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 was 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 [27].

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 (1). 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. For a dilemma to exist the payoffs are subject to constrains (2) and (3).

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

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

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

Due to rational behaviour and constrains (2), (3) it would never benefit a player to cooperate. It can 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 outcomes matter, 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 when players interact for more than one time. One of the most important of these works has been R. Axelrod's as described in his book [18] "The Evolution of Cooperation".

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 **C** and **D** again and again while having memory of their previous encounters. Academics from several fields were invited to design computer strategies to compete in the tournament. 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.

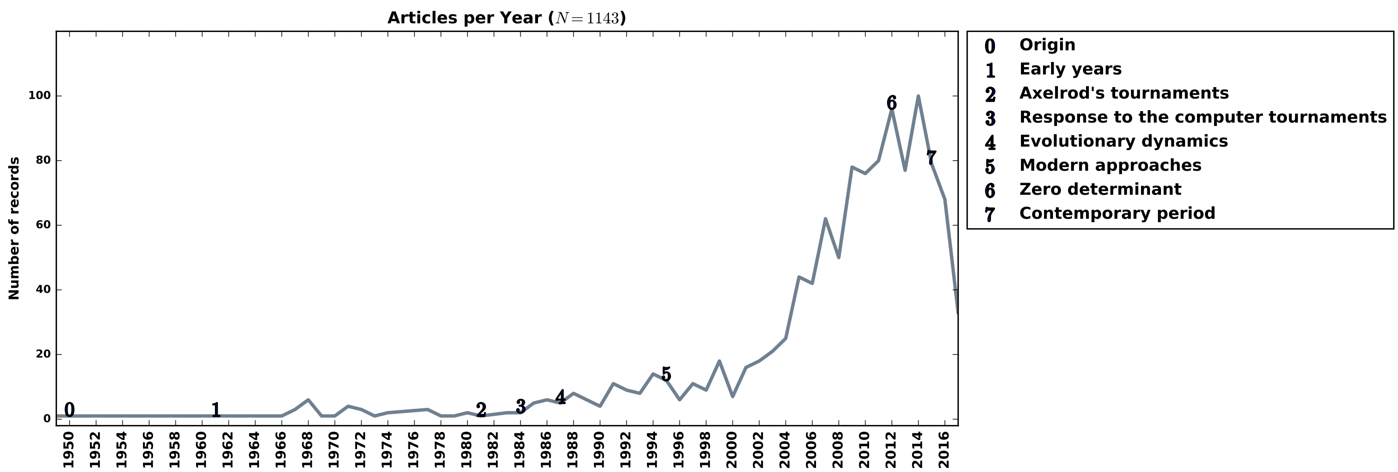

Figure 1: A timeline of the prisoner's dilemma research.

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 Origin and Primal research (1961-1972)

The origin of the prisoner's dilemma goes back to the 1950s in early experiments conducted at the RAND [27] to test the applicability of games described in [65]. Although the name behind the game was given later the same year. According to [61], A. W. Tucker (the PhD supervisor of J. Nash), in an attempt to delivery the game with a story during a talk used prisoners as players 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 is Prof A. Rapoport, a mathematical psychologist, whose work focused on peacekeeping. In his early work [53]

Rapoport conducted experiments using humans to simulate a play of the prisoner's dilemma. Experimental groups have not been used only by Rapoport. They are a common mean of studying the game [25, 28, 38, 39, 56] used until today.

The early experiments were exploring the conditions under which altruist behaviour emerges in human groups. Conditions explored were the gender [25, 38, 39] of individuals, the representation of the game [25], the distance between players [56], the initial effects [60] and whether the experimenter was biased [28].

The community however soon came to ask a different question: "what is the best way to play the game?" Inspired by the work of Rapoport a political scientist R. Axelrod conducted the first ever, to the author's knowledge, computer tournaments of the IPD. According to Axelrod [16], he decided to use computers because human participant were known to act very randomly even though the aim of the game is clear to them.

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 [16], 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 [12]. 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 13 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 (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 the player it was interacting with.

In order to further test the results Axelrod performed a second tournament [13] later in 1980. The results of the first tournament were publicized 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' [17]) 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, the winner was once again Tit for Tat. The conclusions made from the 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 [13], 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 [13], 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. In Axelrod's ecological tournament the Tit for Tat continued to thrive and the simulation of the ecological tournament had shown that the Tit for Tat rule was used by everyone. Ecological tournament does not offer any evolutionary perspective. It's only an estimation of the strategies' frequencies based on the tournament payoff matrix.

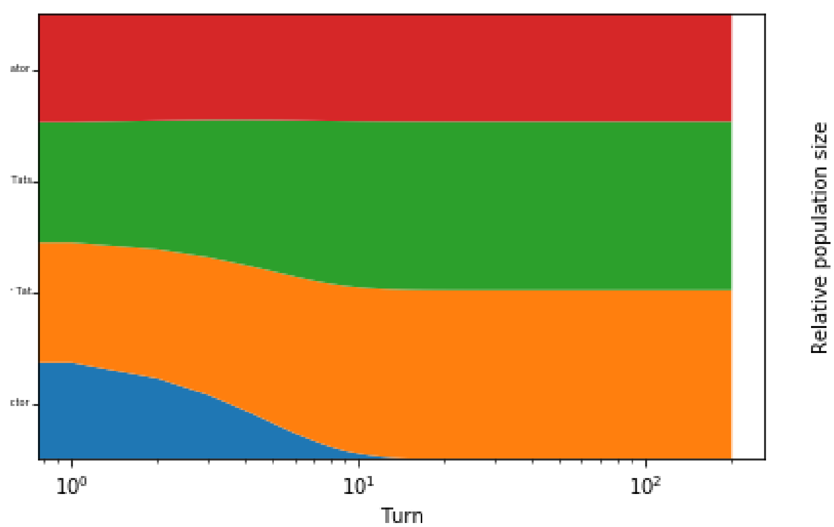

Figure 2: Results on an ecological tournament with Tit for Tat, Tit for Two Tats, Cooperator and Defector.

In 1981, Axelrod also study the prisoner's dilemma in an evolutionary context based on the evolutionary approaches introduced by an attendant of the second tournament and a well known evolutionary biologist, John Maynard Smith [41, 57, 58]. John Maynard Smith alongside his collaborator George Price are considered fundamental figures of evolutionary game theory. In [41] 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 the strategy B is an population that interacts only with itself, the concept of invasion is equivalent to a single mutant being able to outperform the average population. This leads to the concept of the evolutionary approach. Thus for a strategy to be ESS it must be able to resists any invasion.

The work described in [14], studied the evolutionary stability of Tit for Tat and introduce the definition of a collective stability. Although the strategy was likely to take over the population, its stability was conditional on the importance of the future of the game, which is represented by a discounting factor 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 such players could invade a hostile environment.

Axelrod decided to work on the biological applications of his 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 [18] won the Newcomb-Cleveland prize of the American Association for the Advancement of Science.

Overall, all of Axelrod's tournaments offered many insights and new concepts to the field. These were not only limited on Tit for Tat. For example several strategies of these tournaments are still being used in literature to date, 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 [37]. Though the strategy goes by many names in the literature such as, Spite [20], Grim Trigger [19] and Grim [64].

Though not all strategies of Axelrod's tournament are retrievable. The author had given a explanation of all 13 strategies of the first tournament. The size of the second tournament did not allow him this time to go into details for every single participant. The author mainly focused on the high ranked participants. However, the source code of the 63 strategies can be found on Axelrod's personal website [1]. Figure 3 serves as an example of the source code for the winning strategy Tit for Tat.

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

Figure 3: Source code for Tit for Tat in Fortran. Provided by [1].

Note though that the code of the second tournament only includes the strategies. The code for running a round robin tournament or an ecological tournament is not given. Moreover, the source code of the first 13 strategies is not available, as stated in Axelrod's personal website [1].

2.3 Response to the computer tournaments (1984-1993)

The pioneering work of 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 the research the falls in the following categories, over the time period between 1984 and 1993, are considered in this section:

- Applications of the iterated prisoner's dilemma, the famous strategy Tit for Tat and the reciprocal behaviour.

- Computer tournaments under stochastic uncertainty and the introduction of new strategies.

One of the scientific disciplines that immediately employed Axelrod's work has been the ecological field. More specifically, the works of [23, 24, 29, 42, 66]. In [24, 42] the behaviour of fish when confronting a potential predator was studied. Conflicts can arise within pairs of fish in these circumstances. In both works experiments were held using a system of mirrors where sticklebacks and guppies respectively, would be accompanied by a cooperating companion or a defecting one. In both cases the hypothesis that the fish would behave according to Tit for Tat and that cooperation would evolve was supported. The works of [29, 66] looked at food sharing between vampire bats and explained behaviour based on known strategies.

Another quick response to the tournaments was that of [44]. It was argued that Axelrod's work assumed that each player had a perfect information of the opponent's actions. In real life situations this is not always the case. Interactions often suffer from measures of uncertainty and this was not captured in the original tournaments. Molander studied the performance of Tit for Tat in an uncertain environment by introducing noise. A probability that a player's move will be flipped. It was proven that when two strategies playing Tit for Tat met in a noisy match the average payoff of each strategy would be the same as that of a Random player (with a probability 0.5 of cooperating).

In 1988 publications from other disciplines were using the iterated prisoner's dilemma and Axelrod's work for teaching and social studies. In [35] a version of the prisoner's dilemma which set the problem in an ordinary business context was used as a pedagogic instrument within graduate business students. The work of [52] considered non zero sum games, specifically the prisoner's dilemma, and illustrated the impact they have on societal problems such as war.

In 1989 reactive strategies are introduced in [46]. Reactive strategies are a set of players that take into account only the last move of the opponent. Thus, they can be represented by the probability of cooperating after an opponent's **C** or **D**. The same author a year later in 1990 gave a formal definition of a memory one [47]. 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, 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. The utility of a player can be given by multiplying the steady states of M by the payoffs of equation(1). Thus [47] offered a mathematical framework to calculate the utility of two players without simulating the game. Memory one strategies are considered an important family of strategies and the formulation given in 1990 is still being used to date.

A player called **Handshake** was presented by [55] in 1989. Handshake is a strategy that starts with cooperation, defection. If the opponent plays in a similar way then it will cooperate forever, otherwise it will defect forever. Handshake has a property that will be revisited in this literature review which is it's recognition property.

From 1991 to 1992, further research explored the performance of Tit for Tat in uncertain environments. These works include [29, 32, 45]. In [32] 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 and from a university seminar. Each match would last a random number of turns, with a probability of 0.0067 of ending in the next turn. The probability of noise was a random variable, distributed normally with a mean of zero and a standard deviation of eight. 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**.

The work of [45] aimed to also investigate stochastic effects. This is was done using an evolutionary setting of a heterogeneous population with noise. The author investigate which strategies would manage to take over the population and would be evolutionary stable. The strategies were explored over the space of reactive strategies. 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 presented as $(0, \frac{2}{3})$.

The author of [45, 46, 47] also introduced another important strategy of the literature. It was presented in [48] and is it known as **Pavlov**. Pavlov is a strategy with 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. Otherwise it shifts it's last move.

2.4 Evolutionary Dynamics (1987-1999)

Determining the evolutionary stability of strategies for the iterated prisoner's dilemma as we discussed is not an easy task. Methods can be use to deal with the difficulty. In [22] the author restricted the possible strategies that could be adopted to a relatively narrow set and resulted that no pure strategy is evolutionary stable, including Tit for Tat. Arguing with the results presented in [14]. The list of strategies used included strategies such as Defector and **Suspicious Tit for Tat**, a strategy that plays Tit for Tat but starts by defecting.

The results were questioned by [40], stating that much was still no fully explored and more research had to be put into the results. Farrel and Ware in 1989 [26] extended the result to include finite mixture of pure and mixtures of Tit For n Tats as well. On the same year the work of [21] looking again at a narrow set of strategies extended their results to noisy environments.

Evolutionary dynamics have been highly useful in the research of the prisoner's dilemma. In [15], an evolutionary process, called the genetic algorithm, was used to discover effective strategies. The author introduced lookup tables as a mean of representing a strategy in a gene format. A lookup table is a set of deterministic responses based on the opponents m last moves; [15] considered $m = 3$.

An extension to the natural selection was introduced in the 1992 [50], recommending a different type of topology. A population of two deterministic strategies, Defector and Cooperator, were placed on a 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 4. The authors claimed that the essential results remain true of all topologies; the results also hold whether self interactions are taken into account.

Thus each cell of the lattice is occupied by a 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.

Nowak studied the population dynamics as a function of the temptation payoff. It was shown that for different values of the temptation payoff, cooperators and defectors could persist together.

This work dealt with dealt with symmetric spatial lattices in two dimensions, deterministic winning and discrete time. The authors in later work [49], that the results remain valid in more realistic situations. Such as situations where the spatial distributions of cells are random in two or three dimensions, and where winning is partly probabilistic.

2.5 Modern approaches (1995-2015)

In this section we will cover several research projects published between 1995 and 2015. The research reviewed here focuses on computer tournaments and serves as an introduction to various strategies that have made an impact in the literature.

Initially in 1995 a combination of tournament studies, ecological simulations and theoretical analysis was used in [67] to demonstrate approaches on coping with noise. The first approach was *generosity*. A more generous version of Tit for Tat, the strategy called the Generous Tit for Tat, was proven to be highly effective against players that were not adapted to

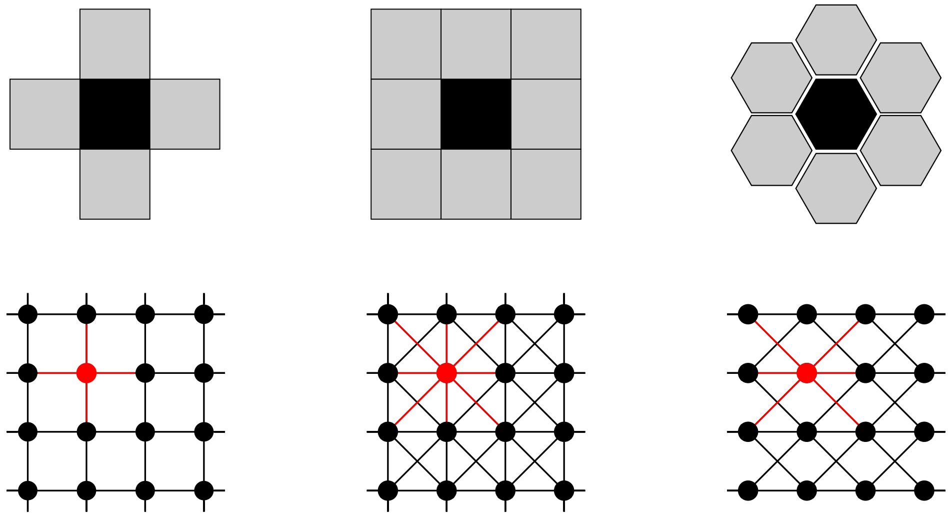

Figure 4: Spatial neighbourhoods

noise. The second approach introduced was *contrition*. The authors stated that when interacting with strategies that have been adapted to noise a contrite version of Tit for Tat is even more effective at quickly restoring mutual cooperation without the risk of exploitation. The strategy introduced as the contrite variant was called **Contrite Tit for Tat**. A strategy which 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. The final approach has been using strategy Pavlov. Moreover, in the analysis an variant of Pavlov was also included called Generous Pavlov, a variant that cooperates 10% of the times when it would either wise had defected. Though all approaches were effecting in coping with noise in a standard tournament, the authors argued that Pavlov was not suited for ecological tournaments.

An interesting approach of to capture promising strategies for the game was written in 1996 by [43]. Strategies represented by finite automata were learning to update their choices through an explicit evolutionary process modelled by a genetic algorithm.

Strategies based on finite state machines are described by the number of states. The strategy selects the next action in each round based on the current state and the opponent's last move, transitioning to a new state each time. Figure 5, illustrates the finite state representation of Tit For Tat.

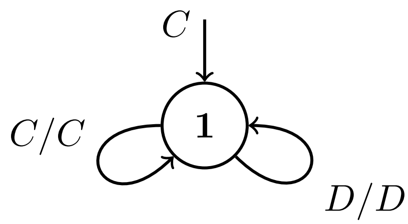

Figure 5: Finite state machine representation of Tit for Tat.

In 1997, **Gradual** a well performed strategy and a commonly known in the literature was proposed by [20]. 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. Gradual had managed to outperform strategies such as Tit for Tat and Pavlov.

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. The project includes a good number of strategies from the literature but unfortunately the last update of the project dates back in 2004.

Another measure of uncertainty discussed in 1998 [31] is that of mis perception. Mis perception is the probability that the opponent's current move is flipped before being recorded to the history. Results of [31] indicated that noise effected the emergence of cooperative behaviour in the populations. In 1999, [62] uses evolutionary game theory to study the spread of virus.

Following the success of Gradual the authors of [63] conducted a tournament of 13 strategies just to outperform it. Their strategy was the **Adaptive Tit for Tat** and the algorithm used by it is given by 1.

Algorithm 1 Adaptive Tit for Tat.

if opponent played **C** in the last cycle **then**
  world = world + $r(1 - \text{world})$, r is the adaptation rate
else
  work = world + $r(0 - \text{world})$
end if
if world ≥ 0.5 **then**
  play **C**
else
  play **D**
end if

Adaptive Tit for Tat ranked first in it's tournament surpassing Gradual. However, the results were drawn from a tournament of 13 strategies.

In [11] the author presented two new strategies that have been trained using a finite state machine representation. They are called, **Fortress3** and **Fortress4**. Figure 6 illustrates their diagrammatic representation where the transition arrows are labelled O/P where O is the opponent's last action and P is the player's response.

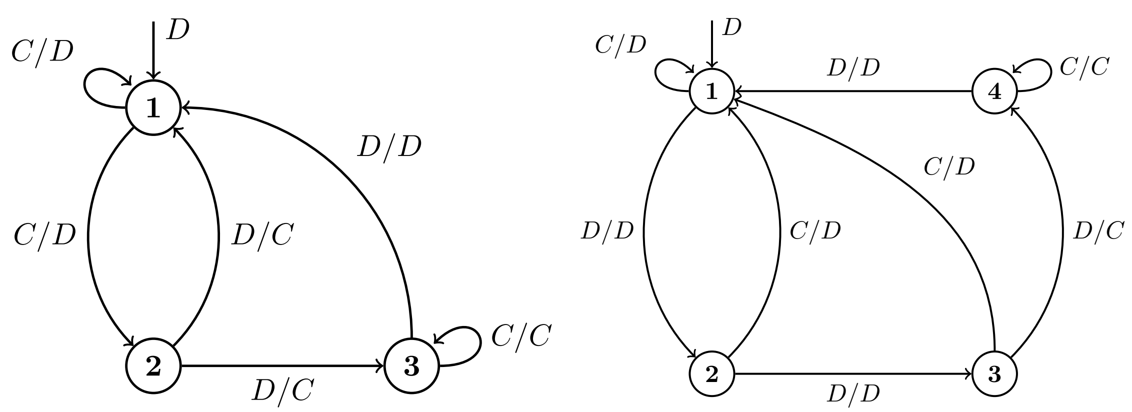

Figure 6: Representations of Fortress 3 and Fortress 4. Note that the strategy's first move, enters state 1, is defection for both strategies.

Optimisation methods return a spectrum of strategies. In order to distinguish the strategies and assuring that they are indeed different [6] introduced a method called fingerprinting.

The method of fingerprinting is a technique for generating a functional signature for a strategy [7]. 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 [6] Tit for Tat is used as the probe strategy. Fingerprint functions can then be compared to allow for easier identification of similar strategies. In Figure 7 an example of Pavlov's fingerprint is given. Fingerprinting has been studied in depth in [7, 8, 9, 10].

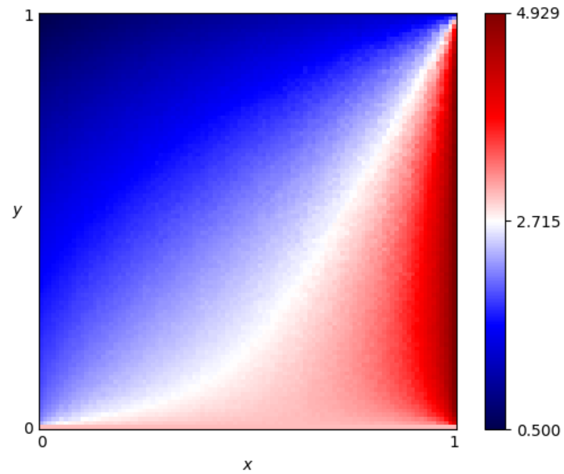

Figure 7: Pavlov fingerprinting with Tit for Tat used as the probe strategy. Figure was generated using [5].

The adaptive behaviour that was tried to be capture by [63] was not constrained only on Tit for Tat. In [36] **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 [36] 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.

2.6 Zero determinant (2012 - 2015)

In 2012 Press and Dyson [51] 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:

$$s_p = vS_p$$
$$s_q = vS_q$$

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" [59]. In [59], 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 [51] and revealed a set of generous ZDs the **Generous ZD**.

In [34], 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.

2.7 Contemporary period (2015 - 2017)

In the following section we report research and publications from 2015 to 2017. It will focus on sophisticated strategies and computer software.

In 2015, an open source library, called the Axelrod project [5] was launched. 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 [30] and a paper describing it and it's capabilities was published in 2016 [33]. The source code for Tit for Tat as implement within the library is shown in Figure 8. Furthermore, performing a tournament with a selection of strategies is possible in five lines of code, shown in Figure 9.

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

Figure 8: Source code for Tit for Tat in Python as implemented in Axelrod Python library [5]

.

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

Figure 9: Performing a computer tournament using [5].

In [54], the authors claim that they have managed to re-run the first tournament that Axelrod performed. They tried to push his work further by altering aspects such as, the format of the tournament, the objective and the population. One of the authors claimed to have been a contributor to the first tournaments, which would explain how it was managed to reproduce the tournament.

A number of strategies based on artificial neural networks are introduced by [30], in 2017. Artificial neural networks provide a mapping function to an action based on a selection of features computed from the history of play.

These strategies are refereed to as **EvovlvedANN** strategies and are based on a pre-trained neural network with the following features,

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

A representation of **EvovlvedANN 5** is given in Figure 10. The inputs of the neural network are the 17 features as listed above. Number 5 reefers to the size of the hidden layer.

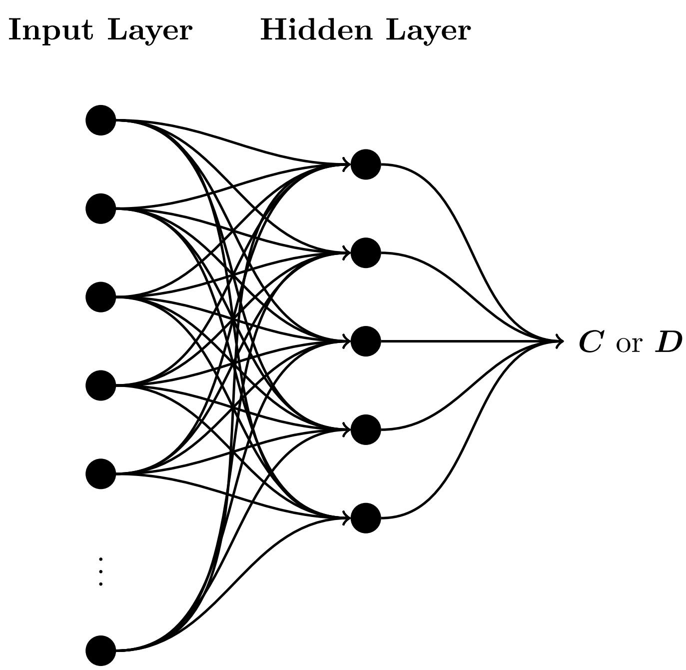

Figure 10: Neural network representation of EvovlvedANN 5.

In [30], these representing methods are refereed to as archetypes. Finite state machines and artificial neural networks are included in the work but also new archetypes are introduced, such as hidden Markov models. A variant of a finite state machine that use probabilistic transitions based on the prior round of play to other states and cooperate or defect with various probabilities at each state. Finite state machines and hidden Markov models based strategies are characterized by the number of states. Similarly, artificial neural networks based players are characterized by the size of the hidden layer and number of input features.

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

In [30] evolutionary algorithms are used to introduce as stated by the authors the best performing strategies for the iterated prisoner's dilemma. These strategies will be refereed as **Evolved** strategies. Several successful new strategies are,

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

Though several papers have claimed before to have discovered the dominant strategies for the game the work of [30] is promising. This is due the fact that the introduced strategies have been trained using different types of evolutionary algorithms in a pool of 176 well known strategies for the literature. Including all the strategies that have been discussed in this section.

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 match and tournaments configurations. Such as noise.

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

[7] 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.

[8] 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.

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

[10] 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.

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

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

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

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

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

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

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

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

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

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

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

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

[23] John L Craig. Are communal pukeko caught in the prisoner's dilemma? *Behavioral Ecology and Sociobiology*, 14(2):147–150, feb 1984.

[24] L A Dugatkin. Do guppies play TIT FOR TAT during predator inspection visits? *Behavioral Ecology and Sociobiology*, 23(6):395–399, dec 1988.

[25] 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.

[26] Joseph Farrell and Roger Ware. Evolutionary stability in the repeated prisoner's dilemma. *Theoretical Population Biology*, 36(2):161–166, 1989.

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

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

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

[30] 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.

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

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

[33] 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.

[34] 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.

[35] Leon Levitt. A business version of the prisoner's dilemma: A teaching technique. *International Journal of Value-Based Management*, 1(2):83–90, jun 1988.

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

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

[38] 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.

[39] 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.

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

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

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

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

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

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

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

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

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

[49] Martin A Nowak, Sebastian Bonhoeffer, and Robert M May. Spatial games and the maintenance of cooperation. *Proceedings of the National Academy of Sciences*, 91(11):4877–4881, 1994.

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

[51] 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.

[52] G. Rabow. The social implications of nonzero-sum games. *IEEE Technology and Society Magazine*, 7(1):12–18, March 1988.

[53] 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.

[54] 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.

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

[56] 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.

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

[58] 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.

[59] 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.

[60] 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.

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

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

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

[64] 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.

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

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

[67] 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.