Literature review paper for the iterated prisoner's dilemma.

N. 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 of game theory 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 (**C**) and defection (**D**).

The fitness of each player is influenced by its own behaviour, and the behaviour of the participants. Players will receive a **reward** if both of them choose **C**, even so each of the players has a **temptation** to deviate and defect. Due to rational behaviour and the knowledge that an individual is tempted, the game's equilibrium lies at a mutual defection and players receive a **punishment** payoff.

The game's payoff are defined by,

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

where,

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

and

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

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 [8] to test the applicability of games described in [24]. In [8] 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 [23].

In the following section several milestones on the research of the prisoner's dilemma are presented and discussed.

2 Timeline

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 [19] conducted experiments using humans playing the game. This was not only done by Prof Rapoport but several other researchers perform similar experiments [7, 9, 11, 12, 22]. This is done until today (ask Martina for a reference?).

These early experiments explored the conditions under which altruist behaviour emerges. Furthermore, researchers were in search of a dominant way to play the game. Inspired by the work of Rapoport and the idea that AI was trained to play the game of chess [4], the political scientist R. Axelrod performed the first ever computer tournament of the iterated prisoner's dilemma [6].

In 1980 [2], the first computer tournament of the 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 opponents previous move. his is illustrated diagrammatically in Figure 1. To further test the robustness of the results a second tournament was performed later with a total of 63 strategies [3]. 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 [5]. 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.

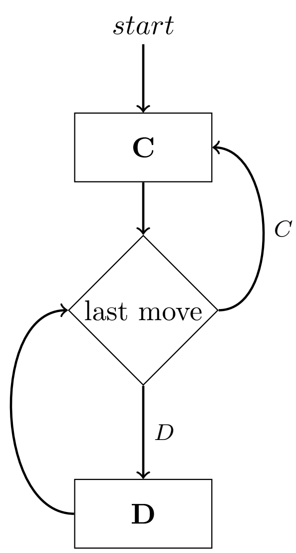

Figure 1: Diagrammatic representation of Tit for Tat.

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 [6], 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.

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 are unprotected and often suffer from a measure of uncertainty. In the original tournaments there was no possibility of misimplementation or misunderstanding. These are implemented within the iterated prisoner's dilemma as noise and mis-perception. The performance of Tit for Tat was proven to suffer from such stochasticity in the tournament environment, especially against itself [1, 10, 14, 15].

An interesting result was introduced by [14]. 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 [15], they used evolutionary dynamics and showed that to cope with noise a more generous version of Tit for Tat is needed. 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 [16]. Memory one strategies, are a set of strategies that consider only the last turn of the game to decide on the next action [17]. 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 $p_2 = p_1$. A few examples of strategies that have been discussed can be in their reactive representation are the following,

- Tit for Tat - $(p_1 = 1, p_2 = 0)$,
- Generous Tit for Tat - $(p_1 = 1, p_2 = \frac{1}{3})$.

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 [15], has the same representation as Tit for Tat but plays **D** in the first round.

In 1993 [18], 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.

A memory one representation of strategies are the following,

- Tit for Tat - $(p_1 = 1, p_2 = 0, p_3 = 1, p_4 = 0)$ or for simplicity $(1, 0, 1, 0)$,
- Pavlov - $(1, 0, 0, 1)$.

A diagrammatic representation of Pavlov is given in Figure 2, using a 2 state finite machines. Finite state machines are a common mean of representing iterated prisoner's dilemma strategies [21, 13].

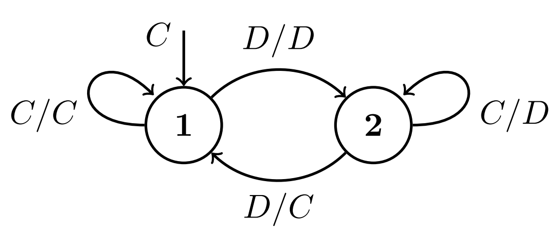

Figure 2: Diagrammatic representation of Pavlov using a finite state machine.

Note that the transition arrows are labeled O/P where O is the opponents last action and P is the players response. The initial move of the strategy, enters state 1, is **C**. An example, of how several strategies can be represented in a similar manner is given by Figure 3, illustrating the strategy Tit for Tat.

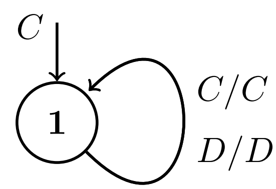

Figure 3: Diagrammatic representation of Tit for Tat using a finite state machine.

In his works [5] 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.

In [20], they claim the re-run the first tournament. They altered aspects such as, the format of the tournament, the objective and the population. They claim one of the authors was also a contributor to the first tournaments. Source code?

References

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

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

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

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

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

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

[7] 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.

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

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

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

[11] 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.

[12] 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.

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

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

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

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

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

[18] 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.

[19] 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.

[20] 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.

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

[22] 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.

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

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