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 ever 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 1980's after the publication of the "The Evolution of Cooperation" and has been a topic of pioneering research ever since. In this work we aim to provide a chronological literature review of the field. This is achieved by partitioning the timeline into five different time periods. Furthermore, a comprehensive data set of literature is analysed using network theoretic approaches in order to explore the collaborative behaviour and identify the influencers of the field.

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 [30].

Each player has two choices, to either be selfness and cooperate or to act in a selfish manner and chose 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 (6). Both players receive a reward for mutual cooperation, R, and a payoff P for mutual defection. A player that defected while the other cooperates receives a payoff of T, whereas the cooperator receives S. The dilemma exists due to constrains (2) and (3).

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

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

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

Constrains (2), (3) and rational behaviour guarantee that it never benefits a player to cooperate. It can been shown mathematically that defecting is the dominant strategy for the one shot prisoner's dilemma. 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 for more than one time. The most important of these works has been R. Axelrod's "The Evolution of Cooperation" [20]. 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. This is shown in Figure 1, which illustrates the number of publications on the prisoner's dilemma per year from the following sources:

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

Each point of Figure 1 marks the starting year of a time period. Each of these time periods is reviewed and presented in Section 2, as subsections of an extensive literature review. Furthermore, in Section ?? a comprehensive data set of literature regarding the prisoner's dilemma will be presented and analysed. This allow us to review the amount of published academic articles as well as measure and explore the collaborations within the field.

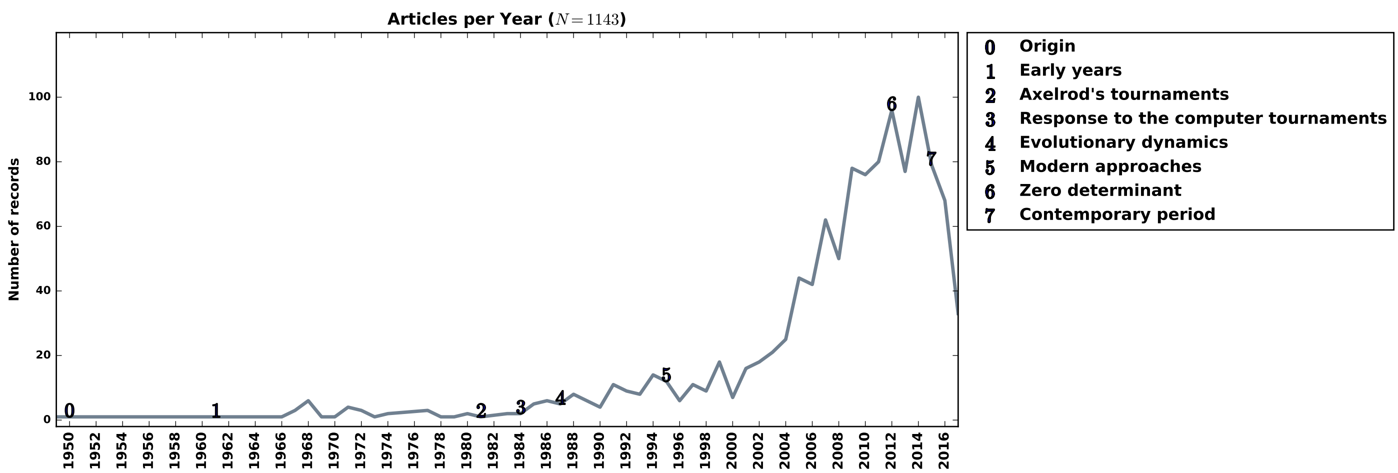


Figure 1: A timeline of the prisoner's dilemma research.

2 Timeline

In this section we review a large amount of literature regarding the prisoner's dilemma. We start from the year the game was formulated all the way to today.

2.1 Original research (1961-1972)

The origin of the prisoner's dilemma goes back to the 1950s in early experiments conducted at the RAND [30] to test the applicability of games described in [81]. The game received it's name later the same year. According to [78], A. W. Tucker (the PhD supervisor of J. Nash [54]), 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 study of the prisoner's dilemma has attracted people from various fields across the years. An early figure within the field of game theory is Prof A. Rapoport, a mathematical psychologist, whose work focused on how to promote international and national cooperation. Rapoport sought to conceptualize strategies that could promote international cooperation. It should come to no surprise that in his teaching and research he used the prisoner's dilemma. In [66] Rapoport conducted experiments using a group of humans to simulate rounds of the iterated prisoner's dilemma. He sought out to understand the conditions under which altruist behaviour emerged using human groups, and he was not the only one in his time. Conditions explored were the gender [29, 45, 46] of individuals, the representation of the game [29], the distance between players [69], the initial effects [77] and whether the experimenter was biased [31].

Several results argued about the importance of these factors. The early experiments were very constrained. The only source of simulation was through experimental groups and those came with disadvantages. In the next section we will introduce the evolutionary computer tournaments of Axelrod and in the next few sections it will be come evident how the field has not stood still ever since.

2.2 Axelrod's Tournaments (1981-1984)

Before the 1980s a great deal of research was done in the field, as discussed in Section 2.1. However, as described in his article [18], 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. This section is dedicated to a series of computer tournaments he performed in the early 1980s.

The first computer tournament was performed in 1980 [14]. 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;
13. Name not given.

Each competed in a 200 turn match against all 12 opponents, itself and a player that played randomly. This type of tournament is referred to as a round robin and corresponds to a complete graph from a topological point of view. 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 for equation (6) 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 the player it was interacting with.

In order to further test the results Axelrod performed a second tournament [15] later in 1980. 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 one exception; the number of turns was not specified instead a fixed probability (refereed to as 'shadow of the future' [19]) of the game ending on the next move was used. The fixed probability was chosen to be 0.0036 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. Each of the five matches had a length of 63, 77, 151, 308 and 401.

Several entries 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. However none of the variants managed to outperform the pure version and 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 Axelrod wanted to further test the robustness of the strategy. In the later sections of [15], he discusses about an ecological tournament he performed using the 62 strategy of the second tournament. The 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 [15], 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. Cooperator and Defector are two deterministic strategies that will always cooperate and defect equivalently. 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 2 illustrates a simulation of our hypothetical 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 started 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.

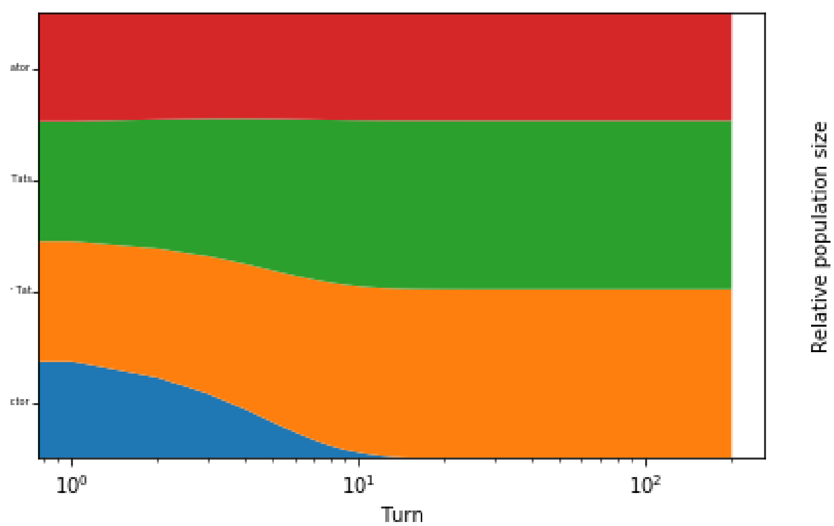


Figure 2: Results on an ecological tournament with Tit for Tat, Tit for Two Tats, Cooperator and Defector.

In 1981, Axelrod also studied the prisoner's dilemma in an evolutionary context based on the evolutionary approaches of John Maynard Smith [49, 71, 72]. John 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 [49] 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.

The work described in [16], 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. In [20], Axelrod decided to work on the biological applications of his evolutionary results in collaboration with the biologist William Donald Hamilton. According to Richard Dawkins he was the one to introduce Axelrod to Hamilton's work. Their collaboration [20] won the pair the Newcomb-Cleveland prize of the American Association for the Advancement of Science.

Axelrod's work was evolutionary and offered many insights to the field. Several other strategies, apart from Tit for Tat, are still being used in research today. Such as Tit for Two Tats and **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 [43]. Though the strategy goes by many names in the literature such as, Spite [23], Grim Trigger [21] and Grim [80].

Axelrod was neither the first to demonstrate that cooperation is possible in the iterated prisoner's dilemma, neither to provide proof that reciprocity can be advantageous, however, he was the first to conduct these experiments in a such a well designed framework. As describe in [67], "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." It is important to mention however that the only source code available in the code for the 62 strategies of the second tournament, found on Axelrod's personal website [1].

His work offered him more than 9000 citations and opened a new avenue of research. In a collaboration [19] with a colleague, Douglas Dion, they summarized a number of works inspired from the "Evolution of Cooperation". Several of these are presented and discussed in the following section where we cover the response of the academic community to the computer tournaments.

2.3 Response to the computer tournaments (1984-1993)

The pioneering work of the computer tournaments and the results of the reciprocal behaviour in the prisoner's dilemma spread the knowledge of the game worldwide and across disciplines. Several researchers responded immediately to Axelrod's tournaments and the study of cooperation became of critical interest once again. This section focuses on the research that was carried out after the initial computer tournaments. More specifically, research that introduced new strategies in order to surpass several limitations of Tit for Tat over the time period between 1984 and 1993.

One the biggest assumptions in Axelrod's tournaments had been that each player had a perfect information of the opponent's actions. However, stochastic uncertainty severely undercuts the effectiveness of reciprocating strategies, [52] 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. Hammerstein [68], pointed out another weakness of Tit for Tat that if by mistake, one of two Tit for Tat players makes a wrong move, this locks the two opponents into a hopeless sequence of alternating defections and cooperations. A year later in 1986, [26] ran a computer tournament with a 10 percent chance of noise and Tit for Tat finished sixth out of 21 strategies.

A second type of noise is misperception, where a player's action is made correctly but it's recorder incorrectly by the opponent. In 1986, [75] introduced a strategy called **Contrite Tit for Tat** that was more successful than Tit for Tat in such environments. Contrite Tit for Tat has three states: *contrite, content, provoked*. It begins by cooperating and stays there unless there is a unilateral defection. If it was the victim of a defection while content the strategy becomes provoked and defects until the opponent cooperates, and causes it to become content. If it was the defector while content, it becomes contrite and cooperates. When contrite it becomes content only after there has been a mutual cooperation.

Such mistakes, such as noise and misperception, are unlikely to occur in a computer tournament, but have to be expected in real life situations. Actual biological situations are fraught with errors and uncertainties. The answer to the opponents last move is to increase or decrease the readiness to cooperate. This emerges quite clearly from [75] introduction of Contrite Tit for Tat, and [50] experiments on sticklebacks.

Another limitation of [16] was that the interactions have been between pairs of players, as argued by [36]. In several applications, however, interactions involve more than two players. This can be modelled using the corresponding n-player prisoner's dilemma (NPD), in which players make a choice (cooperate or defect) which they play with all other players.

In evolutionary settings [36] found that if individuals play a "hard" Tit For Tat, meaning that they will cooperate until one player defects, and the number of individuals playing hard Tir Fot Tat passes a certain threshold, then hard Tit For Tat can dominate a population of Defectors. But this threshold rises as the number of individuals in the society increases. "hard" Tit for Tat means that they will cooperate only when all remaining opponents cooperate.

The work of [48] pointed out that it is important 'to take more account of intrinsic stochasticities'. This suggested considering stochastic strategies and [57] studied such strategies. An iterated prisoner's dilemma strategy was 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. These are a very specific set of strategies that only remember their opponent's last move, not their own and they are called reactive strategies. 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)

This framework was used in [57] to study game dynamical aspects of the iterated prisoner's dilemma, the results of which will be presented in the next section which is dedicated to such research. Another outcome of the framework came the next year. In 1990, [58] 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 [32], which is described by a matrix M.

$$M = \begin{bmatrix} p_1q_1 & p_1(-q_1+1) & q_1(-p_1+1) & (-p_1+1)(-q_1+1) \\ p_2q_3 & p_2(-q_3+1) & q_3(-p_2+1) & (-p_2+1)(-q_3+1) \\ p_3q_2 & p_3(-q_2+1) & q_2(-p_3+1) & (-p_3+1)(-q_2+1) \\ p_4q_4 & p_4(-q_4+1) & q_4(-p_4+1) & (-p_4+1)(-q_4+1) \end{bmatrix} \tag{4}$$

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 (6). Thus [58] 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,

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

In 1992 reactive strategies were used to investigate which strategies would manage to take over the population and would be ESS in an environment with noise. The results demonstrated that though a small fraction of Tit for Tat players have been essential for the emergence of cooperation, more generous strategies took over the population. More specifically the re-active strategy known as **Generous Tit for Tat** which is give by the triplet $(1, 0, \frac{2}{3})$.

Generous Tit for Tat was not the only strategy to outperform Tit for Tat in noisy environment, same conclusions were made by [33, 35] In [35] a similar tournament to that of Axelrod's was performed, but this time it was a noise tournament. Bendor had invited researchers from several departments across his university. Each match would last a random number of turns, with a probability of 0.0067 of ending in the next turn. The results of his tournament demonstrated that Tit for Tat performed rather poorly and the highest ranked strategies were generous ones. The top ranked strategy was **Nice and Forgiving**. Nice and Forgiving, differs in significant ways from Tit for Tat.

Initially, Instead of reciprocating by returning an unbiased estimate of its opponent's action, Nice's generosity took the form of a benign indifference. It would continue to play cooperation as long as its rival's cooperation level exceeded 80%. Secondly, although it would retaliate if its rival's observed cooperation fell below 80, it was willing to revert to full cooperation before its partner did, so long as the partner satisfied certain thresholds of acceptable behaviour.

The next section will focus a bit more on the evolutionary settings and dynamics that the iterated prisoner's dilemma offered.

2.4 Evolutionary Dynamics (1987-1992)

The complex nature of iterated prisoner's dilemma strategies makes their evolutionary stability more complex to study. In this section we will cover several works in the 80's and early 90's that studied the dynamics of the iterated prisoner's dilemma strategies.

In [25] 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 [65]. Furthermore, Boyd and Lorderbaum, in [24], showed that cooperation can be started without a population structure if the correct combination of strategies is presented. This remark argued with Axelrod's result that only a cluster of cooperative strategies would succeed. The example they provided was the following:

Suppose that a population consists of Tit for Tat, Suspicious Tit for Tat and Tit for Two Tat. Suspicious Tit for Tat will end up cooperating after the first move, Suspicious Tit for Tat and Tit for Tat will each continue cooperating while the other defects (and vice versa), and Tit for Tat and Tit for Two Tat will cooperate on every move. In such a situation, Tit for Two Tat will be able to invade both Tit for Tat and Suspicious Tit for Tat, even without clustering.

Instead of using a given number of strategies, Nowak and Sigmund in [57], studied the dynamics of the evolutionary iterated prisoner's dilemma with a spectrum of stochastic strategies. The strategies they considered, as discussed in the previous section, were the reactive strategies. They managed to prove that there can be multiple fixed points that there can be an evolutionary stable coexistence among multiple such strategies.

There is limitation to all of these dynamic treatments. That is their inability to develop new strategies. A way of overcoming is to use a genetic approach. An evolutionary process called the genetic algorithm was used to discover effective strategies in [17]. A genetic algorithm is a search heuristic that is inspired by the theory of natural selection. In a population of candidate solutions, the fittest individuals are selected for reproduction. They produce offsprings that will replace the weakest members of the population as long as they do better than them.

In order to use a genetic algorithm [17] needed to represent strategies in a format such that the algorithm could optimise. Axelord 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. Axelord therefore used a list of 64 C's and D's to represent different strategies. We refer to this representation structure as a lookup table, which is a set of deterministic responses based on the opponents m last moves; for example [17] considered $m = 3$. In later section we discuss some more recent work that has been done using lookup tables.

Another dynamic approach that was considered in 1992 is that of using very simplistic strategies but in more complex topological structures. An extension to the natural selection where who meets whom is not random anymore. In [60], a population of 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 3, 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 (6):

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

where ($b > 1$).

It was shown that for different values of the temptation payoff b, this purely deterministic spatial version could generate chaotically change patterns in which cooperators and defectors could persist together in a mix population. Though it was known that in unstructured populations natural selection would favour defection, [56], provided evidence that in structured populations the results can be wildly different. The authors claimed that the essential results remain true of all topologies; the results also hold whether self interactions are taken into account.

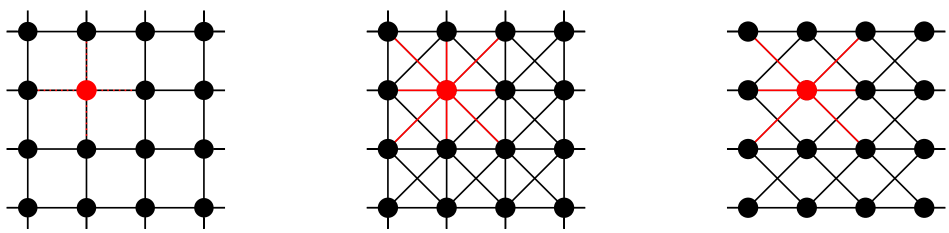


Figure 3: Spatial neighbourhoods

Later in 2003, the authors [47] decided to consider small world networks instead of regular graphs. More specifically they used the well known Watts and Strogatz [82] graphs. Watts and Strogatz are not regular graphs with a k regular degree. The experiment starts and each node has $k/2$ nearest neighbours on each side. Then a proportion p of the total edges is rewired by removing $pkn/2$ edges and creating $pkn/2$ new edges each of whose initial vertex is the initial vertex of a removed edge and its terminal vertex is randomly chosen so that the generated graph does not have multiple edges or once removed edges.

The authors argued that by applying these experimental rules we re closer to capturing real life behaviour where we constantly change who we interact with. The results of their work had been that:

- For small b, it is not so tempting for players to exploit cooperators. Thus cooperators converge regardless the value of p.
- The number of cooperators highly depends on p roughly for for higher values of the temptation payoff b.
- Lastly once temptation if very strong, even cooperators happen to form tight clusters, they cannot survive once they face defectors. Finally, the cooperators eventually extinguish.

In [62] where they considering regular k degree graphs introduced a rule regarding the favour of cooperation in spatial tournaments. They used the following game matrix,

$$\begin{pmatrix} b - c & c \\ b & 0 \end{pmatrix} \tag{6}$$

Ohtsuki etc all proved that natural selection favours cooperation, if the benefit of the altruistic act, b, divided by the cost, c, exceeds the average number of neighbours, k, which means $b/c > k$. Their result, however, holds for weak selection. That means that the fitness of an individual is only proportional to their payoff. Later the same year, Ohtsuki and Nowak [63], managed to approximate the dynamics of a population on graphs using an approximation of the replicator equations. The replicator equation was introduced as the first closed form differential equation to describe the dynamics of natural selected populations.

A lot of work has been done on evolutionary games on graphs. Though we have covered a number of academic publications [76] have also conducted a comprehensive review.

Another topic worth mentioning is that of coevolution on graphs. In [85] 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.

2.5 Modern approaches (1993-2017)

In this section we will cover several research projects published between 1993 and 2017. The research reviewed here focuses on computer tournaments and serves as an introduction to various strategies that have made an impact in the literature.

Another protagonist in the literature is a strategy called **Pavlov**, introduced in 1993 [59]. Pavlov has the tolerance of Generous Tit for Tat but also the capability of resisting and invading an all out cooperators population. 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. A variant of Pavlov was quickly introduced by [83]. That variant was Generous Pavlov, a variant that cooperates 10% of the times when it would either wise had defected.

An interesting approach of capturing promising strategies for the game was written in 1996 by [51]. Strategies represented by finite automata were learning to update their choices through a genetic algorithm. The specific type of finite automata that were used were Moore machines [53]. Finite state machine consist of a set of internal states. One of these states is the initial state of the machine. A machine also consits of transitions arrows associated with the states. Each arrow is labelled with A/R where A is the opponents last action and R is the players response. For example let us consider a graphical representation of the famous Tit for Tat given by a finite machine, Figure 4.

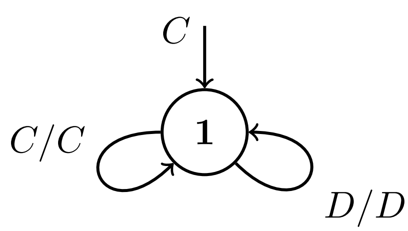


Figure 4: Finite state machine representation of Tit for Tat.

Miller used the genetic algorithm to train finite state machines in enviroments 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.

In 1997, **Gradual** another well performed strategy was proposed by [23]. 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 [23], Gradual had managed to outperform strategies such as Tit for Tat and Pavlov.

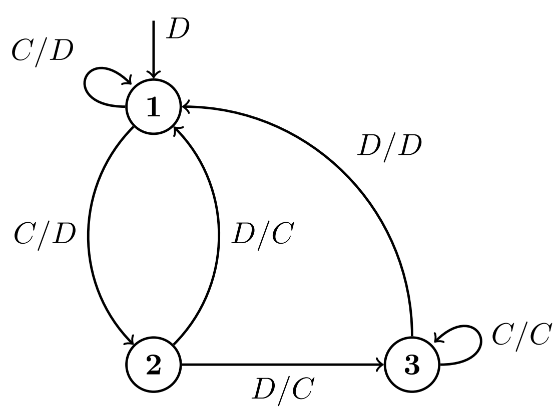
Following the success of Gradual the authors of [79] conducted the same tournament, with now 13 strategies, desinged to outperform Gradual. Their strategy was the **Adaptive Tit for Tat** and the algorithm used by it is given by Algorithm 1. Adaptive Tit for Tat ranked first in it's tournament surpassing Gradual. It should be noticable that the number of opponents was low and might have influenced the perfomance of Adaptive Tit for Tat.

Algorithm 1 Adaptive Tit for Tat.

```
if opponent played C in the last cycle then
  world = world + r(1 − world), r is the adaptation rate
else
  work = world + r(0 − world)
end if
if world ≥ 0.5 then
  play C
else
  play D
end if
```

In 2006, Slany and Kienreich, tried to deal with one of the problems of Tit for Tat. This problem was described above and its that Tit for Tat players would go into a chain of mutual defections if noise were to be added in the environment. Furthemore, the 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 behavour to act as a Defector. They called their strategy OmegaTFT [70].

Similar to Miller's work, in [12] the author presented two new strategies that have been trained using a finite state machine representation. These strategies are called, **Fortress3** and **Fortress4**. Figure 5 illustrates their diagrammatic representation.



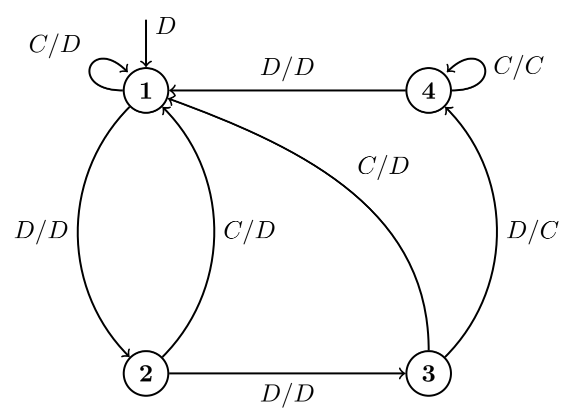


Figure 5: Representations of Fortress 3 and Fortress 4. Note that the strategy's first move, enters state 1, is defection for both strategies.

During the experiments that introduced Fortress3 and Fortress4 a large diversity of strategies were found. Differencianting between strategies represented as a finite machine is not an easy task. It is not obvious looking at a finite state diagram how a machine will behave, and many different machines can represent the same strategy. In order to distinguish the strategies and assuring that they are indeed different [7] introduced a method called fingerprinting. 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 6 an example of Pavlov's fingerprint is given. Fingerprinting has been studied in depth in [8, 9, 10, 11].

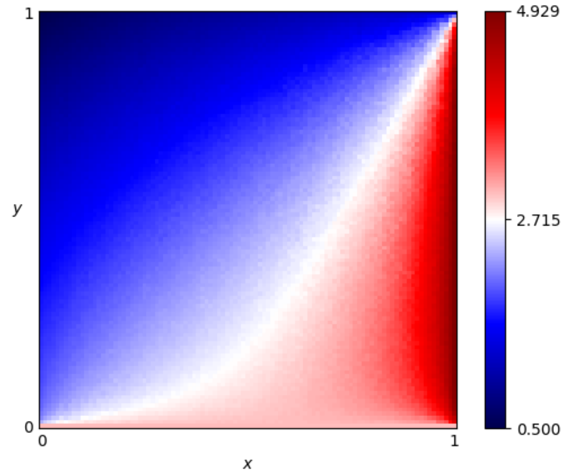


Figure 6: Pavlov fingerprinting with Tit for Tat used as the probe strategy. Figure was generated using [5].

In [42] **APavlov**, which stands for adaptive Pavlov, made an appearance. 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. In 2011 the authors of [42] 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 2012 Press and Dyson [64] studied the iterated prisoner's dilemma and 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. In Section 2.3 it was described how the payoffs of two players could be retrieved by formulating their interactions using a Markov chain. Let us denote the payoffs of players p and q 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 results suggested that the best strategies were selfish ones that led to extortion, not cooperation. Arguing with Axelrod's reports. All the more, their work stated that in the iterated prisoner's dilemma, memory is not advantageous.

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" [73]. In [73], they ran a variant of Axelrod's tournament with 19 strategies to test the effectiveness of ZD strategies. While conducting their tournament they have implement several strategies discussed by [64] and revealed a set of generous ZDs the **Generous ZD**.

In [41], 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. Complex strategies were also studied by [34, 39]. This was done using an open source package, called the Axelrod project [5] which launched on 2015.

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 [34] and a paper describing it and it's capabilities was published in 2016 [38].

The two paper using the Axelrod project [34, 39] 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 [74]. A number of strategy representations, referred as archetypes, were used to train strategies. These included, lookerup tables, finite state machines, artificial neural networks [84] and hidden Markov models [28].

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 [39] 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 [34], 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" [13].

These tournaments are, to the authors knowledge, the biggest tournaments and evolutionary progress that have been done in the field.

2.5.1 Software

Though several of this tournament discussed so far were generated using computer code not all of the source code was made available by the authors. Several open projects were created and published through the year. The first one discussed in this work, excluding the code for Axelrod's strategies, is PRISON [4]. 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 [23] which introduced Gradual and [22]. The project includes a good number of strategies from the literature but unfortunately the last update of the project dates back in 2004.

Other recent software 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 matches and tournament configurations.

Finally, as described in Section 2.5, the open source package Axelrod. Axelrod package is a software written 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 [55] and the package has not been used only by the contributors for academic research but from several such as: [40].

2.6 Contemporary period (2017 - 2018)

In research years the study of the iterated prisoners' dilemma is still active and papers are still being published. New strategies, new variants of the game and new applications are being introduced every year. In this section we will briefly review some articles that have been published between 2017 and 2018.

The iterated prisoner's dilemma serves as a model in a wide range of applications. For example in [37] cancer cells and how they can resist treatment has been modelled using evolutionary approaches described in Section 2.4. Furthermore, in [27] they explore whether learning in social situations can be driven by rewards.

A lot of work has been done on evolutionary dynamics on structured populations. This is mainly because their applications in social studies. However, understanding evolutionary game dynamics in structured populations remains difficult. In [44] the authors consider the evolutionary spatial prisoner dilemma with memory one strategies and their results indicate that a Pavlov like behaviour is stable and dominant. But how cooperation evovle in such structure situations and more specifically in situations where the number of neighbours can vary, still remains answered. In [6] the authors tried to shed some light to the question. However this was done for week selection. Their results argue that by considering the the coalescence times of random walks of any given graphs they can approximate when cooperation will emerge.

Several variations to the actual game are still being introduced. Ohtsuki in [61], studies the NPD, driefly introduced in section 2.3. Ohtsuki propose a model called coordinated cooperation. It's an NPD game which starts with is a negotiation before an actual game is played. Each individual can flexibly change their decision, either to cooperation or to defection, according to the number of those who show the intention of cooperation/defection. This NPD model was introduced in order to provide one explanation of why people tend to take into account others' decisions even when doing so gives them no payoff consequences at all.

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

[6] 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.

[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] 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.

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

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

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

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

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

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

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

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

[22] 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.

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

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

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

[26] 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.

[27] 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.

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

[29] 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.

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

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

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

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

[34] 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.

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

[36] N. V. Joshi. Evolution of cooperation by reciprocation within structured demes. *Journal of Genetics*, 66(1):69–84, Apr 1987.

[37] 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.

[38] 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.

[39] 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.

[40] Julian F. Latorre, Juan Pablo Ospina, and Jorge E. Ortiz. A coalitional game for achieving emergent cooperation in ad hoc networks through sympathy and commitment. In Juan Carlos Figueroa-García, Eduyn Ramiro Lopez-Santana, and Jose Ignacio Rodriguez-Molano, editors, *Applied Computer Sciences in Engineering*, pages 352–362, Cham, 2018. Springer International Publishing.

[41] 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.

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

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

[44] 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.

[45] 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.

[46] 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.

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

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

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

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

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

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

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

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

[55] 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.

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

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

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

[59] 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.

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

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

[62] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502, 2006.

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

[64] 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.

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

[66] 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.

[67] 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.

[68] 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.

[69] 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.

[70] 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.

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

[72] 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.

[73] 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.

[74] 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.

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

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

[77] 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.

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

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

[80] 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.

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

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

[83] 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.

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

[85] 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.