Prisoner's Dilemma Agents with Phenotypic Plasticity

Daniel Ashlock, EunYoun Kim, and Amanda Saunders

Abstract—**This study compares an adaptive and a non-adaptive implementations of Prisoner's Dilemma playing agents. The adaptive agents implement three interlinked strategies and choose which strategy to use based on environmental cues, in this case the mean score of the agents in the previous generation. The hypothesis under test is that phenotypic plasticity can grant a competitive advantage to agents possessing it. The interlinked strategies are implemented as finite state machines with a thread for each environmental condition; the thread corresponding to the current environmental condition generates the agent's response, but all threads are updated throughout play. It is found that agents with phenotypic plasticity can be superior to agents without it but need not be. Three variations of phenotypic plasticity are studied. One outcompetes the control agents while the control agents outcompete the other two types of plastic agents. Two of the agents with phenotypic plasticity are found to exhibit enhanced levels of cooperation. Other possible implementations of phenotypic plasticity are discussed.**

I. Introduction

Prisoner's dilemma is a good test environment for many ideas about mathematical games and, in some cases, games in general. In this study we examine agents that are granted a more complex form of environmentally triggered adaptability called phenotypic plasticity. This term is defined in Section II. The triggers for use of the different adaptions available to the agents are fixed for a given agent type, but the agents are permitted to evolve their response to those triggers and even, potentially, to discard them.

Agents evolved with three versions of the additional adaptability are compared, for competitive ability, with one another and with control agents that lack the adaptability. The agent types are placed into competition and tested for cooperation with and dominance over other agent types. The results find that the added ability can be a boon or a handicap, depending on the trigger thresholds, demonstrating that the added adaptive ability is worth investigation, but is somewhat complex.

The remainder of this study is structured as follows. Section II surveys past work and defines phenotypic plasticity. Section III gives the experimental design. Section IV gives and discusses results, while Section V draws conclusions and outlines possible next steps.

Daniel Ashlock and Amanda Saunders are at the University of Guelph, in Ontario, Canada, {dashlock|asaunder}@uoguelph.ca

Eun-Youn Kim is at the Hanbat National University, School of Basic Sciences, South Korea, eunykim@hanbat.ac.kr

978-1-7281-1884-0/19/$31.00

II. Background

This section defines and gives background on the added adaptive ability and reviews past work on the evolution of prisoner's dilemma agents.

A. Phenotypic Plasticity

Phenotypic plasticity consists of any environmentally triggered change in the phenotype of an organism. A good survey of the biology appears in [33] and good overviews of recent developments and theories regarding phenotypic plasticity and its role in evolution can be found in [23], [29]. The phenomenon of plasticity was originally viewed as an annoying distraction in evolutionary theory [34] but has recently been recognized as a target quality for evolving adaptability, although the degree to which it accounts for evolution and speciation is still under discussion [23]. Instances thought to be rapid evolution in some species are, instead, environmental triggering of latent potentials that are a form of phenotypic plasticity.

Phenotypic plasticity can be controlled at the developmental level, by epigenetics [28], or via genetically determined systems responding to an environmental signal like drought or extreme temperatures. Examples of phenotypic characters in plants that can be plastic include the size of seeds and leaves, the shape of leaves, and the thickness of leaves. An example of regulatory plasticity in animals is temperature dependent sex selection in reptiles [31]. Many species of reptile do not have genetically determined sex, rather the temperature at which eggs are incubated determines sex. Diet can also affect phenotype. For example, the tadpoles of the Mexican spadefoot toad *Spea multiplicata* can be either large carnivores or small omnivores depending on whether they have ingested fairy shrimp [32].

The advantage of phenotypic plasticity is that a single chromosome can incorporate multiple phenotypes, beyond differences in phenotype caused by brute force effects in the environment where the organism develops. In this study we will make a small modification to a finite state game playing agent to permit it to encode three game strategies that are executed in the same evolving structure and couple this with an environmental sensor that permits the agent to change between the strategies. As the agent evolves, the hope is that these changes will become strategically appropriate.

In [26] the authors state the following:

> Phenotypic plasticity can be broadly defined as the ability of one genotype to produce more than one phenotype when exposed to different environments, as the modification of developmental events by the environment, or as the ability of an individual

organism to alter its phenotype in response to changes in environmental conditions. Not surprisingly, the study of phenotypic plasticity is innately interdisciplinary and encompasses aspects of behavior, development, ecology, evolution, genetics, genomics, and multiple physiological systems at various levels of biological organization.

In this study we devise a type, one of many possible, of phenotypic plasticity for simple agents. Lacking the biochemistry of living organisms, we cannot reasonably try to simulate the biological mechanisms of phenotypic plasticity. Rather, we are experimentally determining if and when plasticity might yield better game playing agents.

Many systems used to simulate biological phenomena use simple non-adaptive representations. In this study we take a representation that can already adapt to its opponent's moves and augment it with an additional layer of adaptive ability. The results of this study will show that incorporating adaptive ability, as a form of phenotypic plasticity, can grant a substantial advantage or disadvantage: the details of the plasticity matter. Given that actual biological organisms use and benefit from phenotypic plasticity [34], it seems natural to include it as a means of making game playing agents more effective.

This study uses a single representation, a finite state transducer, for its agents. This is one of the most studied and among the most effective types of game playing agents for the prisoner's dilemma, but this choice defers examination of the issue of representation for the future [6], [5].

B. Prisoner's Dilemma and Past Work

The prisoner's dilemma [17], [16] is a classic model in game theory. Two agents each decide, without communication, whether to cooperate (C) or defect (D). The agents receive individual payoffs depending on the actions taken. The payoffs used in this study are shown in Figure 1. The payoff for mutual cooperation C is the *cooperation* payoff. The payoff for mutual defection D is the *defection* payoff. The two asymmetric action payoffs S and T, are the *sucker* and *temptation* payoffs, respectively. In order for a two-player simultaneous game to be considered prisoner's dilemma, it must obey the following pair of inequalities:

$$S \leq D \leq C \leq T \tag{1}$$

and

$$2C \leq (S+T). \tag{2}$$

In the *iterated prisoner's dilemma* (IPD) the agents play many rounds of the prisoner's dilemma. IPD is widely used to model emergent cooperative behaviors in populations of selfishly acting agents and has been used to model systems in biology [37], sociology [25], psychology [36], and economics [24].

Previous work has shown that the introduction of noise substantially alters the course of evolution when prisoner's dilemma is studied in the context of an evolutionary algorithm or in a tournament. In particular, the introduction of noise has been shown to affect the evolution of cooperation and the complexity of winning strategies [21], [30], [38], [20]. It is also shown in [21] that more diverse populations are affected less by the introduction of noise. There are two possible types of noise in prisoner's dilemma: misunderstanding an opponent's action or mistakenly making an unintended play. This study investigates the impact of the first type.

$$\begin{array}{cc|cc}
 & & \mathcal{S} & \\
 & & C & D \\
\hline
\mathcal{P} & C & 3 & 5 \\
 & D & 0 & 1
\end{array}
\qquad
\begin{array}{cc|cc}
 & & \mathcal{S} & \\
 & & C & D \\
\hline
\mathcal{P} & C & C & T \\
 & D & S & D
\end{array}$$

(1) (2)

Fig. 1. (1) The payoff matrix for prisoner's dilemma used in this study – scores are earned by strategy $\mathcal{S}$ based on its actions and those of its opponent $\mathcal{P}$. (2) A payoff matrix of the general two player game – C, T, S, and D are the scores awarded.

The choice of representation is critical for prisoner's dilemma; the study in [9], continued in [6], investigates the impact of this choice. The representations covered by the two studies are two versions of feed forward neural nets (one biased at the neuron level toward cooperation), Boolean parse trees [18], with and without a one-step time-delay operation, a linear genetic programming representation called an ISAc list [13], lookup tables, a type of Markov chain [35], and both a direct and cellular [6] representation of finite state machines. The change of representation, with other factors held as near to constant as possible, yielded a change from 0% to 95% in the probability that final populations were cooperative. This study uses one of these representations, finite state machines, investigating the probability of cooperative behavior in the presence of noise. The duration of evolution in [9], [6] was 250 generations with samples taken at 50, 100, 150, 200, and 250 generations. Little effect on the cooperativeness of agents was observed at different epochs. This study demonstrates a different result for noisy strategies evolved for a longer time.

In [14] it was found that evolving agents to play the iterated prisoner's dilemma for a long time gave them a substantial competitive advantage against agents evolved for less time from different evolutionary lines. This phenomenon, called *non-local adaptation*, suggests that agents are gaining not just skill playing the agents with whom they are co-evolving but general skill at playing the prisoner's dilemma; they have a statistically significant advantage against agents they have never been evaluated against before. In [11] it was demonstrated that non-local adaptation takes place in a steady fashion across much of evolution. This study extends measurement of competitive advantage to include the effects of noise as well. Non-local adaptation is also observed in other co-evolutionary contexts. In [22] it was observed in competitive exclusion in a spatial model of plant growth. In [3] it was observed in populations of virtual robots evolving to paint a floor in two competing colors. The effect was observed in predator-prey models in [1] and in populations of agents in a game called *divide the dollar* in [4].

Other recent results on Prisoner's Dilemma include a demonstration that the strategies located change with the choice of payoff matrix [8]. The average competitive ability of strategies changes when they are evolved using different payoff matrices. The resources permitted an agent, the number of states in a finite state representation, the number of neurons in an artificial neural net, or the granularity of a Markov chain also influence which strategies arise and the competitive ability of those strategies [27]. Even the details of the evolutionary algorithm, such as elite fraction and population size, influence the type of agents that arise [7].

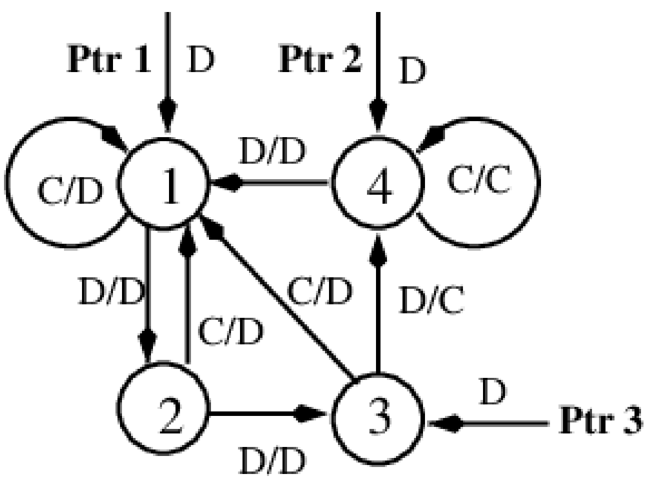


Fig. 2. An example of a modified FSR for playing IPD with three pointers to active states. One pointer is used to generate responses for each environmental condition.

III. Experimental Design

The representation used in this study is a simple modification of the finite state representation (FSR) used in [27], this is a Mealy architecture finite state transducers that encode an initial action and condition their transition to a new state and action on the opponent's last action. This representation is described in detail, below. Normally, this sort of FSR maintains a pointer to the currently active state. The modification is to have three such pointers, permitting the agent to run three threads simultaneously. All three threads are run for each input, and each generates a next action. We implement phenotypic plasticity by using a simple method of selecting which of the three actions, generated by the three threads, is actually used.

An example of a threaded agent appears in Figure 2. The isolated **D**s are the initial actions of the threads; each is on a sourceless arrow that points to the starting state of the thread. Transitions are shown with arrows that have labels with two prisoner's dilemma moves separated by a slash. The first of these is the opponent's last action, the second is the agent's response.

We use a pair of *plasticity thresholds* to partition the agent's environment into three zones. These thresholds are numbers, T_1 and T_2 so that $1 \leq T_1 \leq T_2 \leq 3$. The numbers 1 and 3 are the minimum and maximum average score possible in a round robin prisoner's dilemma tournament using the payoff matrix given in Section II. If the previous generations average score S_{avg} is below T_1, the action generated by the first thread is used; if S_{avg} is between the thresholds the action generated by the second thread is used; otherwise the action generated by the third thread is used.

TABLE I

Plasticity thresholds used for sets of runs in this study.

Experiment	T_1	T_2
1	2.25	2.8
2	1.2	2.25
3	1.2	2.8
4 (control)	1.00	1.00

The thresholds are fixed for a collection of experiments. Control agents are implemented by setting $T_1 = T_2 = 1$ so that only the third thread is ever used. Table I gives the values of the thresholds used. The value 2.25 is the average of all four payoffs and is what two players using random moves get against one another. The value 1.2 represents the boundary below which an agent is believed to be mostly defecting; likewise 2.8 is the lower bound for behavior termed mostly cooperative.

The first experiment has the ranges "worst than random", "better than random", and "cooperative" delineated by its plasticity thresholds. The second experiment uses the ranges "defecting", "better than defecting but worse that random" and, "better than random". The third experiment uses the ranges "defecting", "neither", and "cooperating". It is not clear what good choices of plasticity thresholds are and these three, together with the control agents, are sufficient to permit a first assessment.

A. Agent Representation

The agent representation used in this study is 24-state finite state machines with actions associated with transitions between states (Mealy machines). State transitions are driven by the opponent's last action. Access to state information permits the machine to condition its play on several of its opponent's previous moves, for each of the adaptive threads. The machines are stored as linear chromosomes listing the states. The initial action for each thread is stored with and undergo crossover with the first state in this linear chromosome. The initial states for each thread are states 0, 8, and 16, numbering the 24 states with zero based counting.

Two variation operators are employed, a binary variation operator and a unary variation operator. The binary variation operator used is two-point crossover on the list of states. Crossover treats states as atomic objects. The mutation operator changes a single state transition 40% of the time, the initial action 10% of the time, or an action associated with a transition 50% of the time. The unary variation operator replaces the current value of whatever it is changing with a valid value selected uniformly at random.

B. *The Evolutionary Training Algorithm*

Agent's fitness is evaluated through a round robin tournament lasting 150 rounds. The tournament length is borrowed from earlier studies, retained for consistence. The average score of the population in the last round is used to determine which thread each agent uses to generate its next action and, for the first generation, the "random behavior" value of 2.25 is used in place of the previous generation value.

The population is updated by an elitist method. The algorithm uses a population of 36 agents. The highest scoring 24 agents, breaking ties uniformly at random, are retained into the next generation. This structure of 36 agents with a 24 agent elite is chosen for consistency with several earlier studies, to permit comparison. The twelve lowest scoring agents are replaced by selecting pairs of parents, without replacement, from the agents to be retained. Parents are copied and the copies undergo two point crossover and a mutation. The mutation changes the agent's initial action 10% of the time, a state transition 40% of the time, and an action 50% of the time. This choice of reproductive algorithm and variation operators are also retained from earlier studies [8], [27] for consistency.

The agents are evolved for 250 generations and then the elite population is saved for evaluation for competitive ability. The population average and maximum fitness over the course of evolution for each population is also saved. Each experiment consists of thirty independent runs of the evolutionary algorithm, providing 30 final elite populations of 24 agents in each experiment.

C. *Competitive Evaluation*

Comparison is made of agents from all six pairs of experiments. In order to compare two experiments, the populations are loaded, one at a time, from each experiment. The first population competes against the first, the second against the second, and so on. Re-using populations would potentially cause bias if exceptional populations arose, so only thirty pairs are compared. In order to compare two populations, each pair of agents with one member in each population play 150 rounds of IPD, a binary round robin tournament. This is done 20 times, generating an average fitness for use with the plasticity thresholds. Each pair of agents, in each of the 20 iterations, is evaluated for win, lose, or draw.

If two agent's scores in 150 rounds of play differ by ten or less they *draw* and are judged to be coordinating the scores. Otherwise the status win goes to the higher scoring agent, lose to the lower scoring agent. The probability $p = win/(win + loss)$ is estimated using the normal approximation to the binomial distribution and reported. If $p = 0.5$ is not within a 99% confidence interval, then one of the experiments is judged to have generated competitively superior agents.

IV. Results and Discussion

The competition results appear in Table II. In all pairs of experiments, one or the other was competitively superior. The control run, which did not benefit from phenotypic plasticity, beat two of the sets of agents that possessed phenotypic plasticity. The agents in experiment one, however, were competitively superior to all others, beating the control agents with a higher probability than either of the other plastic agents. This suggests that, while the who-beat-who dominance pattern in this study was transitive, some sub-collections of agents might exhibit intransitive behavior. The subject of transitivity, or lack thereof, in agent behavior is itself one worth study [2].

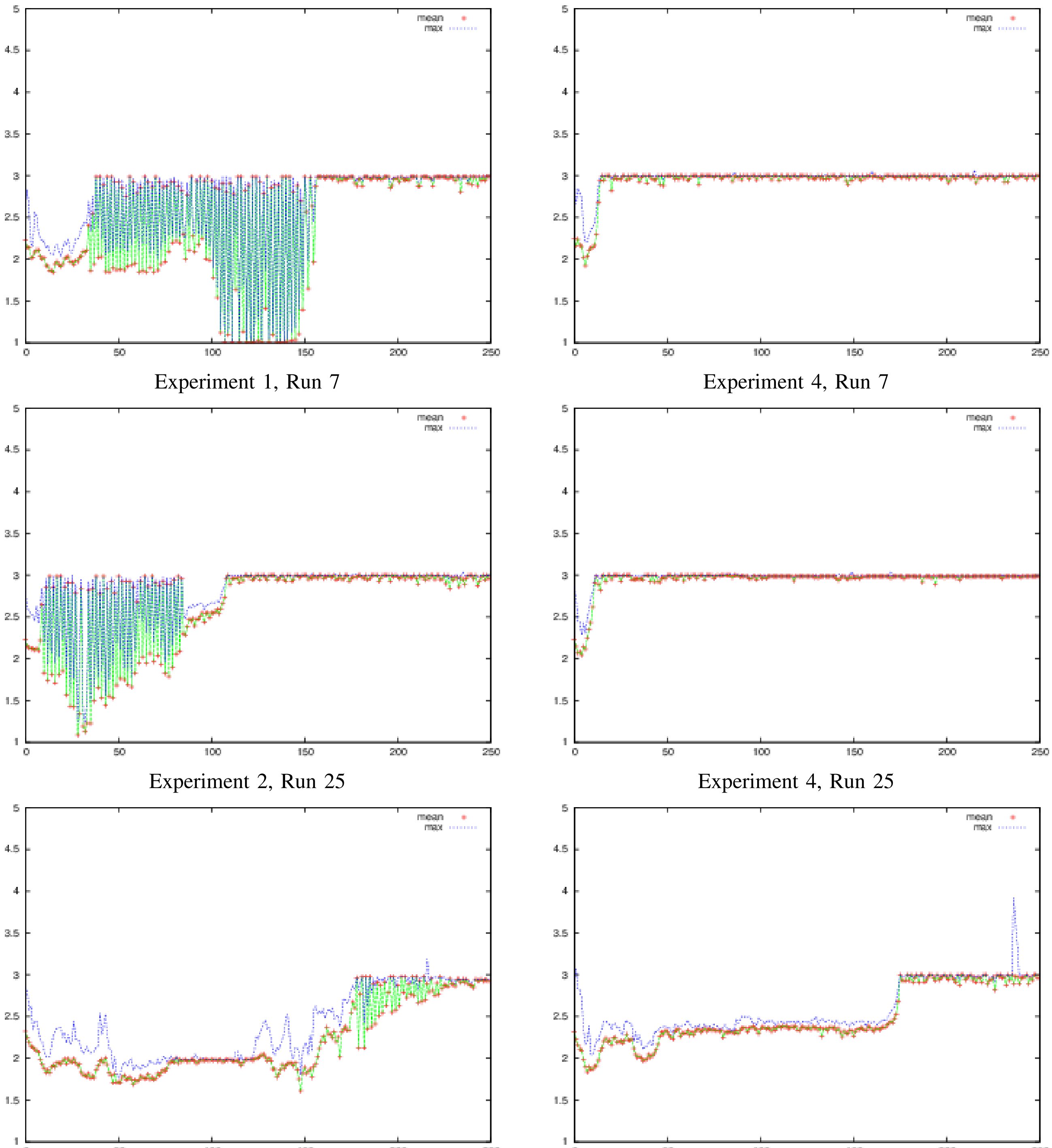
TABLE II
Shown are 95% confidence intervals on the probability of agents evolved in the experiment labelling the first column have of beating agents labelling the second in a tournament. The last column shows the percentage of actions where agent pairs that were coordinated during tournament play .

Pairs of Experiments	Confidence Interval	Coordination
1 vs 2	0.575±0.003	51.9%
1 vs 3	0.526±0.003	42.7%
1 vs 4	0.632±0.003	48.9%
2 vs 3	0.636±0.003	48.5%
2 vs 4	0.463±0.003	49.6%
3 vs 4	0.469±0.003	42.0%

Figure 4 diagrams the pattern of dominance between agents in different experiments. This figure illustrates the information in Table II in graphical form. The experiment that exhibited the worst competitive ability, Experiment 4, is the one that segregated "cooperation" and "sustained defection" off from all other behaviors. Experiment 1, with the best competitive ability, groups worse than random behaviors as its low-score category and cooperation as its high score category. Sustained defection, also a region in Experiment 2, which came in third, is very easy to recognize and so may not be worth spending one of the two threshold values to detect.

The researcher's expectation was that the control agents would be beaten by all three types of plastic agents. The actual result verifies that phenotypic plasticity *can* grant a competitive advantage but need not. The experiments presented here are a first attempt, using an extremely simple form of plasticity, and it seems that the adaptive ability built into the FSR's transition diagram is still providing the majority of the adaptive ability.

A natural choice would be to make the thresholds, and even the number of thresholds, part of the agents genes rather than imposing them as exogenous behaviors. Between them



Experiment 1, Run 7

Experiment 4, Run 7

Experiment 2, Run 25

Experiment 4, Run 25

Experiment 3, Run 3

Experiment 4, Run 25

Fig. 3. Shown are the evolution of fitness over time in populations from Experiments 1, 2, and 3 that exhibit thrashing and the corresponding runs from the control experiments. Correspondence is of run index number, to provide unbiased comparison. The tracks display the population average and maximum fitness.

the authors have published dozens of papers on evolving game playing agents and on the prisoner's dilemma specifically. The decision to do a very simple experiment is a lesson learned from burned hands: sorting out the effects of an evolvable plasticity mechanism is likely to be quite difficult without the background context provided by this study.

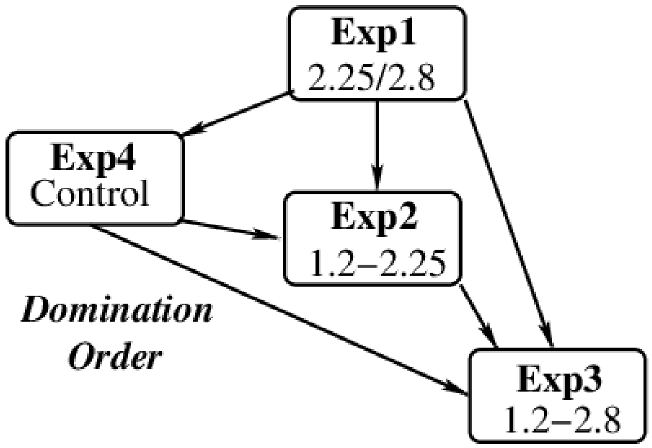


Fig. 4. Shown is the dominance order of the evolved agents. The agents from the experiment at the base of the arrow dominate the agents at the head; the numbers in the lower part of each box are the plasticity thresholds for the experiment

A. Behaviors Observed During Evolution

We begin with an overall comparison of the fitness in the three experiments. Figure 5 shows an average of population averages of the fitness tracks. This shows an inverse relationship between fitness, which usually corresponds to cooperation, and competitive ability. This inverse correlation has been observed in other studies [14], [8]. Examination of the individual runs in Experiment 1 showed that 23 of the runs ended in a cooperative state while only one located the all-defect Nash equilibrium for two-player single shot prisoner's dilemma. The competitive ability of the agents evolved in Experiment 1 does not rely on the simple mechanism of using the always defect strategy.

Experiments two and three achieve higher levels of cooperation that the control experiment. The issue of cooperation, while a central one in the study of prisoner's dilemma, was not a core focus in the design of this study. When considering competitive ability, we concluded that the threshold values 1.2 and 2.8, which detect sustained defection and cooperation for the score matrix used in this study, were not a good use of the adaptive resources represented by plasticity. It may be, however, that the plasticity mechanism permitted the agent to bin its reactions to sustained defection and cooperation that made both the emergence and maintenance of cooperation easier. This is a phenomenon that deserves additional study.

Figure 3 shows the evolution of fitness over evolutionary time for examples drawn from the three experiments with plasticity and the runs with the same index number (1-30) in the control runs. Since the plasticity mechanism has a sharp threshold, it is possible to jump back and forth across that threshold in successive generations as the strategies triggered by the last generation's average score caused the population to shift across a threshold value, in opposite directions in successive generations. This behavior is common in the plastic runs and absent in the control runs.

A similar sort of rapid strategic shifting was found as an evolved behavior in agents being trained to play the game *divide the dollar* [12]. In this game, two agents bid and, if their bids total at most a dollar, they are paid what they bid. In that study, the rapid behavioral change was adaptive because it made invading populations using the oscillating strategy difficult to invade. The fact that the thrashing behavior terminates in almost all examples suggests that it is possible to invade, but the adaptive value of this behavior, if any, should be investigated if only because it was so common in the plastic runs.

V. Conclusions and Next Steps

This study demonstrated that implementation of a simple mechanism for endowing game playing agents with phenotypic plasticity yielded substantial changes in the agents behavior and adaptive ability. The goal of increasing competitive ability is a qualified success: one of the plastic agent types gained a strong competitive ability, the other two were out-competed by control agents. An unexpected but pleasant result is that the implementation of phenotypic plasticity proved able to change the level of cooperation, and in two of the three experiments with plastic agents improved it.

The results of implementing phenotypic plasticity show a need for additional study, even for the relatively simple mechanism implemented in this study. A large number of pairs of plasticity thresholds need to be tested and the ability of the plasticity mechanism to enhance cooperation is also in need of additional study. One of John Nash's original specifications for an effective prisoner's dilemma agent was that if be cooperative but able to defend itself. The plasticity mechanism seems to be able to control evolution's implementation of their criterion, but not in an obvious manner.

There is a point that bears repeating about the well-known Nash equilibrium for the prisoner's dilemma: always defect. The theory of Nash equilibria does not apply to evolutionary systems which update their population more than one agent at a time. The definition of a Nash equilibria is that no *one* player may improve their score by changing strategies. The updating used in the evolutionary algorithm in this study updates twelve agents simultaneously. The emergence of cooperation in a majority of the interactions of both the control and phenotypically plastic agents in this study demonstrates that the Nash equilibria of prisoner's dilemma in classical game theory lacks predictive power in the investigation of game playing agents trained by evolution.

When applied to evolutionary computation in optimization, phenotypic plasticity may provide a road to robust optimization [19]. Adaptive representations for optimization have been devised [10] and the representation cited exhibits

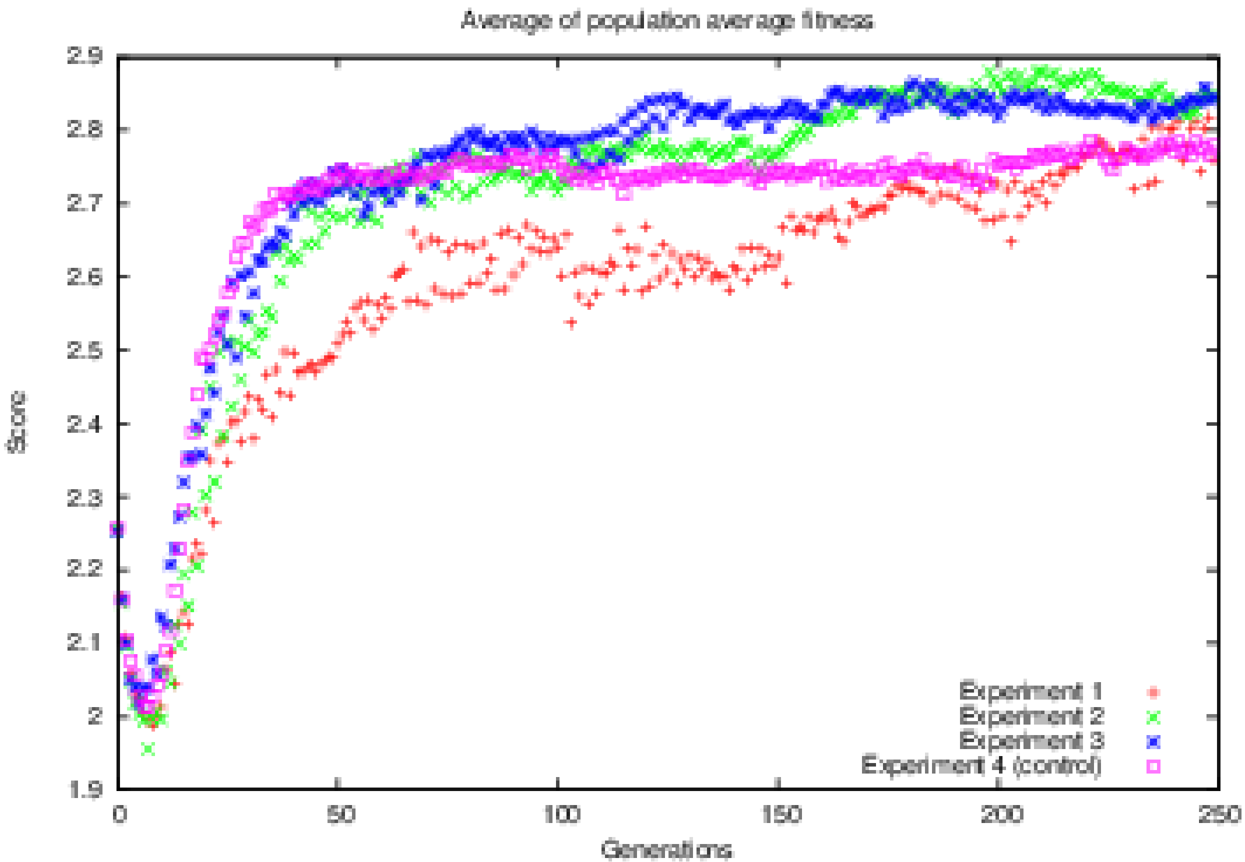


Fig. 5. Shown are the average, over all 30 evolutionary runs, of the population average fitness values for the four experiments over the course of evolutionary time.

a different form of phenotypic plasticity than the one studied here. This is a rich area for future investigation.

A. Improving the Experimental Design

An obvious improvement in the experimental design would be to implement hysteresis so as to prevent thrashing. This would require distinct but similar thresholds for shifting strategies. Instead of 1.2, an agent would shift to the first strategy thread at 1.15 or below and to the second one and 1.25 or above. Thrashing is still possible in this scheme, but would be less likely. This is separate from the idea of investigating what adaptive benefit thrashing might produce.

The next logical step in this research is to make the adaptive thresholds, and even the number of adaptive thresholds, part of the agents genome. This may permit the discovery of "good" adaptive thresholds or may help demonstrate that different adaptive thresholds grant an intransitive form of advantage.

B. Other Representations, Other Games

There are a number of issues that affect the strategy space available to evolving prisoner's dilemma agents and the strategies actually selected from those spaces by evolutionary competition. These include the agent representation, the resources (number of states, neurons, probability levels in a Markov chain), and the details of the algorithm including population size, geographic structure, and duration of evolution. The strategy *Fortress 4* which cooperates only after three mutual defections, was discovered only after 16,000 generations of evolution [15].

Phenotypic plasticity has the potential to enhance agent robustness. This means that incorporating it into simulations may *reduce* the impact of the many factors that have been found to affect cooperation level, competitive ability, and the mix of strategies that arise during evolutionary training.

Finally, prisoner's dilemma is a very interesting game, but there are so many others. In addition to mathematical games, like the snowdrift game or divide the dollar, phenotypic plasticity may be valuable for creating better bots or opponents in video games. A simple example of this application of phenotypic plasticity would be to have an agent that, when substantially wounded, goes into a different fighting more. It might flee or it might go berserk style.

The mechanism of phenotypic plasticity used in this study was relatively proscriptive. The thresholds that triggered plastic behavior were given at the top level of an experiment and the number of threads, three, was determined ahead of time. There are a vast number of other ways that phenotypic plasticity could be implemented. Moving closer to the implementation at the level of basic physiology interacting with the environment is a goal of future research. In particular, permitting agents to self-organize their plasticity to a greater degree would be an excellent goal.

The ability of the agents in this study to organize their own plasticity comes in their genetic control over the composition of the transition and response functions encoded by the agent. An interesting analysis not performed would be to see if evolution left the transition diagram accessible to the three pointers connected, or if in fact evolution chose to give different pointers their own individual transition diagrams.

Another possible mechanism for phenotypic plasticity would be to have agents with a mechanism for modifying response labels in its finite state machine, flipping them between the values **C** and **D**. This mechanism could be conditioned on average score with a higher rate of response modification when the score is lower. The original values would be retained in the genotype with a digital "enzyme" whose activity levels were conditioned on payoff level making the modifications again in each session of play.

REFERENCES

[1] D. Ashlock and A. Aherk. Non-local adaptation of artificial predators and prey. In *Proceedings of the 2005 Congress on Evolutionary Computation*, volume 1, pages 41–48, Piscataway, NJ, 2005. IEEE Press.

[2] D. Ashlock, W. Ashlock, S. Samothrakis, S. Lucas, and C. Lee. From competition to cooperation: Co-evolution in a rewards continuum. In *Proceedings of the IEEE 2012 Conference on Computational Intelligence in Games*, pages 33–40, Piscataway, NJ, 2012. IEEE press.

[3] D. Ashlock, E. Blankenship, and J. D. Gandrud. A note on general adaptation in populations of painting robots. In *Proceedings of the 2003 Congress on Evolutionary Computation*, pages 46–53, Piscataway, NJ, 2003. IEEE Press.

[4] D. Ashlock and K. M .Bryden. Non-local adaptation in bidding agents. In *Smart Engineering System Design: Neural Networks, Evolutionary Programming, and Artificial Life*, volume 15, pages 193–199. ASME Press, 2005.

[5] D. Ashlock, K. M. Bryden, S. Corns, and J. Schonfeld. An updated taxonomy of evolutionary computation problems using graph-based evolutionary algorithms. In *Proceedings of the 2006 Congress On Evolutionary Computation*, pages 403–410, Piscataway, NJ, 2006. IEEE Press.

[6] D. Ashlock and E. Y. Kim. The impact of cellular representation on finite state agents for prisoner's dilemma. In *Proceedings of the 2005 Genetic and Evolutionary Computation Conference*, pages 59–66, New York, 2005. ACM Press.

[7] D. Ashlock and E. Y. Kim. The impact of elite fraction and population size on evolved iterated prisoner's dilemma agents. In *Proceedings of the IEEE 2016 Congress on Evolutionary Computation*, pages 364–371, 2016.

[8] D. Ashlock and E.Y. Kim. Prisoner's dilemma: the payoff values matter. In *Proceedings of the 2010 IEEE Conference on Computational Intelligence in Games*, pages 219–226, 2010.

[9] D. Ashlock, E.Y. Kim, and N. Leahy. Understanding representational sensitivity in the iterated prisoner's dilemma with fingerprints. *IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews*, 36(4):464–475, 2006.

[10] D. Ashlock and J. Montgomery. An adaptive generative representation for evolutionary computation. In *Proceedings of the IEEE 2016 Congress on Evolutionary Computation*, pages 1578–1585, Piscataway NJ, 2016. IEEE Press.

[11] Dan Ashlock and Brad Powers. The effect of tag recognition on non-local adaptation. In *Proceedings of the 2004 Congress on Evolutionary Computation*, volume 2, pages 2045–2051, Piscataway, NJ, 2004. IEEE Press.

[12] Daniel Ashlock. Gp-automata for dividing the dollar. In *Genetic Programming 1997, Proceedings of the Second Annual Conference on Genetic Programming*, pages 18–26, 1997.

[13] Daniel Ashlock. *Evolutionary Computation for Opimization and Modeling*. Springer, New York, 2006.

[14] Daniel Ashlock and John Mayfield. Acquisition of general adaptive features by evolution. In *Proceedings of the 7th Annual Conference on Evolutionary Programming*, pages 75–84, New York, 1998. Springer.

[15] W. Ashlock and D. Ashlock. Changes in prisoner's dilemma strategies over evolutionary time with different population sizes. In *Proceedings of the 2006 Congress On Evolutionary Computation*, pages 1001–1008, Piscataway, NJ, 2006. IEEE press.

[16] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[17] Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[18] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone. *Genetic Programming : An Introduction*. Morgan Kaufmann, San Francisco, 1998.

[19] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009.

[20] J. Bendor, R.M. Kramer, and S. Stout. When in doubt: Cooperation in a noisy prisoner's dilemma. *Journal of Conflict Resolution*, 35(4):691–719, 1991.

[21] S. Y. Chong and X. Yao. Behavioral diversity, choices and noise in the iterated prisoner's dilemma. *IEEE Transaction on Evolutionary Computation*, 9:540–551, 2005.

[22] David Doty. Non-loca evolutionary adaptation in grid plants. In *Proceedings of the 2004 Congress on Evolutionary Computation*, volume 2, pages 1602–1609, Piscataway, NJ, 2004. IEEE Press.

[23] Anders Forsman. Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity*, 115(4):276, 2015.

[24] Michael Hemesath. Cooperate or defect? Russian and American students in a prisoner's dilemma. *Comparative Economics Studies*, 176:83–93, 1994.

[25] John M. Houston, Judy Kinnie, Bernice Lupo, Christeine Terry, and Sandy S. Ho. Competitiveness and conflict behavior in simulation of a social dilemma. *Psychological Reports*, 86:1219–1225, 2000.

[26] S. A. Kelly, T. M. Panhuis, and A. M. Stoehr. Phenotypic plasticity: molecular mechanisms and adaptive significance. *Comprehensive Physiology*, 2(2):1417–1439, 2012.

[27] E. Y. Kim and D. Ashlock. Changing resources available to game playing agents: Another relevant design factor in agent experiments. *IEEE Transactions on Computational Intelligence and AI in Games*, 9(4):321–332, 2016.

[28] J. E. Krebs, E.S. Goldstein, and S. T. Kilpatrick. *Genes XII*. Jones and Barttlett Learning, Burlington MA, 2017.

[29] Courtney J Murren, Josh R Auld, H Callahan, Cameron K Ghalambor, Corey A Handelsman, Mary A Heskel, JG Kingsolver, Heidi J Maclean, Joanna Masel, Heather Maughan, et al. Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity*, 115(4):293, 2015.

[30] Daniel B. Neill. Optimality under noise: Higher memory strategies for the alternating prisoner's dilemma. *Journal of Theoretical Biology*, 211:159–180, 2001.

[31] N. Pezaro, M. B. Thompson, and J. S. Doody. Seasonal sex ratios and the evolution of temperature-dependent sex determination in oviparous lizards. *Evolutionary Ecology*, 30(3):551–565, Jun 2016.

[32] D. Pfennig. The adaptive significance of an environmentally-cued developmental switch in an anuran tadpole. *Oecologia*, 85(1):101–107, 1990.

[33] T. Piersma and J.A. Van Gils. *The Flexible Phenotype: A Body-Centred Integration of Ecology, Physiology, and Behaviour*. Oxford University Press, 2011.

[34] Massimo Pigliucci. Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution*, 20(9):481 – 486, 2005.

[35] Sidney I. Resnick. *Adventures in Stochastic Processes*. Birkhauser, Boston, 1992.

[36] Duncan Roy. Learning and the theory of games. *Journal of Theoretical Biology*, 204:409–414, 2000.

[37] Karl Sigmund and Martin A. Nowak. Evolutionary game theory. *Current Biology*, 9(14):R503–505, 1999.

[38] J. Wu and R. Axelrod. How to cope with noise in the iterated prisoner's dilemma. *Journal of Conflict Resolution*, 39(1):183–189, 1995.