Social Trends in the Iterated Prisoner's Dilemma

Jessica Finocchiaro
Florida Southern College
Department of Computer Science
Lakeland, Florida 33801
jfinocchiaro@mocs.flsouthern.edu

H. David Mathias
Florida Southern College
Department of Computer Science
Lakeland, Florida 33801
hmathias@flsouthern.edu

ABSTRACT

In this paper, we utilize a multi-objective genetic algorithm (GA) to investigate the Iterated Prisoner's Dilemma problem with a population of players that don't have uniform objectives. Each of the members of our population has one of four objective pairs. We simulate a tournament similar to those in previous work to investigate patterns of convergence in objective pairs when they are free to change. We also consider the most successful objective pair within a population when members' objective pairs are fixed.

CCS CONCEPTS

•**Theory of computation** →*Algorithmic game theory;* •**Computing methodologies** →*Genetic algorithms;*

KEYWORDS

Prisoner's Dilemma, Multi-objective Genetic Algorithms

ACM Reference format:
Jessica Finocchiaro and H. David Mathias. 2017. Social Trends in the Iterated Prisoner's Dilemma. In *Proceedings of GECCO '17 Companion, Berlin, Germany, July 15-19, 2017,* 2 pages.
DOI: http://dx.doi.org/10.1145/3067695.3082037

1 INTRODUCTION

Academic ethicists often make the assumption that all people desire to be good. Evil occurs when there is a conflict of "good" decisions, when a good decision for one contradicts the good of another. In this case, a person makes a decision depending on the "good" they want to accomplish. In a society that starts with equal numbers of "selfishly good", "communally good", "cooperatively good", and "selflessly good" members, we investigate the convergence patterns and discover their causes through simulation of populations playing the Iterated Prisoner's Dilemma (IPD) over time. Additionally, we investigate the measures of success and which member of the population is most successful (in terms of its objectives) in Leave-One-Out testing against Axelrod's population [1].

The typical motivation for evolving cooperation in IPD is Mutually Assured Destruction, the Cold War policy. The motivation for our work lies instead in applications involving cooperation, specifically related to national policy. Consider environmental regulation. Some countries may be in a better position to take on the initial burdens of becoming more environmentally friendly. However, such countries may not want to be the first to make the sacrifices required, selflessly paying for the benefit of all. Parties in complex scenarios such as these may have differing goals. We find no work that has approached IPD using GAs with non-uniform objectives.

GECCO '17 Companion, Berlin, Germany
 978-1-4503-4939-0/17/07...$15.00
DOI: http://dx.doi.org/10.1145/3067695.3082037

2 PREVIOUS WORK

Our work is inspired by that of Mittal and Deb [7]. We use their genome and their probabilities of crossover and mutation, to allow comparison of results. Mittal and Deb [7] use a multiobjective GA to evolve an optimal strategy for maximizing personal score and minimizing opponent's score, running against Axelrod's population [1]. We use this framework, and NSGA-II [3], to expand on their results. Some previous results [2] describe attempts to apply Iterated Prisoner's Dilemma to real world problems, such as missile defense, but do so using only one objective. Coevolution [6] has also been used in an attempt to improve an evolved solution to IPD. Coevolution with random sampling has been applied to *Othello* [4]. A degree of dynamism has been used in allowing fitness functions to evolve [5]. While we keep fitness functions static, we allow members to change their objectives in some of our experiments.

P1 \ P2	Cooperate	Defect
Cooperate	3 : 3	0 : 5
Defect	5 : 0	1 : 1

Table 1: Scoring for the version of IPD in this work.

3 IMPLEMENTATION DETAILS

We simulate a population of 60 members, each having one of four objective pairs. Each pair is comprised of two of the following objectives: maximize personal score, minimize opponent's score, maximize opponent's score, and maximize cooperation. Some combinations of these objectives do not make sense, such as maximizing and minimizing opponent's score simultaneously. Due to such contradictions, we propose the objective pairs shown in Table 3.

The genome contains 70 bits. We use 0, 1 to represent cooperation and defection, respectively. Bits 0-63 determine the next decision based on the history of the past three moves. Six history bits, in positions 64-69, store the outcomes of the past three rounds in the form of the opponent's decision and personal decision for each. The next decision made by a player is determined by indexing into bits 0-63 based on the binary number represented by the six history bits.

Objective	Value
Max Own	Sum of own rewards
Min Opp	Additive inverse of sum of opponent rewards
Max Opp	Sum of opponent rewards
Max Co-op	Count of mutual co-op ∗ scaling factor

Table 2: Fitness values are divided by number of rounds played. The scaling factor for cooperation allows comparison with other objectives.

Name	Max Own	Min Opp	Max Opp	Max Co-op
Selfish	•	•		
Communal	•		•	
Cooperative	•			•
Selfless			•	•

Table 3: Objective pairs used in experiments reported.

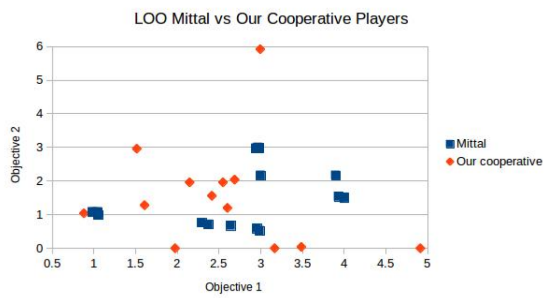

Figure 1: Our selfless LOO results vs Mittal's results per round.

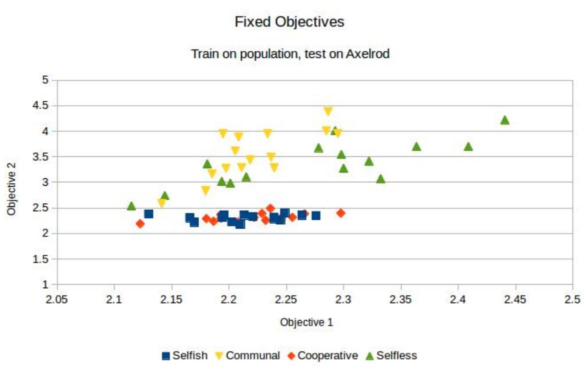

Figure 2: Distribution for objective pairs against Axelrod's population.

4 RESULTS

We first simulate our population playing IPD for 150 rounds against other members of the population. We repeat experiments both with fixed objective pairs to see which members of the population succeeded, as well as free objective pairs to see if there is a pattern of convergence to an objective pair. We are able to extract learned decisions by evaluating trends in repeated experiments. Figure 2 shows an example of the objectives of a population graphed against each other. In another set of experiments, we perform Leave-One-Out testing in which we train a population on 15 of the 16 members of Axelrod's population [1] and test our best player on the 16th member to evaluate robustness of the results. An example of the Leave-One-Out testing results is depicted in Figure 3.

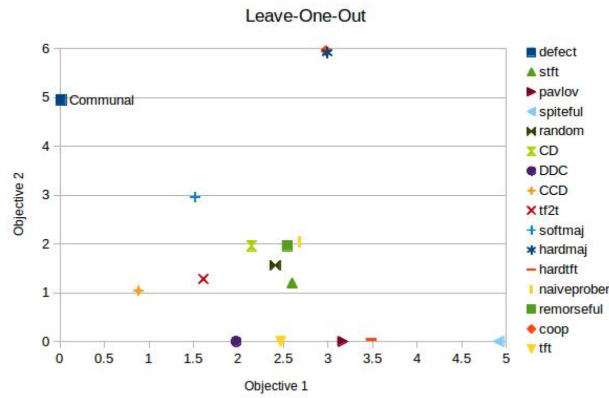

Figure 3: Scores for best of each LOO trial against Axelrod's 16th player. The best player in each trial doesn't always have the same objective pair.

5 CONCLUSIONS

In experiments in which members are permitted to change objectives, there is typically a convergence to one or two objective pairs that thrive. When two pairs flourish, they typically complement each other (*ex.* selfless and communally good players.) In Leave-One-Out testing, we conclude that members do learn specific patterns of behavior and that our selfless player is typically the most "successful," in terms of its own objectives, against the members of Axelrod's population. We are able to extract common patterns of decisions given a specific history of the past three moves. With developing foreign policy in mind, our model suggests that cooperating with other nations is attainable, and once cooperation starts to take place, other nations will follow suit.

REFERENCES

[1] Robert Axelrod. 1987. The evolution of strategies in the iterated prisoner's dilemma. *The dynamics of norms* (1987), 1–16.
[2] Paul Darwen and Xin Yao. 2002. Co-evolution in iterated prisoner's dilemma with intermediate levels of cooperation: Application to missile defense. *International Journal of Computational Intelligence and Applications* 2, 01 (2002), 83–107.
[3] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation* 6, 2 (2002), 182–197.
[4] Wojciech Jaśkowski, Pawe l Liskowski, Marcin Szubert, and Krzysztof Krawiec. 2013. Improving coevolution by random sampling. In *Proceedings of the 15th annual conference on Genetic and evolutionary computation*. ACM, 1141–1148.
[5] Jason D Lohn, Gary L Haith, Silvano P Colombano, and Dimitris Stassinopoulos. 1999. A comparison of dynamic fitness schedules for evolutionary design of amplifiers. In *Proceedings of the First NASA/DoD Workshop on Evolvable Hardware*. IEEE, 87–92.
[6] John H. Miller. 1993. *The Coevolution of Automata in the Repeated Prisoner's Dilemma*. Technical Report. Carnegie Mellon, Department of Decision Sciences. http://EconPapers.repec.org/RePEc:wop:carnds:_007
[7] Shashi Mittal and Kalyanmoy Deb. 2009. Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives. *IEEE Transactions on Evolutionary Computation* 13, 3 (2009), 554–565.