

# Power of memory. In interactions both social and biological, is longer memory advantageous?

December 27, 2017

In interactions both social and biological how does cooperation emerges? Why cells sacrifice themselves and why humans behave in an altruist manner? In game theory, the prisoner's dilemma ([https://en.wikipedia.org/wiki/Prisoner%27s\\_dilemma](https://en.wikipedia.org/wiki/Prisoner%27s_dilemma)) has been used for decades to explain the emergence of altruistic behaviour.

The prisoner's dilemma is a two players non cooperative game. Both players can choose to Cooperate or Defect with each other. Both players are better off choosing Cooperation (3), even so here is always a temptation for a player to Defect (5). This is described by the game matrix given below,

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \left( \begin{array}{cc} (3, 3) & (5, 0) \\ (0, 5) & (1, 1) \end{array} \right) \end{array} \quad (1)$$

In the 1980's, a political scientist called Robert Axelrod carried out a computer tournament (<https://science.sciencemag.org/content/211/4489/1390>) of the iterated prisoner's dilemma. In the iterated version of the game the players interact for an infinite number of time and they have access to the full history of the matches. Axelrod's results showed that defecting strategies performed very poorly in the tournament. On the other hand, strategies that begun a match with a cooperation and where forgiving ranked on the top. Axelrod's findings were from the earliest that explain how cooperation can be evolutionarily advantageous.

In 2012 Press and Dyson(<https://www.pnas.org/content/109/26/10409.abstract>) studied the iterated prisoner's dilemma in a similar manner with Axelrod. Although their results were different. They suggested that the best performing strategies were selfish ones that led to extortion, not cooperation. These strategies by Press and Dyson are called **zero determinant strategies**.

They stated that in a two players interaction, a player with the shortest memory in effect sets the rules of the game. A player with a good memory-one strategy can force the game to be played, effectively, as memory-one. Thus a memory one player cannot be undone by another player's longer-memory strategy. Thus, in the iterated prisoner's dilemma, memory is not advantageous.

The purpose of this work is to consider a given memory one strategy  $q = (q_1, q_2, q_3, q_4)$ , (in a similar fashion to [?]). However whilst [?] found a way for the opponent of  $q$  to manipulate  $q$ , this work will consider an optimisation approach to identify the best response  $p^* = (p_1, p_2, p_3, p_4)$  to a strategy  $q$ . In essence answering the question: what is the best memory one strategy against a given other memory one strategy.

This approach will then be expanded on to consider multiple players  $q^{(1)}, q^{(2)}, \dots, q^{(N)}$  which in turn allows for the optimisation of a given memory one strategy in a group of memory one strategies.

But what are memory one strategies? Memory one strategies consider one the previous turn in order to make a decision on their next action. Memory one strategies were introduced by M. Nowak in 1990(<https://rd.springer.com/article/10.1007%2FBF00049570>).

					<span style="color: orange;">memory-1</span>	
	<b>1</b>	<b>2</b>	<b>3</b>	$\cdots$	<b><math>n-1</math></b>	<b><math>n</math></b>
player $p$	$C$	$C$	$C$	$\cdots$	$D$	$\cdots$
player $q$	$C$	$C$	$D$	$\cdots$	$D$	$\cdots$
					<span style="color: blue;">memory size-<math>m</math></span>	

Depending on the simultaneous moves of two players the states of the game, when only the previous round is considered, a state where both cooperated, both defected or either of them defected. These states are represented as  $CC, CD, DC, DD$ . A memory one strategy can be written as the probability of cooperating after each of these states. Thus as a vector of four probabilities  $p$  where  $p = (p_1, p_2, p_3, p_4) \in \mathbb{R}_{[0,1]}^4$ .

The above formulation offered a new framework of studying strategies. Consider that two memory one strategies are in a game of the prisoner's dilemma. Their interaction can be written as the following markov chain,

$$M = \begin{bmatrix} p_1 q_1 & p_1(-q_1 + 1) & q_1(-p_1 + 1) & (-p_1 + 1)(-q_1 + 1) \\ p_2 q_3 & p_2(-q_3 + 1) & q_3(-p_2 + 1) & (-p_2 + 1)(-q_3 + 1) \\ p_3 q_2 & p_3(-q_2 + 1) & q_2(-p_3 + 1) & (-p_3 + 1)(-q_2 + 1) \\ p_4 q_4 & p_4(-q_4 + 1) & q_4(-p_4 + 1) & (-p_4 + 1)(-q_4 + 1) \end{bmatrix}$$

where the opponent is denoted as  $q = (q_1, q_2, q_3, q_4) \in \mathbb{R}_{[0,1]}^4$ . The expected state that two opponents will end up can be estimated by calculating the steady states of the markov chain.

The players are assumed to move from each state until the system reaches a state steady. There after, the scores for each player can be retrieved by multiplying the steady states with the payoffs matrix. Thus, the utility for player  $p$  against  $q$ , denoted as  $u_q(p)$ , is defined by equation (??).

$$u_q(p) = v \times S_p \quad (2)$$

where  $v$  denotes the stationary vector of  $M$  and  $S_p$  the payoffs of player  $p$ . Subsequently, the steady states are defined and equation (??) is re written as:

$$u_q(p) = \frac{N}{D} \quad (3)$$

where,

The first theoretical result of this work shows that  $u_q(p)$  is a ratio of quadratic forms, Lemma ??.

Defining a memory one strategy  $p = (p_1, p_2, p_3, p_4)$ , the utility of the player  $u_q(p)$  can be re written as a ratio of two quadratic forms:

$$u_q(p) = \frac{\frac{1}{2}p^T Q + c^T p + a}{\frac{1}{2}p^T \bar{Q} + \bar{c}^T p + \bar{a}} \quad (4)$$

where  $Q, \bar{Q}$  are matrices of  $4 \times 4$  defined with the transition probabilities of the opponent's transition probabilities  $q_1, q_2, q_3, q_4$ .