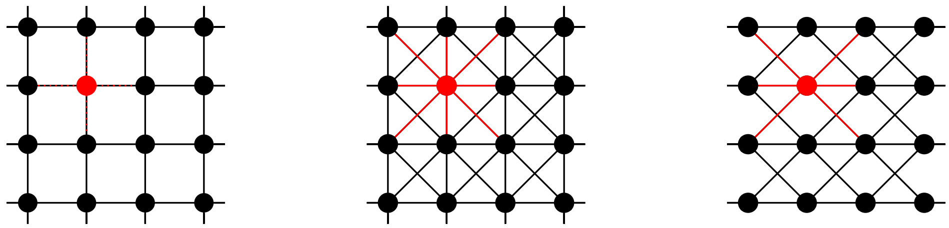


Figure 2: Spatial neighbourhoods

Recent studies have scrutinized another mechanism through which altruists interact with each other so preferentially that they can receive a disproportionate share of the benefit of altruism. 'Cues' such as reputation (Nowak and Sigmund 1998; Suzuki and Akiyama 2005; Janssen 2006) and communication tokens (Miller et al. 2002) can increase the likelihood of assortative interactions among cooperative agents. These models are concerned with partner identification, but agents should be able to have relatively high cognitive capacities. 2.6 Another related approach highlighting tag-based partner identification has contributed to a better understanding of the evolution of cooperation among minimally cognitive agents (Riolo 1997; Riolo et al. 2001; Hales 2000; Hales 2004a; Hales 2004b; Hales and Edmonds 2003; Edmonds and Hales 2005; Choi et al. 2006). Human agents have observable markers called 'tags' as the phenotype of "memes" (Dawkins 1976) to form "memetic kin" (Heylighen and Campbell 1995). In other words, human agents, albeit genetically unrelated, have perceived social distances to others reading tags to distinguish 'us' from 'them.'

2.4 Structured strategies and training

In Section 2.3 several evolutionary dynamic approaches used in the iterated prisoner's dilemma research were covered. All of these approaches have a limitation, and that is their inability to develop new strategies. In evolutionary settings strategies can learn to adapt their actions over time based upon what has been effective and what has not. In essence reinforcement learning techniques can be used to train strategies for playing the iterated prisoner's dilemma. A strategy must be represented in a generic strategy archetype so it can be trained. Strategies that are discovered via strategy archetypes are refereed to as structured strategies. They are the opposite to that of intelligent designed discussed in Section 2.2. In this section papers that proposed new structured strategies as well as papers that explore new training algorithms are covered.

In Axelrod, [20] having realised the limitations of his evolutionary work, decided to use reinforcement learning to demonstrated how to fine tune the responses of an IPD strategy and obtain a trained strategy. can be trained. The algorithm used in [20] is called the genetic algorithm and the archetype is refereed to as a lookup table. Axelrod decided to considered deterministic strategies that took into account the last 3 turns of the game. 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. A more generic lookup definition is a deterministic strategy which take into account the last m actions.

In 1989 [69], Nowak and Sigmund proposed a structure for studying sophisticated strategies instead of deterministic ones. Moreover, they studied a set of very simple strategies that remember only the previous turn, and moreover, only record the move of the opponent. They 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. Using the above notation a strategy can now be defined by a triple. For example,

- Defector: (0, 0, 0)
- Cooperator: (1, 1, 1)
- Tit for Tat: (1, 1, 0)
- Pavlov: (0, 1, 1, 0)

In 1992 [68] the same authors reduced the spectrum of strategies to 99 reactive. They were placed in the same population and the reactive strategy which managed to take over was called **Generous Tit for Tat** given by the triplet $(1, 0, \frac{2}{3})$.

Moreover in 1990 [70] they expanded their structure and gave a formal definition of a memory one strategy. Memory one strategies consider the entire history of the previous turn to make a decision (thus reactive strategies are a subset of memory one). If only a single turn of the game is taken into account and depending on the simultaneous moves of two players there are only four possible states that players could possibly be in. These are CC, CD, DC and DD. A memory one strategy is denoted by the probabilities of cooperating after each of these states, $p = (p_1, p_2, p_3, p_4) \in \mathbb{R}^4_{[0,1]}$. A match between two memory one players p and q can be modelled as a stochastic process, where the players move from state to state. More specifically, it can be modelled by the use of a Markov chain [36], which is described by a matrix M.

$$M = \begin{bmatrix} p_1q_1 & p_1(-q_1+1) & q_1(-p_1+1) & (-p_1+1)(-q_1+1) \\ p_2q_3 & p_2(-q_3+1) & q_3(-p_2+1) & (-p_2+1)(-q_3+1) \\ p_3q_2 & p_3(-q_2+1) & q_2(-p_3+1) & (-p_3+1)(-q_2+1) \\ p_4q_4 & p_4(-q_4+1) & q_4(-p_4+1) & (-p_4+1)(-q_4+1) \end{bmatrix} \tag{6}$$

The players are assumed to move from each state until the system reaches a state steady, let the steady states vector be denoted as $\bar{v}$. The utility of a player can be given by multiplying the steady states of M by the payoffs of equation (1). Thus [70] offered a mathematical framework to calculate the utility of two players without actually simulating the game. The payoff of a player p can be obtained by,

$$s_p = \bar{v} \times \begin{pmatrix} R \\ S \\ T \\ P \end{pmatrix}$$

The family of memory on strategies have been proven rather useful in the terms of exploring strategies. The most famous work of memory one strategies is that of Press and Dyson [75]. In 2012, [75] presented a new set of strategies called **zero determinant (ZD)**. The ZD strategies are memory one strategies that manage to force a linear relationship between their score and that of the opponent. The payoffs of players p and q are denoted as:

$$\begin{aligned} s_p &= vS_p \\ s_q &= vS_q \end{aligned}$$

where v is a vector of the steady states of matrix M and S_p, S_q are the equivalent payoff values of the players for each state CC, CD, DC, DD. Using linear algebra, Press and Dyson showed that the dot product of the stationary distribution of v with any vector f can be expressed as a 4×4 determinant. In which one column is f, one column is entirely under the control of player p and another column is entirely under the control of player q. This meant that either p or q could independently force the dot product of v with some other chosen vector f to be zero by choosing their strategy so as to make the column they control be proportional to f. In particular, by $f = \alpha S_p + \beta S_q + \gamma$, any player can force a given linear relation to hold between the long-run scores of both players. Press and Dyson's suggested that these extortionate strategies are the dominant family of strategies and moreover memory would not benefit any strategy.

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" [85] and as stated in [44] the American Mathematical Society's news section said that "the world of game theory is currently on fire". In [7, **?**, 43, 44, 85, 52, 53] they question the effectiveness of ZD strategies. In [85], they revealed a more generous set of ZDs the **Generous ZD**, [7] showed that ZD strategies were not evolutionarily stable. In [53], 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.

Memory one structure only allows simple strategies to be consider, whether they are deterministic or sophisticated, they only can take into account a single history. Other structures allow for both simple and complex strategies to be explored, such as neural networks and finite automata. In 1996 Harrald and Fogel [40], considered a neural network that used memory with length of 3 and the actions were encoded as continuous values in $[-1, 1]$, where 1 meant complete cooperation. The input nodes represented 3 previous steps of the player and the opponent and there was a single hidden layer of N fully connected nodes and an output node that produced values from the range $[-1, 1]$. Same year [63] consider finite state automata as a structure. The specific type of finite automata that were used were Moore machines [65]. Finite state machine consist of a set of internal states. One of these states is the initial state of the machine. A machine also consists of transitions 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.

Miller used the genetic algorithm to train finite state machines in environments with noise. His results showed that even a small difference in noise (from 1% to 3%) significantly changed the characteristics of the evolving strategies. Three machines described in his paper are the following:

- **Punish Twice**: A strategy that punishes defection with 2 defections.
- **Punish Once for Two Tats**: A strategy which will defect only if the opponent has defected twice in a row.
- **Punish Twice and Wait**: A variant of Punish Twice which will answer defection with 2 defections and will cooperate if an only if the opponent cooperated.

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 [39] and a paper describing it and it's capabilities was published in 2016 [51].

The two paper using the Axelrod project [39, 52] present several powerful strategies created using reinforcement learning techniques. Reinforcement learning refers to a collection of algorithms that trains a model by exploring a space of actions and evaluating consequences of those actions. In these papers the authors used genetic algorithms and particle swarm optimisation algorithms [86]. A number of strategy representations, referred as archetypes, were used to train strategies. These included, lookerup tables, finite state machines, artificial neural networks [96] and hidden Markov models [32].

Hidden Markov models, are 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.

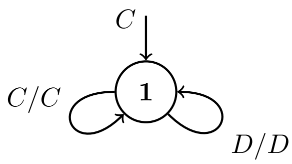
Additionally a variant of a look up table is also presented called the lookerup archetype. The lookerup archetype responses based on the opponent's first n_1 moves, the opponent's last m_1 moves, and the players last m_2 moves. Taking into account the initial move of the opponent can give many insights. 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.

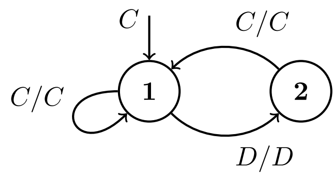
The training of these archetypes was done in two following settings:

- 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.
- A round robin tournament.

The result of [52] show that the trained strategies evolve an ability to recognise themselves by using a handshake. This characteristic of the strategies was an important one because in a Moran process this recognition mechanism allowed these strategies to resist invasion. In [39], they performed a a standard tournament with 200 turns but also a noisy tournament. For the standard tournament the newly introduced trained strategies outperform the designed ones. In the case of noise there is one particular strategy that has not seen much attention in the literature called "Desired Belief Strategy" [16]. These experiments are, to the authors knowledge, the biggest ones done in the field in terms of different strategies.

Training can return a series of strategies and different archetypes make Differentiating between strategies is not an easy task. 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 Figures ref and ref are both finite automata representation of Tit for Tat.





In order to distinguish the strategies and assuring that they are indeed different [11] introduced a method called fingerprinting. The method of fingerprinting is a technique for generating a functional signature for a strategy [12]. 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 [11] 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 3 an example of Pavlov's fingerprint is given. Fingerprinting has been studied in depth in [12, 13, 14, 15].

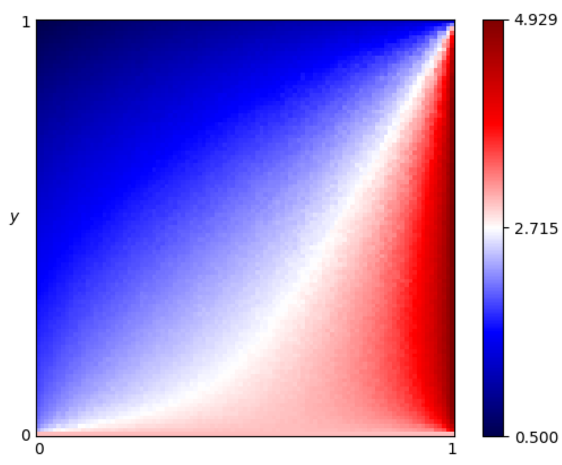


Figure 3: Pavlov fingerprinting with Tit for Tat used as the probe strategy. Figure was generated using [6].

2.5 Software

The research of the iterated prisoner's dilemma heavily relays on software. This is to be expected as the pioneer computer tournaments have become the main mean of simulating the interactions in an iterated prisoner's dilemma game. Many academics fields suffer from the lack of source code availability and the prisoner's dilemma is not an exception. Though several of the tournaments that discussed so far were generated using computer code not all of the source code was made available by the authors. 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 [1]).

Several projects, however, are open, available and have been used as research tools or educational platforms over the years. Two research tools are briefly mentioned here [4, 6] and two educational tools [2, 3]. Both [4, 6] are open source projects used as research tools. PRISON is written in the programming language Java and preliminary version was launched on 1998. It was used by it's authors in several publications, such as [26] which introduced Gradual and [25]. The project includes a good number of strategies from the literature but unfortunately the last update of the project dates back in 2004. Axelrod-Python is a package used by several papers covered in Section 2.4. It is written in the programming language Python following best practice approaches and contains the larger to date data set of strategies. The strategy list of the project has been cited by publications [9, 42, 67] and the package has not been used only by the contributors for academic research but for several works such as: [37, 94].

The 'Game of Trust' [2] is an on-line, graphical user interface educational platform for learning the basics of game theory, the iterated prisoner's dilemma and the notion of strategies. It attracted a lot of attention due to being "well-presented with scribble-y hand drawn characters" [45] and "a whole heap of fun" [50]. Finally [3] is a personal project written in PHP. It's graphical user interface platform that offers a big collection of strategies and allows the user to try several matches and tournament configurations.

2.6 Conclusion and Contemporary period

This section of the paper has focused on reviewing articles that have been published on the prisoner's dilemma. This review has partitioned the literature into five different sections focusing on different aspects of the research. Section 2.1 covered the early years of research. This was when scientists mainly grouped their students into pairs and asked them to simulate turns of the game. An early figure discussed in the section was Rapoport. A brilliant scientist that sought out the answer behind the emergence of cooperation and introduced to the literature a number of successful strategies.

Following the early years the pioneer tournaments of Axelord were introduced in Section 2.2. Axelrod's work offered the field an exceptional agent based game theoretic framework to study the iterated prisoner's dilemma. His original work asked researcher to develop their own intelligent design of strategies, and the winning strategy of both his tournaments was the Tit for Tat. The strategy however came with limitations which were explored by other researchers and new strategies of intelligent design were introduced in order to surpass Tit for Tat with some exceptional 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 dynamics methods were applied to many works in the fields, and factors under which cooperation emerges were explored, as described in Section 2.3. This was not done only for unstructured populations, where all strategies in the population interacted with each other, but also in population where interactions were limited to only the closes strategies. In structured setting it was proven that even in the simplest game of Cs and Ds cooperation can indeed emerge.

Evolutionary approaches can offer many insights in the study of the prisoner's dilemma. In evolutionary settings strategies can learn to adapt and take over population by adjusting their actions; this has been referenced to as training. Such algorithms can be applied so that evolutionary robust strategies can emerge. Algorithms and structured used to train strategies in the literature were covered in Section 2.4. Several strategies can emerge from such processes, and to be able to Differentiate between strategies fingerprinting was introduced. The research of dominance and cooperation has been going on since the 1950s, and several computer software have been developed along the way. Few have been briefly discussed in Section 2.5.

A large scale of articles has been covered in each of the corespomding sections of the section, not all in the next section. in recent years the study of the iterated prisoners' dilemma is still active and papers are still being published. This is going to be verified in the following section were a data set containing a number of publications is analysed. The iterated prisoner's dilemma now serves as a model in a wide range of applications, for example in medicine and the study of cancer cells [10, 49], as well as in social situations and how they can be driven by rewards [31]. A lot of work is still being published on structured populations trying to shed more light into evolutionarily dynamics on graphs [8, 41, 55]. The study of the prisoner's dilemma is still an ongoing field of pioneer and innovating research, where new variants and new structures of strategies are continuously being explored [73].

3 Analysing a large corpus of articles

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] Plos public library of science.

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

[7] Christoph Adami and Arend Hintze. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications*, 4:2193, 2013.

[8] Benjamin Allen, Gabor Lippner, Yu-Ting Chen, Babak Fotouhi, Naghmeh Momeni, Shing-Tung Yau, and Martin A Nowak. Evolutionary dynamics on any population structure. *Nature*, 544(7649):227, 2017.

[9] Nicolas Anastassacos and Mirco Musolesi. Learning through probing: a decentralized reinforcement learning architecture for social dilemmas. *arXiv preprint arXiv:1809.10007*, 2018.

[10] Marco Archetti and Kenneth J Pienta. Cooperation among cancer cells: applying game theory to cancer. *Nature Reviews Cancer*, page 1, 2018.

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

[12] 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.

[13] 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.

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

[15] 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.

[16] T. C Au and Dana S. Nau. Accident or intention: That is the question (in the noisy iterated prisoner's dilemma). pages 561 – 568, 2006/// 2006.

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

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

[19] Robert Axelrod. The emergence of cooperation among egoists. *American political science review*, 75(2):306–318, 1981.

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

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

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

[23] Robert Axelrod and William D. Hamilton. The evolution of cooperation, 1984.

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

[25] Bruno Beaufils, Jean-Paul Delahaye, and Philippe Mathieu. Complete Classes of Strategies for the Classical Iterated Prisoner's Dilemma. In V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben, editors, *Evolutionnary Programming VII (EP'7)*, volume 1447 of *Lecture Notes in Computer Science*, pages 33–41, undef, France, 1998. Springer.

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

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

[28] J.P. Delahaye. L'altruisme perfectionn. *Pour la Science (French Edition of Scientific American)*, 187:102–107, 1993.

[29] J.P. Delahaye. Logique, informatique et paradoxes, 1995.

[30] Christian Donninger. *Is it Always Efficient to be Nice? A Computer Simulation of Axelrod's Computer Tournament*, pages 123–134. Physica-Verlag HD, Heidelberg, 1986.

[31] Slimane Dridi and Erol Akay. Learning to cooperate: The evolution of social rewards in repeated interactions. *The American Naturalist*, 191(1):58–73, 2018. PMID: 29244562.

[32] Sean R Eddy. Hidden markov models. *Current opinion in structural biology*, 6(3):361–365, 1996.

[33] 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.

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

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

[36] Dani Gamerman and Hedibert F Lopes. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference.* Chapman and Hall/CRC, 2006.

[37] Aaron Goodman. Learning implicit communication strategies for the purpose of illicit collusion. *arXiv preprint arXiv:1802.06036*, 2018.

[38] Nature Publishing Group. Nature, 1869.

[39] 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.

[40] Paul G. Harrald and David B. Fogel. Evolving continuous behaviors in the iterated prisoner's dilemma. *Biosystems*, 37(1):135 – 145, 1996.

[41] David Hathcock and Steven H Strogatz. Fitness dependence of the fixation-time distribution for evolutionary dynamics on graphs. *arXiv preprint arXiv:1812.05652*, 2018.

[42] Victoria Hayes. *The Evolution of Cooperation: A Recreation of Axelrod's Computer Tournament.* The University of North Carolina at Greensboro, 2017.

[43] Christian Hilbe, Martin A Nowak, and Karl Sigmund. Evolution of extortion in iterated prisoners dilemma games. *Proceedings of the National Academy of Sciences*, page 201214834, 2013.

[44] Christian Hilbe, Arne Traulsen, and Karl Sigmund. Partners or rivals? strategies for the iterated prisoner's dilemma. *Games and Economic Behavior*, 92:41 – 52, 2015.

[45] Samuel Horti. The evolution of trust is a cute explain-o-game about cooperation. `http://ncase.me/trust/`.

[46] IEEE. Ieee xplore digital library.

[47] Bendor Jonathan, Kramer Roderick M., and Stout Suzanne.

[48] M. Jurii, D. Kermek, and M. Konecki. A review of iterated prisoner's dilemma strategies. pages 1093–1097, May 2012.

[49] Artem Kaznatcheev, Jeffrey Peacock, David Basanta, Andriy Marusyk, and Jacob G. Scott. Fibroblasts and alectinib switch the evolutionary games that non-small cell lung cancer plays. *bioRxiv*, 2017.

[50] Nathan Kinch. Game theory and the evolution of trust. `https://medium.com/greater-than-experience-design/game-theory-and-the-evolution-of-trust-6da95b33407a`.

[51] 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.

[52] Vincent A. Knight, Marc Harper, Nikoleta E. Glynatsi, and Owen Campbell. Evolution reinforces cooperation with the emergence of self-recognition mechanisms: an empirical study of the moran process for the iterated prisoner's dilemma. *CoRR*, abs/1707.06920, 2017.

[53] 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.

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

[55] Xu-Sheng Liu, Zhi-Xi Wu, Michael Z. Q. Chen, and Jian-Yue Guan. Evolutionary fate of memory-one strategies in repeated prisoner's dilemma game in structured populations. *The European Physical Journal B*, 90(7):138, Jul 2017.

[56] 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.

[57] 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.

[58] Naoki Masuda and Kazuyuki Aihara. Spatial prisoner's dilemma optimally played in small-world networks. *Physics Letters A*, 313(1):55 – 61, 2003.

[59] J. Maynard Smith and G.R. Price. The logic of animal conflict. *Nature*, 246(5427):15–18, 1973.

[60] Gerry McKiernan. arxiv. org: the los alamos national laboratory e-print server. *International Journal on Grey Literature*, 1(3):127–138, 2000.

[61] Mannheim Media. Springer publishing, 1950.

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

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

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

[65] Edward F Moore. Gedanken-experiments on sequential machines. *Automata studies*, 34:129–153, 1956.

[66] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[67] Shai Neumann, Suraj Sood, Markus Hollander, Freda Wan, Alexis-Walid Ahmed, and Monte Hancock. Using bots in strategizing group compositions to improve decision–making processes. In Dylan D. Schmorrow and Cali M. Fidopiastis, editors, *Augmented Cognition: Users and Contexts*, pages 305–325, Cham, 2018. Springer International Publishing.

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

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

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

[71] Martin Nowak and Karl Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(6432):56–58, 1993.

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

[73] Hisashi Ohtsuki. Evolutionary dynamics of coordinated cooperation. *Frontiers in Ecology and Evolution*, 6:62, 2018.

[74] Hisashi Ohtsuki and Martin A Nowak. The replicator equation on graphs. *Journal of theoretical biology*, 243(1):86–97, 2006.

[75] 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.

[76] P. Pudaite. On the initial viability of cooperation. *Merriam Laboratory for Analytic ARTICLES 1389 Political Research*, 1985.

[77] 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.

[78] 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.

[79] A. Rogers, A Rogers, RK Dash, SD Ramchurn, P Vytelingum, and NR Jennings. Coordinating team players within a noisy iterated prisoners dilemma tournament. *Theoretical computer science.*, 377(1-3):243–259, 2007.

[80] Reinhard Selten and Peter Hammerstein. Gaps in harley's argument on evolutionarily stable learning rules and in the logic of tit for tat. *Behavioral and Brain Sciences*, 7(1):115116, 1984.

[81] 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.

[82] 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.

[83] J. Maynard Smith. The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology*, 47(1):209 – 221, 1974.

[84] J. Maynard Smith. Game theory and the evolution of behaviour. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 205(1161):475–488, 1979.

[85] 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.

[86] Ponnuthurai N Suganthan. Particle swarm optimiser with neighbourhood operator. In *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, volume 3, pages 1958–1962. IEEE, 1999.

[87] Robert Sugden et al. *The economics of rights, co-operation and welfare*. Springer, 2004.

[88] György Szabó and Gabor Fath. Evolutionary games on graphs. *Physics reports*, 446(4-6):97–216, 2007.

[89] 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.

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

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

[92] 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.

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

[94] Shiheng Wang and Fangzhen Lin. Invincible strategies of iterated prisoner's dilemma. *arXiv preprint arXiv:1712.06488*, 2017.

[95] Duncan J Watts and Steven H Strogatz. Collective dynamics of small-worldnetworks. *nature*, 393(6684):440, 1998.

[96] B Yegnanarayana. *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.

[97] Chunyan Zhang, Jianlei Zhang, Guangming Xie, Long Wang, and Matja Perc. Evolution of interactions and cooperation in the spatial prisoner's dilemma game. *PLOS ONE*, 6(10):1–7, 10 2011.