No winning strategy in the Iterated Prisoner's Dilemma: Game Theory and Simulated Evolution

Julian García
Faculty of Information Technology
Monash University
Melbourne, Australia
julian.garcia@monash.edu

Matthijs van Veelen
Department of Economics
Universiteit van Amsterdam
Amsterdam, The Netherlands
c.m.vanveelen@uva.nl

ABSTRACT

The iterated prisoner's dilemma (IPD) is a test bed for adaptation and cooperation. Computational experiments are regularly used for studying the competition of IPD strategies in multi-agent settings. This experimental work rarely links their results to game theoretical results with the potential to enlighten the analysis and the questions being asked. Here we focus on simulated evolution and results from Evolutionary Game Theory (EGT) and the IPD. The theory implies that all Nash equilibria can be upset by a sequence of mutants. If strategies are not restricted, populations of agents should move between Nash equilibria with different levels of cooperation. We argue this instability is inescapable, regardless of how strategies are represented. We present algorithms that show that simulated evolution perfectly aligns with EGT predictions. This implies that cognition itself may only have limited impact on the cycling dynamics of cooperation and defection. We argue that the role of mutations or exploration is more important in determining levels of cooperation.

1 INTRODUCTION

Humans are an exceptionally cooperative species, and this capacity for coordination and cooperation is an essential part of what makes us intelligent [11]. It is therefore not surprising that research in artificial intelligence has taken an interest in understanding how to promote and maintain cooperation in groups of self-interested agents [4, 13].

Cooperation means paying a cost in order to help somebody else. This implies selfish agents have no incentive to behave cooperatively, in spite of everyone being better off in cooperative groups. This tension between individuals and groups is best captured by the prisoner's dilemma (PD) [12].

In the simplest version of the PD, two players are given the option to cooperate or defect. Cooperators pay a cost c to bestow a benefit b on the opponent. Defectors pay no cost and provide no help, but may still benefit from others cooperating with them. The dilemma arises because, given $b > c > 0$, defection is the only dominant strategy but not Pareto optimal.

One way out of this dilemma is to allow players to interact repeatedly, thereby providing them with the chance to reciprocate cooperative acts. Thus, a strategy is a map from histories of the game to actions. The most well-known strategy is "Tit-for-Tat" (`TFT`), which cooperates on the first round and repeats the previous move of its opponent thereafter. The simplest strategies are `ALLC`, always cooperate; and `ALLD`, always defect. While many strategies can be named and described, the total number of strategies in the game is uncountably infinite [6].

The game is most interesting when players are uncertain about the number of rounds or repetitions. This is often captured with a continuation probability δ. For large δ, repeated interactions can lead to cooperative equilibria, but strategies that do not cooperate can also be strategically stable – this is known as the folk theorem[5].

With infinite possibilities for equilibria, a natural question is whether (or when) some equilibria are more stable than others. A reasonable answer relies on a learning, or evolutionary process to try to answer this question. While `TFT` and `ALLD` are both equilibria, if a learning or adaptive process spends more time in `TFT` that is good news for cooperation.

EGT assumes the competition of strategies in a large population, where strategies that do well reproduce more or faster. A strategy is said to be Evolutionarily Stable (ESS) if no mutant in small quantities can invade. Likewise, a Neutrally Stable Strategy (NSS) is able to perform at least as good as – but not necessarily strictly better than – any mutant arising in small enough quantities [9]. ESS and NSS are sometimes able to single out equilibria, but they are unfortunately inconclusive in the IPD, which has no ESS and infinitely many NSS's [3]. In [6] we show that no NSS is more stable than others, and an evolving population should see cycles of defection and cooperation, with neutral mutants being the catalysts for these transitions, in what is known as indirect invasions [14]. This is shown without restricting the strategy space.

2 SUMMARY OF RESULTS

How does the game theory of the IPD relate to computational experiments? Here we summarise the results in [7], where we show that in the case of simulated evolution these results are perfectly aligned when considering a rich strategy set, and large but finite populations. We perform an evolutionary simulation with unrestricted finite state automata, similar to some others found in the literature (e.g., [2, 8, 10]). We present a method to analyse the results of computational experiments. This methodology involves:

- Addressing the noise inherent to the simulations.
- Verifying that prevalent strategies are Nash equilibria – we present an efficient algorithm to perform this verification.
- Counting indirect invasions via neutral mutants, accounting for identical strategies whose representation is different.

As expected in theory we observe cycles of cooperation and defection (1). This is a prevalent feature of many simulations, but we are also able to show that the reason behind the cycles is the

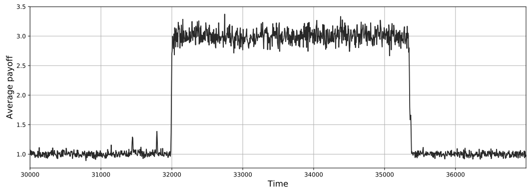

Figure 1: Typical cycles of cooperation and defection.

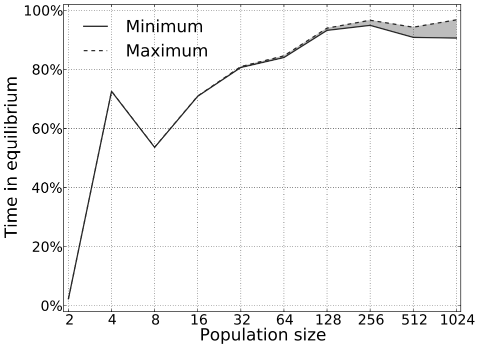

Figure 2: Prevalence of Nash equilibria.

prevalence of indirect invasions. We find that indirect invasions account for the vast majority of transitions between different levels of cooperation, specially when the population is large. This kind of transition dominates the dynamics, taking the population from one Nash equilibria to another (2), with varying different levels of cooperation.

Our work also shows that computational experiments are compatible with game theory analysis. Theory can in fact, facilitate meaningful data analysis of simulation results, but simulations can also help us push game theory forward. Our view is that more attention should be paid to how representations and exploration may affect cooperation. While cycles are unavoidable in repeated games, different exploration schemes—and strategy representations—may lead to more or less cooperation.

Our analysis shows that cycles are the norm when using the most general space of deterministic strategies. This implies that cognition itself may have minimal impact in changing the dynamics. The collapse of cooperation cannot be avoided with evolutionary learning, regardless of how sophisticated strategies are. Research should therefore focus on understanding the process of strategy exploration and implementation. Recent work in reinforcement learning may be a fruitful avenue to explore in population games [1, 4].

REFERENCES

[1] Nicolas Anastassacos and Mirco Musolesi. 2018. Learning through Probing: a decentralized reinforcement learning architecture for social dilemmas. (2018). arXiv:arXiv:1809.10007

[2] Robert M Axelrod. 1997. *The complexity of cooperation: Agent-based models of competition and collaboration.* Princeton University Press.

[3] J. Bendor and P. Swistak. 1995. Types of evolutionary stability and the problem of cooperation. *Proc. Natl. Acad. Sci. USA* 92 (1995), 3596–3600.

[4] Jakob Foerster, Richard Y Chen, Maruan Al-Shedivat, Shimon Whiteson, Pieter Abbeel, and Igor Mordatch. 2018. Learning with opponent-learning awareness. In *Proceedings of the 17th International Conference on Autonomous Agents and Multi-Agent Systems.* International Foundation for Autonomous Agents and Multiagent Systems, 122–130.

[5] D. Fudenberg and E. Maskin. 1986. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54, 3 (1986), 533–554.

[6] Julián García and Matthijs van Veelen. 2016. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. *Journal of Economic Theory* 161 (2016), 161–189.

[7] Julian Garcia and Matthijs van Veelen. 2018. No strategy can win in the repeated prisoner's dilemma: Linking game theory and computer simulations. *Frontiers in Robotics and AI* 5, 102 (2018).

[8] The Anh Han, Luis Moniz Pereira, Francisco C. Santos, and Tom Lenaerts. 2015. Emergence of cooperation via intention recognition, commitment and apology-A research summary. *AI Communications* 28, 4 (2015), 709–715. https://doi.org/10.3233/AIC-150672

[9] J. Maynard Smith. 1982. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge.

[10] J. Miller. 1996. The coevolution of automata in the repeated Prisoner's Dilemma. *Journal of Economic Behavior & Organization* 29, 1 (Jan. 1996), 87–112.

[11] D. G. Rand and M. A. Nowak. 2012. Human cooperation. *Trends in Cognitive Sciences* in press (2012).

[12] A. Rapoport and A. M. Chammah. 1965. *Prisoner's Dilemma.* University of Michigan Press, Ann Arbor.

[13] Yoav Shoham and Kevin Leyton-Brown. 2008. *Multiagent Systems: Algorithmic, Game-theoretic, and Logical Foundations.* Cambridge University Press.

[14] M. van Veelen. 2011. Robustness against indirect invasions. *Games and Economic Behavior* 74, 1 (2011), 382–393.