Cooperative Equilibria in Iterated Social Dilemmas

Valerio Capraro[1], Matteo Venanzi[2], Maria Polukarov[2], and Nicholas R. Jennings[2]

[1] Mathematics, University of Southampton, United Kingdom
V.Capraro@soton.ac.uk
[2] Electronics and Computer Science, University of Southampton, United Kingdom
{mv1g10,mp3,nrj}@ecs.soton.ac.uk

Abstract. The implausibility of the extreme rationality assumptions of Nash equilibrium has been attested by numerous experimental studies with human players. In particular, the fundamental social dilemmas such as the Traveler's dilemma, the Prisoner's dilemma, and the Public Goods game demonstrate high rates of deviation from the unique Nash equilibrium, dependent on the game parameters or the environment in which the game is played. These results inspired several attempts to develop suitable solution concepts to more accurately explain human behaviour. In this line, the recently proposed notion of cooperative equilibrium [5, 6], based on the idea that players have a natural attitude to cooperation, has shown promising results for single-shot games. In this paper, we extend this approach to iterated settings. Specifically, we define the Iterated Cooperative Equilibrium (ICE) and show it makes statistically precise predictions of population average behaviour in the aforementioned domains. Importantly, the definition of ICE does not involve any free parameters, and so it is fully predictive.

1 Introduction

The standard assumption of economic models that players in strategic situations act perfectly rationally has been constantly rejected by numerous experiments over the years. These experiments, typically conducted on the fundamental social dilemmas such as the Prisoner's dilemma, the Traveler's dilemma, and the Public Goods game, have shown that cooperation between players (associated with the deviation from the unique, but inefficient, Nash equilibrium) is frequent, and appears to depend on both the game parameters and the environment in which the game is played. In particular, it has been observed that the rate of cooperation in the Traveler's dilemma depends on the bonus/penalty value, whenever the game is single-shot or iterated [7, 12]; the rate of cooperation in the Prisoner's dilemma depends on the payoff parameters or the way the players are matched to play together [11, 32]; and the rate of cooperation in the Public Goods game depends on the marginal return or on the frequency of interaction between free-riders and cooperators [13, 14, 17].

Considerable research efforts have been made in attempt to explain deviations from Nash equilibria. Some methods developed to this end are based on the idea that humans have bounded rationality and/or can make mistakes in computations[1] [4, 9, 20, 25]; others explain cooperation in terms of evolution [1, 3, 10, 19, 21–23, 29]. Finally,

[1] See [31] for a recent parallelism among these approaches.

B. Vöcking (Ed.): SAGT 2013, LNCS 8146, pp. 146–158, 2013.

much of work has been directed towards defining profoundly different solution concepts [24, 26], especially in the recent algorithmic game theory and artificial intelligence communities [2, 8, 15, 16, 18, 27, 30]. This interest is particularly motivated by the emerging applications of human-agent collectives, where artificial agents interact with humans. To build such systems effectively, it is highly important to understand and find accurate methods to predict human behaviour.

To this end, a new solution concept, termed *cooperative equilibrium*, has been recently proposed for one-shot games [5, 6]. This approach is inspired by the aforementioned experimental findings, which suggest that players are conditionally cooperative—that is, the same player may act more or less cooperatively in the same game scenario, depending on the actual payoffs. In other words, humans have an attitude to cooperation by nature: they do not act a priori as single players, but rather forecast how the game would have been played if they formed coalitions and then select actions according to their best forecast. It turns out, that direct implementation of this idea can predict human behaviour with impressively high precision, as demonstrated in [5, 6] on the aforementioned social dilemmas.

In this paper, we further explore this direction and extend the cooperative equilibrium approach to iterated settings. Specifically, we define the Iterated Cooperative Equilibrium (ICE), that combines this concept with some ideas developed in [7] for iterated games. Importantly, in contrast to other methods, ICE does not use any free parameters, and thus is fully predictive. We then evaluate our method on the iterated Traveler's dilemma, the Prisoner's dilemma, and the Public Goods game. To this end, we make use of the experimental data provided in [7], [32] and [14] for these three domains, respectively.[2] Our results confirm that the ICE makes accurate predictions of population average behaviour in social dilemmas. In particular, it clearly outperforms the Logit Learning Model (LLM) developed in [7] for the Traveler's dilemma.

The paper unfolds as follows. In Section 2 we define the social dilemmas in consideration. In Section 3 we formalise our approach. We then apply it to the iterative Traveler's dilemma in Section 4, to the Prisoner's dilemma in Section 5, and to the Public Goods game in Section 6. Section 7 concludes with directions for future work.

2 Preliminaries

We start with the definitions of the social dilemmas in consideration of this paper.

Prisoner's Dilemma (PD). Two players choose to either cooperate (C) or defect (D). If both players cooperate, each receives the monetary reward, R, for cooperating. If one player defects and the other cooperates, then the defector receives the temptation payoff, T, while the other receives the sucker payoff, S. If both players defect, they both receive the punishment payoff, P. Payoffs are subjected to the condition $T > R > P > S$.

Traveler's Dilemma (TD). Two travelers need to claim for a reimbursement between L and H monetary units for their (identical) luggage that has been lost by the same air company. To avoid high claims, the air company employs the following rule: the traveler who makes a lower claim, say m, gets a reimbursement of $m + b$ monetary units, and the other one gets a reimbursement of $m - b$ monetary units, for a fixed

[2] These were the only sources we could find that reported sufficient data for our purposes.