Edith Cowan University Research Online

ECU Publications Pre. 2011

1-1-2010

Evolving Group Strategies for IPD

Philip Hingston Edith Cowan University

Follow this and additional works at: https://ro.ecu.edu.au/ecuworks

Part of the Computer Sciences Commons

10.1109/CEC.2010.5586197

This is an Author's Accepted Manuscript of: Hingston, P. F. (2010). Evolving Group Strategies for IPD. Proceedings of IEEE Congress on Evolutionary Computation. (pp. 2414-2420). Barcelona International Convention Centre, Barcelona, Spain. IEEE. Available here

© 2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This Conference Proceeding is posted at Research Online.

https://ro.ecu.edu.au/ecuworks/6337

Evolving Group Strategies for IPD

Philip Hingston, Senior Member, IEEE

Abstract— The Iterated Prisoners Dilemma (IPD) is often used to model cooperation between self-interested agents. In an earlier study, we introduced a framework using IPD to study the effects of species-level competition on the evolution of cooperative behaviour. In this paper, we extend the previous work, using co-evolutionary simulations of interactions between species of IPD-playing agents to investigate how group strategies may evolve. We find that the ability to cooperate more with agents of the same species greatly increases the ferocity of competition between species.

I. INTRODUCTION

An important question in biology is how cooperative behavior can evolve in a population of selfish organisms. This question has long been studied using computers to simulate the evolution of agents playing the iterated prisoners dilemma (IPD), beginning, perhaps in 1981 with [1]. IPD is a model that encapsulates the choices an organism faces regarding whether or not to cooperate with another organism, and the payoffs resulting from the choices the two organisms jointly make.

Many hundreds of papers on the topic have examined it from many directions the effect of miscommunication, spatial models, multiple levels of cooperation, multiple players, choice of partners, signaling, selection schemes and so on. But in all these variations, the evolutionary process has been studied at the level of changing proportions of different alleles in a population of players - "microevolution". In an earlier paper ([9]), we proposed a framework for studying the evolution of cooperation at the level of competition between species, sometimes called "macroevolution", which includes phenomena such as speciation, mutualism (cooperation between species), parallel evolution, extinction and so on. In that paper we showed that group strategies can perform well in IPD, but we left open the question of how such strategies might evolve. This paper is a first step in answering that question.

In the rest of this paper, we first briefly introduce the iterated prisoners dilemma and the idea of group-aware strategies for playing it. We then review a simulation framework for simulating species-level evolution of IPD playing agents. In the following two sections, we describe and discuss experiments using this framework to study some scenarios with 0-order and first order group-aware strategies. We conclude with some suggestions for further work on evolving group-aware strategies.

II. BACKGROUND AND RELATED WORK

The Prisoners Dilemma (PD) is an abstract game used to study human and natural systems in which cooperation between self-interested individuals is observed or desired. It was introduced by Flood and Dresher in the early 1950s in studies applying game theory to global nuclear strategies [2]. It has also been applied to problems in psychology [3], economics [4], politics [5], and biology [6].

As PD is widely known, we describe it briefly here, and refer the reader to [7] for a more detailed description. In a game of PD, two players simultaneously choose one of two possible moves, C (cooperate) or D (defect). If both players choose C, they each get a payoff of R (reward). If one chooses C and one chooses D, the cooperator gets S (sucker) while the defector gets T (temptation). If both defect, they get P (punishment). For this work, we chose the common values T = 5, R = 3, P = 1, and S = 0. Defecting is the dominant strategy for PD (i.e. it is the best choice regardless of the opponent's choice). Thus two rational players will defect on each other, each getting a payoff of P. This is the dilemma, for if the players could somehow manage to cooperate, each would be better off, getting a payoff of R. Cooperation becomes viable when players play sequences of rounds of PD against each other, i.e. when they play the Iterated Prisoners Dilemma (IPD). To prevent players anticipating the last move in a game, the sequence continues with some fixed probability (the discount factor).

In IPD, player strategies are rules that determine a players next move following any possible sequence of previous moves. For example, TitForTat is this strategy : cooperate on the first move, and play the opponents previous move after that. Each players aim is to maximize his total payoff over the series of moves.

A. IPD Tournaments

Interest in IPD has been greatly enhanced over the years by IPD tournaments or competitions. Around 1980, Robert Axelrod staged two round-robin tournaments between computer programs designed by participants to play IPD. Many sophisticated programs were submitted. In each case, the winner was Anil Rapaport's submission, a program that simply played TitForTat. In 1987, Axelrod carried out computer simulations using a genetic algorithm to evolve populations of strategies playing the IPD against each other [1]. In these simulations, TitForTat-like strategies often arose, but other, more complicated strategies sometimes evolved that outperformed TitForTat in particular populations. Axelrod used this to illustrate that there is no "best" strategy for playing the IPD in such an evolving population, because

Philip Hingston is with the School of Computer Science and Security, Edith Cowan University, Australia (phone: +61 8 9370 6427; email: p.hingston@ecu.edu.au).

success depends on the mix of other strategies present in the population.

In 2004 at the IEEE Congress on Evolutionary Computation, a set of competitions was run to celebrate the 20th anniversary of Axelrod's 1984 book on the subject [10]. Competitors were allowed to enter more than one strategy, and some took advantage of this to enter group strategies – strategies in which players colluded with each other in order to improve the chances of one of their number winning the competition. The competitions were repeated at the 2005 IEEE Symposium on Computational Intelligence and Games (CIG05), along with an extra competition in which group strategies were not allowed. Many of the successful competitors described their entries in [7].

B. Group-aware strategies for IPD

Motivated partly by the success of the colluding strategies in the 2004 and 2005 competitions, in [9], we introduced a framework for studying how species could evolve the capacity to cooperate. In that initial work, we explored interactions between species of agents playing different fixed strategies some of which were group-aware - agents played differently depending on the species of their opponent. It was found that the dynamics of a multi-species co-evolutionary model are different from those of the traditional single-species model, and that in that context, even simple group-aware strategies can sometime outcompete more sophisticated group-unaware strategies. For example, the very simple Clique strategy (cooperate with your own species, defect on everyone else) can perform better than TitForTat. This is different from other studies such as [8], where cliques form within a single species, and individuals can refuse to interact with others not in their clique - a microevolutionary phenomenon.

This paper extends the work begun in [9], by allowing species to evolve their strategies while competing with each other. Thus, while the earlier study focussed on how certain fixed strategies interact with each other, it did not address the question of how such strategies could evolve in the first place.

Prior to [9], existing evolutionary simulations involving IPD operated at the level of evolving alleles in a single species population. There had also been earlier research involving the use of evolutionary simulatons at the species level (for example, [11], [12], [13]), but not for IPD.

III. FRAMEWORK

We explore the question of how group-aware IPD strategies may evolve using co-evolutionary simulations of evolving populations with several species of IPD playing agents. Specifically, we extended the framework introduced in [9], by allowing the agents to play differently depending on which species their current opponent belongs to. The ground rules are the same as in [9], namely:

- We simulate a population of IPD playing agents;
- The population consists of agents belonging to a number of distinct species;
- Each species has its own specific kind of genome;

- Agents reproduce asexually, with the child's genes derived from its parent's by mutation the child is the same species as its parent;
- An agent's genes determine the strategy the agent uses when playing IPD;
- An agent's fitness is determined as the average payoff it receives from playing IPD games against the rest of the population;
- An agent's reproductive success is proportional to its fitness.

We implemented this framework in Java, representing agents with the class *Organism*:

```
public class Organism
{
   public Genotype genotype;
   public Phenotype phenotype;
   public double fitness;
}
```

Thus, each organism has a genotype (peculiar to its species). The genotype determines the phenotype of the organism, which in this case embodies the strategy that the organism uses when playing IPD, and the fitness value is determined by the accumulated payoffs recieved by the organism using that strategy. In our Java realisation, *Genotype* is an interface:

```
public interface Genotype
{
   public Genotype copy();
   public void mutate();
   public Phenotype develop();
}
```

The intention of these methods is that:

copy() is used during "reproduction" to make a clone of a genotype.

mutate() is used to simulate mutation of the genome during reproduction and should modify the caller. And

develop() is used to simulate morphogenesis by creating a phenotype based on the information in the genotype.

The *Phenotype* interface is as shown below:

```
public interface Phenotype
{
    public Move getFirstMove(
        Class opponentClass);
    public Move getNextMove(
            Class opponentClass,
            Move oppLastMove);
}
```

, where

getFirstMove() is used to determine the player's first move in a game against a new opponent. The class of the opponent is provided to facilitate collusion. getNextMove() is used to determine a subsequent move by the player. The current opponent's previous move is passed, and it is up to the player to remember his own previous move, and the history of previous moves if desired. Again, the class of the opponent is provided to facilitate collusion.

Note that we chose to provide the agents with the species of their opponent (in our implementation, the class of their opponent's Genotype). An alternative would have been to require agents to infer the species of their opponent solely from their moves during a game of IPD. For example, some of the successful group-aware entries in the 2005 CEC and CIG IPD competitions used the first few moves in a game to recognise fellow conspirators. Because the technicalities of this kind of signalling are not our focus here, we decided to make the task of recognising friends and allies explicit and straightforward.

A species of IPD player is specified, in this framework, by a pair of Java classes implementing the Genotype and Phenotype interfaces. In our experiments, we simulate coevolutionary competition between various sets of species, using a procedure described by the pseudo-code below:

```
1. Create initial population of Organisms
2. While not done do
      For each pair of Organisms O1 and O2
3.
4.
        Play O1 against O2 in a game of IPD
5.
        01.fitness += 01's average payoff
        02.fitness += 02's average payoff
6.
7.
      Start a new population
      While new population not complete
8.
9.
        Select a parent O
10.
        C = 0.genotype.copy()
11.
        C.mutate()
12.
        P = C.develop()
13.
        Add a new Organism(C, P, 0.0)
           to the population
14
      End While
15. End While
```

We start, as described in the next section, with the simplest group-aware strategies we could think of, and progress to more complex strategies in the following section.

IV. 0-ORDER STRATEGIES

To establish a baseline for later experiments, we first consider species playing pure 0-order strategies. Although the framework allows for a strategy to specify different moves for players of any number of different species, for simplicity, we will only consider strategies that specify one move for players of ones own species, and one move for players of other species. Additionally, we restrict attention to strategies that play the same first and subsequent moves. Thus, there are four pure 0-order strategies to consider - let's call them C-C (cooperate with everyone), C-D (cooperate with our own species, defect against others), D-C (defect against our own species, cooperate with others) and D-D (defect against everyone). It is easy to see that D-D is the dominant strategy.

What happens in an evolutionary setting if an individual tries to cooperate? If any sort of cooperation is possible, the most promising idea seems to be to cooperate with fellow species members, and defect against the rest. Imagine then two species, A and B in competition, both playing D-D, and consider a mutant individual in species A, who plays C-D. Compared to D-D, this mutant suffers a penalty of 1 whenever he plays another member of A. However, every other member of A gets a benefit of 4. Therefore there is a nett benefit to species A. Unfortunately, the genes that produce this benefit will be selected against, as selection acts on the individual rather than the species as a whole! (Aside: in the 2005 CEC and CIG IPD competitions, the successful colluding entries used a kind of extreme version of this strategy, where everyone in the group played C-D except for one player, who played D-D. This worked well as the competition, unlike evolution, was one in which it was only important to have one very successful individual in the colluding group - the fate of the rest of the group was irrelevant.)

Hence, we have a new kind of dilemma: if a species could evolve to play C-D more often than a competing species playing D-D, it would have an evolutionary advantage (it would have a higher average fitness), but evolution acts against any mutation in that direction.

As a further illustration, imagine that species A somehow became full of C-D players, and consider a mutant D-Dplayer. Now the mutant gets a benefit of 2 every time he plays another member of A (5 instead of 3) while the species as a whole loses out because everyone else pays a penalty of 3 (0 instead of 3). Thus the mutation will spread within A (and the average fitness of A will decrease). A population of C-D players is an unstable configuration.

Given the analysis above, is there any way for a species to evolve whose members cooperate with each other, but defect against others? Looking for mixed strategies doesn't help – the same arguments seem to apply to any mutation that increases cooperation. We will see though that, as in the case of IPD without group-aware strategies, reasonable levels of cooperation can evolve if the players use higher order strategies (see Section V).

But before we leave this discussion on 0-order strategies, we present some simulation results that support our analysis.

A. Two species 0-order simulations

To verify our expectations about 0-order strategies, we created two species, *Group-aware 0-order Species 1* (*GA0-1*) and , *Group-aware 0-order Species 2* (*GA0-2*). These species are identical, except that individuals can be recognised as one species or the other. The genome is a pair of symbols, each either C or D. Possible genotypes are C-C, C-D, D-C and D-D, which are interpreted as above. Mutation switches each symbol with some fixed probability.

For the first simulations with these species, we began with a population of 50 *GA0-1* and 50 *GA0-2* players, each with randomly generated genotypes, and with a mutation probability of 0.05 per symbol (i.e. a 1 in 20 chance). We ran



Fig. 1. Evolution of one species in a simulation with two identical species of GA0 players. Each species starts with 50 randomly generated genotypes, and the mutation rate is 5%. After a few generations, the amount of cooperation (self and other) fluctuates between 0 and 20%, with mean about 5%. The top figure shows a typical run, while the bottom figure shows mean values over 20 runs.

the simulations for 1000 generations. The top plot in figure 1 shows the results of a typical simulation. In the figure, we show one of the species. Both species evolve towards minimal cooperation, with the residual cooperation being maintained by mutation. After an initial settling period of a few generations, the populations of the two species fluctuate around 50%, and the amount of cooperation varies between about 0 and 20%. The mean population for the species shown after generation 50 is 49.8, the mean proportion of cooperation when playing members of the same species is 5.86%, and when playing members of the other species is 5.23%. The other species is similar, with mean selfcooperation of 5.94% and other-cooperation of 5.37%. The bottom plot in the figure shows average values over 20 runs. Mean values over the 20 runs were: population = 49.98, selfcooperation = 5.49% and other-cooperation = 5.63%.

As an additional test, we ran further simulations with *GA0-1* and *GA0-2*, this time giving them different mutation rates. We saw in the first simulation that the mean proportion of

Fig. 2. Evolution of one species in a simulation with two identical species of GA0 players, with different mutation probabilities. The figure shows the species with the higher mutation probability (5%) - the other species has a mutation probability of 1%. As before, each species starts with 50 randomly generated genotypes, the top figure shows a typical run, while the bottom figure shows mean values over 20 runs.

cooperation is similar to the mutation probablity. This is approximately the case for other mutation rates also. What if, say, GA0-1 has a mutation probability of 5%, while GA0-2 has a mutation probability of 1%. In this case, a priori, we might expect GA0-1 to exhibit more cooperation than GA0-2. We have seen that the self-cooperation component of this should benefit the cooperator's brethren, even though harming the cooperator himself. What will the nett effect be?

Figure 2 shows results from a typical run and mean results over 20 runs, as in Figure 1. In the single run, the mean population for GA0-1 is 40, with mean self-cooperation of 6.89% and mean other-cooperation of 4.49%. This is confirmed by the averages over 20 runs: mean population is 37.64, mean self-cooperation is 6.78% and mean other-cooperation is 4.4%. For GA0-2, mean self-cooperation is 0.86% and mean other-cooperation is 1.37%.

This simulation shows that it is possible for a species to evolve preferential cooperation within a species (i.e. greater cooperation within the species than with individuals from other species). However, in our example with 0-order strategies, this is achieved at the cost of that species occupying a smaller proportion of the total population, when compared with less cooperative species.

Perhaps we should not be surprised that cooperation is hard to achieve with 0-order strategies. Although we have simulated games of IPD, as 0-order strategies have no memory, we are in fact simply repeating single-shot prisoner's dilemma. In normal simulations, i.e. simulations that do not involve group-aware strategies, it is necessary to introduce the iterated game and higher order strategies, such as tit-fortat, where defection can be punished in later rounds, in order to evolve cooperative strategies. This is what we explore next in the context of group-aware strategies.

V. 1-ORDER STRATEGIES

In this section, we examine pure first-order strategies, i.e. those where the agent's move in the current round of an IPD game depends on the moves made in the previous round. An example of such a strategy in the standard version of IPD is tit-for-tat, in which the agent plays whatever move was played by his opponent in the previous round. In our version of IPD, a 1-order strategy must specify what move the agent plays depending on the species of his opponent as well as on the previous moves. Once again, we restrict attention to strategies that only differentiate between agents of ones own species, and agents of other species. Therefore, they can be specified as a pair of 1-order strategies for standard IPD one that is used when playing against an opponent of the same species, and one that is used when playing against others. As there are 32 pure 1-order strategies for standard IPD, we are looking at 32×32 strategies. These can be denoted by two sequences of 5 C'd or D's. For example, CCDCD-CCDCD denotes the strategy of playing tit-for-tat against everyone. The first 5 symbols specify the first move, the move following a round where the agent played C and the opponent played D (CD), the move following DC, the move following CD, and the move following DD, when playing against one's own species, and the last 5 specify how to play against an opponent from another species.

In a similar way as for 0-order strategies, we created two species, GA1-1 and GA1-2, which are identical except that it is possible to tell which species an individual belongs to. Possible genotypes for both species are specified by two sequences as described above. When a mutation occurs, one of the 10 symbols is randomly selected to be switched.

1) One species 1-order simulations: First, we repeated an often-seen experiment by simulating a population of 100 agents of species GA1-1. The mutation probability was set to 0.05. Since there is only one species, only the first part of the genome is relevant, and the simulation is one where the agents are playing standard IPD. Figure 3 shows a sample of 5 runs, and the usual pattern of initial defection followed by cooperation is seen. (In contrast, although we have not presented any one species 0-order simulations, cooperation never evolves in those simulations.) What will change when we introduce a second species?



Fig. 3. Five sample runs with a single species of 1-order group-aware agents. As there is no other species, only the part of the strategy that determines play within the species is relevant, and we see the classic pattern of initial defection, followed by domination by cooperative strategies. The time taken to establish cooperation varies between about 15 generations and 35 generations.

2) Two species 1-order simulations: We now move on to the case of two species using first order strategies. We carry out two experiments. In the first, both species ignore the species of the opponent. The setup for the second is identical, except that both species are group-aware.

Two species 1-order but not group-aware

For this experiment, we created two new identical species, UGA1-1 and UGA1-2. The genome for these species is one half of the genome for the species GA1-x, specifying an initial move, and first order responses, ignoring the species of the opponent.

We started each simulation with 50 UGA1-1 and 50UGA1-2 agents, and ran for 1000 generations. We repeated the simulation 10 times. In each simulation, one species or the other eventually gains and maintains a numerical advantage over the other. In 5 out of 10 cases, the less successful species actually went extinct.

As in the previous experiment, both species initially defect but quickly evolve cooperation. The simulation then becomes a kind of random walk, in which the luckier species becomes more numerous. Figure 4 shows mean values of self- and other-cooperation for both species (which all follow a similar path on average), and the mean number of agents in the more numerous of the two species (which we might call the "winner" of the run).

Two species 1-order group-aware

This experiment is like the previous one, except that the agents are group-aware : we started each simulation with 50 GA1-1 and 50 GA1-2 agents, all with randomly generated genomes. The result was a surprise : in almost every simulation, one species or the other went extinct before the 200^{th} generation.

Figure 5 shows mean values as in Figure 4. As before, initially, both species tend towards defection. By about the



Fig. 4. Means for 10 runs with two species of 1-order group-unaware agents. The bold black line is the proportion of the population in the species that became more numerous (the 'winner'). The solid red line shows the proportion of self-cooperation for the winner, while the dotted red line shows the proportion of cooperating moves with agents of the other species. Similarly, the blue lines show the same proportions for the loser in the run.



Fig. 5. Means for 10 runs with two species of 1-order group-aware agents. The bold black line is the proportion of the population in the species that survived the run. The solid red line shows the proportion of self-cooperation for the surviving species, while the dotted red line shows the proportion of cooperating moves with agents of the other species. Similarly, the blue lines show the same proportions for the species that went extinct in the run.

10 generations point, both species start to cooperate more. It seems that the species that is more numerous by this point is able to evolve towards self-cooperation slightly faster than the less numerous one, and from that time onwards accelerates away, so that by about 40 generations, it accounts for 90% of the population and attains a self-cooperation of about 80%, and about 70% cooperation with members of the other species. The less successful species never gets above about 60% self-cooperation, and drops to about 10 - 20% before finally succumbing.

Thus we see that in this scenario, competition between species is so intense that one species or the other is quickly driven to extinction.

VI. CONCLUSION AND FUTURE WORK

In these experiments, we have extended our earlier work on species level evolution of cooperation by allowing the evolution of strategies that explicitly play differently against different species of player. We found that this can significantly alter the dynamics of evolution — in this case, intensifying inter-species competition.

With these examples, we have shown that the proposed framework is suitable for examining group effects in applications modelled using IPD. For example, in economics, individual agents might represent companies and species might represent the countries in which they are based. Would we expect that companies would develop a tendancy to favour other companies based in the same country in their business dealings? Would this be likely to be a successful strategy? Similar questions could be asked concerning, say academics (agents) and their interactions with each other in a university, be they in the same of different departments. The general question : can a group as a whole, and/or individuals in a group, gain an advantage by cooperating differentially with other individuals depending on the group they belong to?

In all our examples to date, agents only distinguish between their own species and other species. It would be straightforward to examine coalitions of, say, two species, acting against a third. Would there be any advantage to be gained in this way? Elaborations of the basic model might be used to explore some of the same IPD variations that have been proposed and studied previously, such as the effects of noise, different levels of cooperation, reputation, and so on.

REFERENCES

- [1] Axelrod, R. and Hamilton, W. The Evolution of Cooperation, Science, Vol 211, Issue 4489, 1390-1396, 1981
- [2] Flood, M. Some experimental games, Research Memorandum RM-789, RAND Corporation, Santa Monica, CA, 1952
- [3] Cox, T.H., Lobel S.A. and McLeod, P.L., Effects of Ethnic Group Cultural Differences on Cooperative and Competitive Behavior on a Group Task The Academy of Management Journal, Vol. 34, No. 4 (Dec., 1991), 827-847, 1991
- [4] Frank, R.H., Gilovich T. and Regan, D.T., Does Studying Economics Inhibit Cooperation? The Journal of Economic Perspectives, Vol. 7, No. 2 (Spring, 1993), 159-171, 1993.
- [5] Grieco, J.M., Realist Theory and the Problem of International Cooperation: Analysis with an Amended Prisoner's Dilemma Model The Journal of Politics, Vol. 50, No. 3 (Aug., 1988), 600-624, 1988
- [6] Brown, J.L., Cooperationa biologists dilemma, Advances in the Study of Behavior, Vol. 13, 1–37, Academic Press, 1983
- [7] Chong, S.Y., Humble, J., Kendall, G., Li, J. and Yao, X. Iterated Prisoner's Dilemma and Evolutionary Game Theory, in X. Yao, G. Kendall and S.Y.Chong Eds. The Iterated Prisoner's Dilemma: Celebrating the 20th Anniversary, World Scientific, 23-62, 2007
- [8] Hruschka,D., Henrich,J. Friendship, cliquishness, and the emergence of cooperation. Journal of Theoretical Biology, Vol. 239, 1-15, 2006
- [9] Hingston, P., Iterated Prisoner's Dilemma for Species, In: Proceedings of the IEEE Symposium on Computational Intelligence and Games, CIG'09, Milano, Italy, September 2009
- [10] Axelrod, R. The Evolution of Cooperation, BASIC Books, New York, 1984

- [11] Watson, A.J. and Lovelock, J.E. Biological homeostasis of the global environment: the parable of Daisyworld. Tellus B (International Meteorological Institute), Vol. 35, Issue 4, 286-9, 1983
- [12] Schmickl, T. and Crailsheim, K. Bubbleworld.Evo: Artificial Evolution of Behavioral Decisions in a Simulated Predator-Prey Ecosystem in From Animals to Animats 9, pp. 594-605, LNCS 4095, 2006
- [13] Rankin, D.J., Lopez-Sepulcre, Foster, K.R. and Kokko, H. Species-level selection reduces selfishness through competitive exclusion. Journal of Evolutionary Biology, Vol. 20, Issue 4, 1459-1468, 2007