Literature review paper for the iterated prisoner's dilemma.

Nikoleta E. Glynatsi

2016

1   Introduction

The emergence of cooperation is a topic of continuing and public interest for social, biological and ecological sciences. The prisoner's dilemma is a fundamental game commonly used in the evolution of altruistic behaviour.

The prisoner's dilemma is a two players no-cooperative game where the decisions of the players are made simultaneously and independently. Both players can choose between cooperation ($\mathbf{C}$) and defection ($\mathbf{D}$).

The fitness of each player is influenced by its own behaviour, and the behaviour of the opponent. If both players choose to cooperate, both do better than if both defected. A player has the temptation to deviate, cause if one player were to defect while the other cooperates, the defector receives more than if both had cooperated. The reward for mutual cooperation is R units, for a mutual defection they receive P, and for cooperation-defection, the cooperator receives S where the defector receives T.

Thus, the game's payoff are defined by,

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

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

Though the one shot game illustrate how players will not trust their opponents and will act selfishly, choosing defection, greater insights can be achieved by studying the game in a manner where the prior outcomes matters. The repeated form is called the iterated prisoner's dilemma, and it will be discussed later on how, it was proven to leave more room for cooperation to emerge.

The origin of the prisoner's dilemma go back to 1950 in early experiments conducted in RAND [17] to test the applicability of games described by [47]. In [17] they introduced a two player non-cooperative game but the story behind the game was given later the same year. A. W. Tucker, who was John Nash's supervisor, in his attempt to tell a story during one of his talks gave the background story of the prisoner's dilemma as we know it today [45].

In the following section several milestones on the research of the prisoner's dilemma are presented and discussed. Figure **??** illustrates a timeline generated using the open source library discussed in Section 3. This illustrates that the prisoner's dilemma has been under continuous research since it's origins. Several insights and applications have been offered in research by the game, these will be discussed in the following sections.
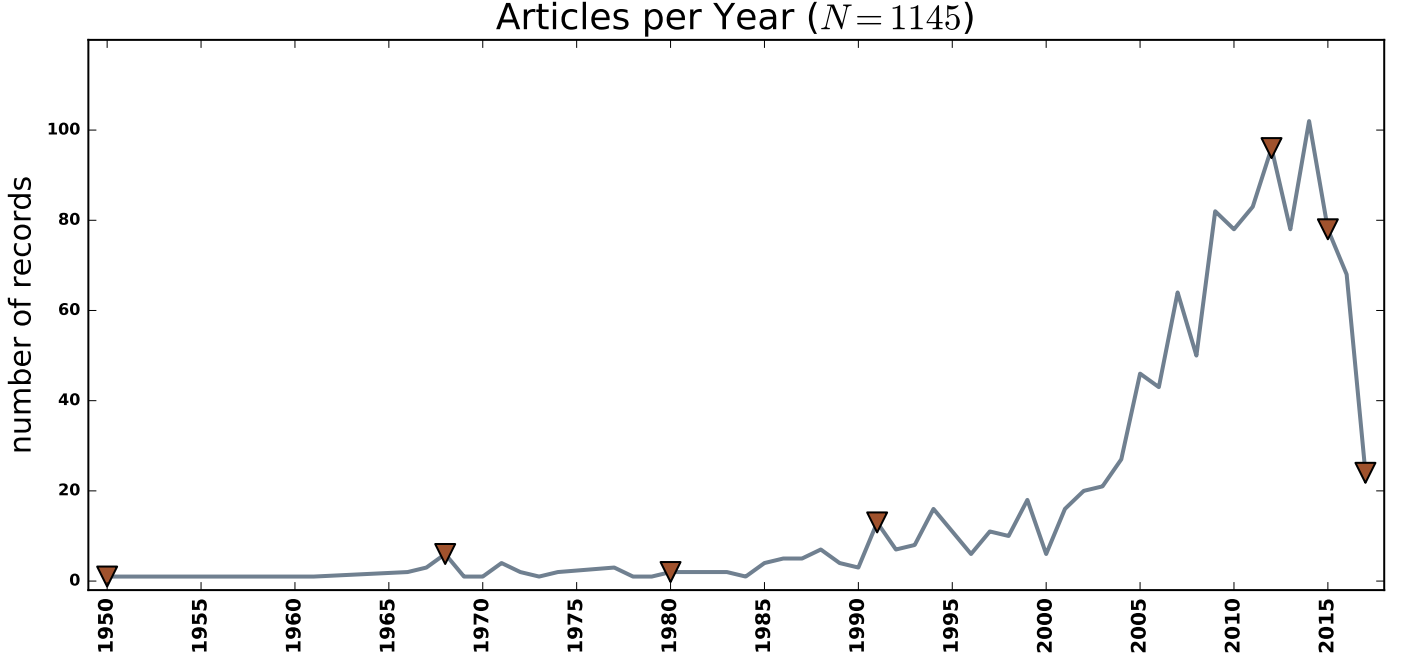
Figure 1: A timeline highlighting the milestones of the prisoner's dilemma.

2 Timeline

2.1 Early Years

The study of the prisoner's dilemma attracted people from various field. An important figure within the field is professor Anatol Rapoport, a mathematical psychologist focused on war and peacekeeping. In his early work [39] he conducted experiments using humans to simulate the interactions of prisoner's dilemma game. Human experiments have not been used only by Prof Rapoport, instead it was a common way of studying the game [16, 18, 29, 30, 43] and it still being used to date.

These early experiments explored the conditions under which altruist behaviour emerges. Researchers were in search of an unbeatable strategy to play the game. Inspired by the work of Rapoport and the idea that AI was trained to play the game of chess [10], the political scientist R. Axelrod performed the first ever computer tournament, known to the author, of the iterated prisoner's dilemma [12].

2.2 Reciprocal Period

In 1980 [7] a computer tournament of the iterated prisoner's dilemma took place with 14 participants. The tournament was of a round robin topology where each strategy played against all the opponents, itself and the Random strategy (a strategy that chooses between **C** and **D** randomly). Each participant knew the exact length of the matches and had access to full history of each match. The payoffs values of (1) used by Axelrod were the following, $R = 3, P = 1, T = 5$ and $S = 0$. These values are the most common used in literature and assume that they are being used in the works referenced from now on unless is stated otherwise. The winner of the tournament was determined by the total average score and not by the number of matches wins. The strategy that was announced the winner was submitted by Prof. Rapoport and was called called Tit For Tat.

Tit for Tat, was a strategy that always cooperated on the first round and then mimics the opponent's previous move. his is illustrated diagrammatically in Figure 2. To further test the robustness of the results a second tournament was performed

later with a total of 63 strategies [8]. All the opponents knew the results of the previous tournament by this time the number of turns was not specified. Instead a probabilistic ending tournament was used. Each match has probability of ending after each move. This is also refereed as 'shadow of the future' is some other works [11]. The winner of the second tournament was once again the strategy Tit For Tat. Tit for tat was an example of how reciprocity behaviour allows for cooperation to emerge in the iterated game.
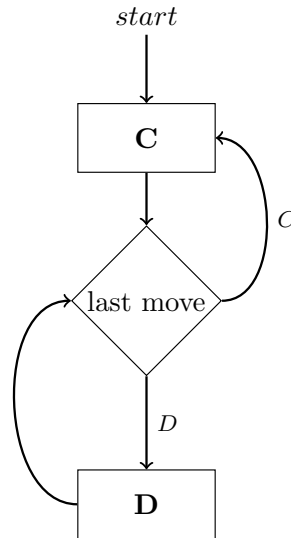

Figure 2: Diagrammatic representation of Tit for Tat.

According to Axelrod, the secrets behind the strategy's success have been 1) that it start of by cooperating 2) it would forgive it's opponent after a defection 3) after opponents identified that they were playing Tit for Tat choose to cooperate for the rest of the game.

In [12], the strategies set of second tournament was used to perform a ecological kind of tournament. The 63 strategies interact generation after generation to a round robin competition where their frequencies is proportional to their payoff in the previous round. Results showed that in a homogeneous population of Tit for Tat invasion by mutant strategies was not successful. The success of Tit For Tat was very soon known world wide and several researcher focused their work on the strategy ever since.

2.3 Stochastic Environments

But success often comes with criticism. Axelrod's tournaments assumed that each player has perfect information of the opponent's actions. In real life situation this is not always the case. Colleagues interactions often suffer from measures of uncertainty. In the original tournaments there was no possibility of misimplementation or misunderstanding. These stochastic variations are often called **noise** and **misperception**.

Noise is the concept of flipping ones moved based on a probability. On the contrary, misperception is the probability that the opponents current move is flipped before being recorded. Noise will flip a players action it will be recorded correctly in the history where mis perception will not have an effect on the player move but it will be recorded wrong [23].The performance of Tit for Tat was proven to suffer from such stochasticity in the tournament environment, especially against itself [4, 19, 34, 35].

An interesting result was introduced by [34]. If two players are both using the Tit for Tat strategy, both players would get the same average payoffs as two interacting Random players with $p = 0.5$. In [35], they used evolutionary dynamics and showed that to cope with noise a more generous version of Tit for Tat is needed.

2.4 Era of Strategies

Following the pioneer work of computer tournaments many researchers have sought to understand which strategies are dominant when playing the repeated prisoner's dilemma. These strategies can be anything, from deterministic to complex. The can make use of the history of the game, the length or nothing at all. The size of history a strategy takes into account is refereed to as memory size of the strategy.

The two most common deterministic strategies used in various works are **All Defector** and **All Cooperator**. Both strategies have been introduced in [12]. All Defector defects in each turn, compared to All Cooperator which is a strategy that always cooperates.

In [35], the space of re-active strategies was explored and the strategy that stand out was a re-active strategy known as **Generous Tit for Tat**. Reactive strategies are denoted by the probabilities to cooperate after a **C** and a **D** of the opponent. Thus, a reactive strategy only considers the previous turn of the opponent.

Reactive strategies are a subset of memory one strategies introduced in 1989 [36]. Memory one strategies, are a set of strategies that consider only the last turn of the game to decide on the next action [37]. They are represented by the four conditional probabilities p_1, p_2, p_3 and p_4 to cooperate after CC, CD, DC and DD respectively (the four possible states a player can be in if only the last turn of the game was to be considered). Reactive strategies are just a constrained version where $p_1 = p_3$ and $p_2 = p_1$.

The first action of the strategy (when the history does not exist yet) is assumed to be **C** unless is stated otherwise. For example, a strategy called **Suspicious Tit for Tat**, studied in [35], has the same representation as Tit for Tat but plays **D** in the first round.

In 1993 [38], an interesting memory-one strategy with the tolerance of Generous Tit for Tat but the capability of resisting and invading an all-out cooperators population was introduced. The strategy is called **Pavlov**, and is based on the fundamental behavioural mechanism win-stay, lose-shift. The strategy starts off with a **C**, then Pavlov will repeat it's last move it was awarder with by R or T but will shift if punished by P or S.

Several other strategies were introduced as more generous versions and described as more dominant than Tit for Tat. These include, **Contrite Tit for Tat** [49] and **Adaptive Tit for Tat** [46]. On the other hand, defector variants have also been studied [22]. **Anti Tit for Tat**, is a strategy that plays the opposite of the opponents previous move. Another limitation of the strategy was discussed in [44]. Tit for Tat was proven to hit a deadlock. Deadlock meaning a loop between cooperation and defection. **Omega Tit For Tat** was introduced and was a strategy capable of avoiding the deadlock [44].

Other strategies that made an impact have been **Gradual** [13] and **Handshake** [41] presented in 1997 and 1989 respectively. Gradual starts off by cooperating, then after the first defection of the other player, it defects one time and cooperates twice. After the second defection of the opponent, it defects two times and cooperates twice. After the n th defection it reacts with n consecutive defections and then two cooperations. Handshake is a strategy that starts with cooperation, defection. If the opponent plays in a similar way then it will cooperate forever, otherwise it will defect forever.

In 2011 [27] performed their own tournament where several interesting strategies made an appearance.

- **Periodic player CCD**, plays **C**, **C**, **D** periodically. Note that variations of a period player also make appearance in the article but will not be listed here.

- **Prober**, starts with the pattern **D**, **C**, **C** and then defects if the opponent has cooperated in the second and third move; otherwise, it play as Tit for Tat.

- **Reverse Pavlov**, a strategy that does the reverse of Pavlov.

In earlier work the same author introduced a strategy called **APavlov**, which stands for adaptive Pavlov [26]. 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 [6] two new strategies are presented. These strategies have been trained using a finite state machine representation. They are called, **Fortress3** and **Fortress4**. Figure 3 illustrates their diagrammatic representation.
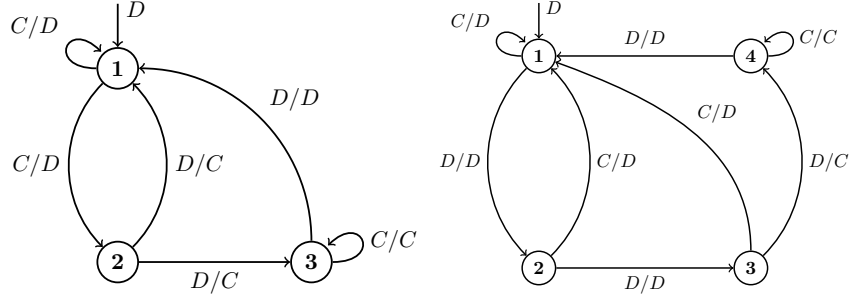

Figure 3: Representations of Fortress 3 and Fortress 4. Transition arrows are labelled O/P where O is the opponents last action and P is the players response. Note that the strategys first move, enters state 1, is defection for both strategies.

Finite state machines are a common mean of representing iterated prisoner's dilemma strategies [33, 42]. Other representation methods include lookup tables [9, 28] and artificial neural networks [21, 25]. A summary of several representation of the iterated prisoner's dilemma strategies is given in [20], the different representations are referred to as archetypes.

2.5 Strategies Stability

A strategy's robustness is tested through out the interactions and the performance of the strategy in a tournament against other strategies. Is the overall success of a strategy based only on it's performance in a round robin tournament, or should the strategy's performances be checked through other ways as well.

Following his initial tournaments Axelrod performed an 'ecological' tournament in 1981 [12]. The ability of strategies to be favoured under natural selection and their ability to withstand invasion from other strategies soon became an new measure of performance.

The ecological approach is based on the payoff matrix of the tournament. The highest performing strategies are adapted by lower scoring individuals within a fixed population. Over time a strategy takes over the population.
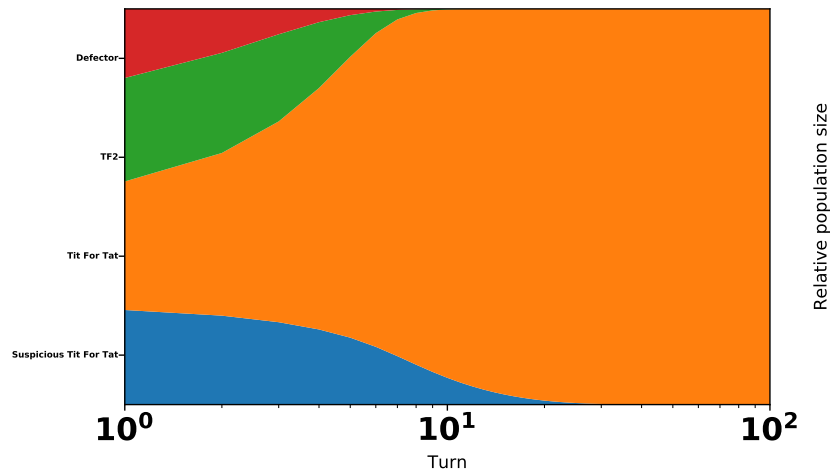

Figure 4: System evolving over time based on natural selection using [5].

Axelrod and his reciprocal claims were put on the test once again. The results of [15] argued that no pure strategy is evolutionary stable in the iterated prisoner's dilemma. This was not proven analytically, instead a series of examples using strategies such as Tit for Tat, Suspicious Tit for Tat and All Defector where explored; a very constrained set of strategies.

The results were questioned by [31], they stated that much was still no fully explored and more research had to be put into the results. Another attempt to explore stability of strategies in the prisoner's dilemma was done in [14]. This time exploring the results in a noisy environment, but similarly a analytical prove was not given.

An extension to the natural selection was introduced in the 1992 [35], recommending a different type of topology. A population of two deterministic strategies, AllD and AllC, are placed on a a two dimensional square array where the individuals can now interact only with the immediate neighbour. In this work it was proven that cooperation will emerge if individuals interact in neighbourhoods. The new topology has been called every since, spatial topology in the literature.

Nowak and moran process.

2.6   Software

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

The earliest source code that can be found, to the authors knowledge, is that of Axelrod's second tournament [8]. The code has been written by Axelrod and several other contributors in the programming language Fortran [1]. The site specifically says that the course code of the initial tournament is not available.

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

More recent projects include [2, 3]. Note that [2] is a user interface project that introduces several concept of the research. Such as the iterated game, computer tournaments and evolutionary dynamics. The list of strategies included is very limited. On the other hand, [3] offers a bigger collection of strategies. The users are also able to introduce the concept of noise in their tournaments. Even so, it remains a user interface project and makes it constrained for a research tool.

In 2015 an open source library, called the Axelrod project was introduced [5]. 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 [20] and a paper describing it and it's capabilities was published in 2016 [24].

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

In [40], 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 more recent work discuss the issues raised in here on the importance of research software and questions several results of the initial tournaments.

2.7   Ecological Applications

The reciprocal period of the prisoner's dilemma spread the knowledge of the game not only worldwide but also across different scientific principles. The study of cooperation was once again a critical issue. The applications of the game soon found their way to ecological studies, for example [32] conducted an experiment using sticklebacks to test the robustness of the strategy Tit for Tat in the interactions of fish. Fish usually travel in pairs and monitor their hunters to gain information on the enemy. Other works that include applications to ecological settings have been those of [19, 48]. There the reciprocal food sharing between vampire bats was studied.

2.8 not sure yet

In his works [11] try to address the criticism by studying Tit for Tat in an evolutionary manner as well. It was shown that Tit For Tat does not perform as well in noisy and in environments with mis-perception, but there are variants of Tit for Tat that do.

3 Analysis

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[16] 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.

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

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

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

[20] 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.

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

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

[23] 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.

[24] 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.

[25] 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.

[26] Jiawei Li. How to Design a Strategy to Win an IPD Tournament. In *The Iterated Prisoners' Dilemma 20 Years On*, World Scientific Book Chapters, chapter 4, pages 89–104. World Scientific Publishing Co. Pte. Ltd., 04 2007.

[27] Jiawei Li, Philip Hingston, Senior Member, and Graham Kendall. Engineering Design of Strategies for Winning Iterated Prisoner ' s Dilemma Competitions. 3(4):348–360, 2011.

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

[29] 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.

[30] 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.

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

[32] M. Milinski. TIT FOR TAT in sticklebacks and the evolution of cooperation. *Nature*, 325:433–435, January 1987.

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

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

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

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

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

[38] 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.

[39] 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.

[40] 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.

[41] A.J. Robson. Efficiency In Evolutionary Games: Darwin, Nash And Secret Handshake. 1989.

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

[43] 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.

[44] 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.

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

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

[47] J Von Neumann and O Morgenstern. Theory of Games and Economic Behavior. *Princeton University Press*, page 625, 1944.

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

[49] 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.