Generalization and Diversity in Co-evolutionary Learning

Siang Yew Chong, Peter Tiňo, and Xin Yao
The Centre of Excellence for Research in
Computational Intelligence and Applications (Cercia)
School of Computer Science
The University of Birmingham, UK.
E-mails: S.Y.Chong@cs.bham.ac.uk, X.Yao@cs.bham.ac.uk, P.Tino@cs.bham.ac.uk

Abstract

Games have long played an important role in the development and further understanding of co-evolutionary learning systems. In particular, the search process in co-evolutionary learning is guided by strategic interactions between solutions in the population, which can be naturally framed as game-playing. We study two important issues in co-evolutionary learning–generalization performance and diversity–using games. On the one hand, one is concerned with the co-evolutionary learning of strategies with high generalization performance, that is, strategies that can outperform against a large number of test strategies (opponents) that may not be seen before during co-evolution. On the other hand, the other one is concerned with diversity levels in the population that may lead to the search of strategies with poor generalization performance. It is not known if there is a relationship between generalization and diversity in co-evolutionary learning. This paper investigates whether there is such a relationship in co-evolutionary learning through a detailed empirical study. We systematically investigate the impact of various diversity maintenance approaches on the generalization performance of co-evolutionary learning quantitatively using case studies. The problem of the iterated prisoner's dilemma (IPD) game is considered. Unlike what was done in past studies, we can measure both the generalization performance and the diversity level of the population of evolved strategies. Results from our case studies show that the introduction and maintenance of diversity do not necessarily lead to the co-evolutionary learning of strategies with high generalization performance. However, if individual strategies can be combined in the form of ensembles, there is the potential of exploiting diversity in co-evolutionary learning to improve generalization performance. Specifically, when the introduction and maintenance of diversity lead to a speciated population during co-evolution, where each specialist strategy is capable of outperforming different opponents, the ensemble can have a significantly higher generalization performance compared to individual strategies when it can sufficiently often choose the best specialist strategy against a test strategy.

Index Terms: Evolutionary Computation, Co-evolutionary Learning, Generalization Performance, Diversity Maintenance, Iterated Prisoner's Dilemma.

1 Introduction

Evolutionary computation (EC) refers to the study of computational approaches that are motivated from and emulate the process of natural evolution. Although current EC studies have since expanded and involved interdisciplinary research with other computational methodologies (such as neural networks) [1, 2], EC studies traditionally and still continue to focus on a broad class of population-based, stochastic search algorithms known as evolutionary algorithms (EAs) [3, 4, 5]. EAs can be described using the unified *generate-and-test* framework for search algorithms [2]. The search process of EAs involves samples (populations) of potential solutions drawn from the space of representation of the problem domain, which undergo iterated applications of variation (generation of new solutions based on solutions in the previous iteration through some pertubation processes) and selection (testing solutions for inclusion in the next iteration based on some quality or fitness measurements) [2].

Although EC has been studied in many problem areas (such as optimization [6, 7] and classification [8, 9]), games have played important roles in the development and understanding of EC [10]. A game has a specific set of rules that constrains strategies to certain behaviors (legal moves), with goals for strategies to meet (to win the game), and rewards for those that better achieve the goals under constraints of finite resources (payoff for a move). A game has enough subtleties to allow representation of a wide-range of complex behaviors (a diverse set of strategies) [11]. Games can capture intrinsic properties of complex, real-world problems where EC methodologies are developed to obtain solutions, which can range from solutions (strategies) for board games to economic games [11, **?**], but are sufficiently simple to enable extensive and in-depth analysis of EC methodologies.

In EC, co-evolutionary learning refers to a broad class of population-based, stochastic search algorithms that involve the simultaneous evolution of competing solutions with coupled fitness [3]. As with EAs, co-evolutionary learning can be described using the generate-and-test framework. For example, early co-evolutionary learning systems are implemented using co-evolutionary algorithms derived from EAs [12, 11]. However, co-evolutionary learning and EAs are fundamentally different in how the fitness of a solution is assigned, which may lead to significantly different outcomes when they are applied to similar problems (different search behaviors on the space of solutions) [13, 14]. Classical EAs are formulated in the context of *optimization* [3, 6, 5], where an absolute fitness function is used to assign fitness values to solutions. The solution fitness is *objective* (remains the same regardless of the composition of the population). Co-evolutionary learning assigns a solution's fitness through its interactions with other competing solutions in the population. The solution fitness is *subjective* (depends on the composition of the population).

Co-evolutionary learning is originally proposed as a viable alternative for problems where it is very difficult to construct an absolute quality measurement for solutions (the fitness function) through which optimization-based search algorithms such as EAs can be used. The evolutionary search is guided by the *strategic* interactions between the competing solutions in the population. In particular, these interactions provide the basis to compute relative fitness values that are used to rank the competing solutions, which would affect the selection process of the population and subsequently the variation in the population. The potential for problem solving through co-evolutionary learning can be realized from an *arms-race* dynamics in the evolutionary search process that leads to solutions that are innovative and of increasingly higher quality [15, 16, 17, 18, 19].

In co-evolutionary learning, games have played an important role not only in the development of various co-evolutionary algorithms, but are crucial for further understanding of its approach for problem-solving. In particular, games provide a natural framework to study co-evolutionary learning as the interactions between the co-evolving solutions in the population can be framed as game-playing. Games are used to study two main properties of a co-evolutionary learning system: decision-making and learning [11]. Decision-making refers to the ability of the evolved strategy to make appropriate responses for the given stimuli (behavior) in light of specific goals that must be achieved. Learning refers to the ability of the system to train (evolve) strategies that can respond to a wide range of environments (opponents). The measure of success of a co-evolutionary learning system in solving games can be viewed as its ability to evolve strategies with behaviors such that they can outperform against a large number of different opponents.

Although co-evolutionary learning has been successfully applied in solving games [12, 11, 20, 21, 19, 22], the methodology has been shown to suffer from *co-evolutionary pathologies* that can potentially have a negative impact on its performance [18, 23, 24, 25, 26, 27]. In the context of game-playing, poor search performance in co-evolutionary learning can be a result of the *overspecialization* of the population to a particular strategy that performs well only against specific opponents rather than a large number of different opponents [25, 26].

As such, there are strong motivations to investigate and understand the conditions in the search process of co-evolutionary learning that have impact on its performance. Generalization performance and diversity are two important and related issues of the study. For the first issue, one is concerned with the co-evolutionary learning of strategies with high generalization performance, that is, strategies that can outperform against a large number of test strategies (opponents) that may not be seen before during co-evolution [25, 28, 29]. For the second issue, one is concerned with diversity levels in co-evolutionary learning [30, 26, 24].

The lack of diversity has been claimed to have a negative impact on the generalization performance of co-evolutionary learning. Studies that have investigated this issue have proposed *diversity maintenance* techniques to introduce and maintain diversity in the population through specific use of selection and variation processes, which have been shown subsequently to improve the performance of co-evolutionary learning [30, 28, 26, 31]. However, *it is not known if there is a relationship between generalization performance and diversity in co-evolutionary learning.* Establishing any direct relationship between generalization performance and diversity theoretically would be very difficult. Although earlier studies have turned to various empirical approaches, there is a lack of rigorous analysis to investigate this issue in terms of generalization performance measures in co-evolutionary learning [28, 32, 31]. Furthermore, these past studies made little attempt to measure diversity levels in co-evolutionary learning despite strong claims that they are increased while investigating their impact on performance [28, 32, 24, 23].

In this paper, we present a detailed empirical study as a first step to investigate whether there is a relationship between generalization performance and diversity in co-evolutionary learning. We have systematically investigated the impact of various diversity maintenance techniques on the generalization performance of co-evolutionary learning quantitatively through case studies involving IPD games. Our case studies address the two shortcomings found in previous empirical studies. First, we have made a series of quantitative measurements of the generalization performance of co-evolutionary learning systems with and without diversity maintenance that are based on the estimation procedure we have developed theoretically in [29]. Second, we have made a series of quantitative measurements of diversity levels in the population of various co-evolutionary learning systems with and without diversity maintenance that are based on relevant diversity measures that have been identified earlier in [33]. These measurements allow us to determine whether the generalization performance and diversity levels in the population have been increased as a result of the application of diversity maintenance in co-evolutionary learning in comparison to the classical co-evolutionary learning without diversity maintenance that we have used as a baseline.

In investigating whether there is relationship between generalization performance and diversity in co-evolutionary learning, we try to answer two related questions: (1) what is the right form of diversity that would be beneficial for generalization, and (2) what amount of the 'right' form of diversity is needed for good generalization. The answers to these two questions may allow us to identify specific conditions whereby diversity can be exploited to improve the generalization performance in co-evolutionary learning. Although our results have shown that not all diversity maintenance techniques that introduce and maintain diversity lead to a significant increase in the generalization performance of co-evolutionary learning, we have identified a condition with which diversity plays a positive role in improving generalization performance. In particular, if individual strategies from the population can be combined in the form of ensembles, there is the potential of exploiting diversity in co-evolutionary learning to improve generalization performance. Specifically, when the introduction and maintenance of diversity lead to a speciated population during co-evolution, where each specialist strategy is capable of outperforming different opponents, the ensemble has a significantly higher generalization performance compared to individual strategies when it can sufficiently often choose the best specialist strategy against a test strategy.

The rest of the paper is organized as follows. Section 2 reviews a number of diversity maintenance techniques in co-evolutionary learning that have been proposed in the past. The case studies that is presented in this paper compare co-evolutionary learning employing these diversity maintenance techniques with the classical co-evolutionary learning without diversity maintenance. Section 3 presents the methodology of the experiments used in the case studies. The section also describes how generalization performance and diversity levels of co-evolutionary learning are measured. Section 4 presents results of the case studies and discusses the observations made from the results. Section 5 concludes the paper with remarks for future studies.

2 Diversity Maintenance Techniques

The search process in evolutionary systems can be described with the notion of how these systems maintain a spread of points (the population) in the search space through *exploration* and *exploitation*. An evolutionary system operating during an exploration phase would spend more effort searching other globally promising regions of the search space that are not covered by the current population. In contrast, evolutionary systems operating during an exploitation phase would spend more effort searching locally in regions covered by the current population to focus on obtaining the required solutions to the problem. The success of an evolutionary search depends on how and when to explore or exploit the search space such that solutions that are of better quality are obtained in subsequent generations. There may be a tradeoff between exploring and exploiting the search space in any generation, and understanding this tradeoff may lead to the design of adaptive control in the search process that leads to improvements in the search performance [34].

One common approach used to study the dynamics of the evolutionary process and the role they play in search is with the notion of diversity. Diversity is concerned with the levels and types of variety in the population [33], which can change depending on the exploration or exploitation phase of the evolutionary process. Many studies [35, 36, 37, 38, 39, 40, 41, 42] have investigated methods that exploit diversity to further improve the search performance of evolutionary systems. Some of these algorithms are designed to introduce and maintain diversity in the population during the evolutionary search process. Early examples of these *diversity maintenance* techniques have been used in EAs to solve difficult optimization problems involving multiple optima [35, 36, 37, 38, 39]. The techniques have since been adapted for co-evolutionary learning with the motivation that diversity may encourage fitness variations in the population for continued search of solutions with higher generalization performance [43, 28].

In our case studies, we categorize the different diversity maintenance techniques into two groups that are based on how the specific algorithm operates in the co-evolutionary learning framework. *Implicit* techniques emphasize diversity maintenance through the selection process. *Explicit* techniques emphasize diversity maintenance through the variation process. This section describes several established diversity maintenance techniques in co-evolutionary learning studies that we use for our case studies.

For clarity, we use examples of game-playing in our descriptions of how a specific diversity maintenance technique is implemented in co-evolutionary learning. The interactions between solutions in co-evolutionary learning that is used to determine the solution fitness (relative to other solutions) can be framed as game-playing. A game is played between two strategies (solutions) and consists of a series of moves that are made by each strategy. During one game-play, both strategies accumulate payoffs that are based on the joint moves they have made, which are then used to determine the outcome when the game finishes. An example would be the case where a strategy wins the game when it has a higher total payoff in comparison to the other strategy, which loses the game. Diversity maintenance techniques differ in terms of how information from game outcomes is obtained and used.

2.1 Implicit Diversity Maintenance

2.1.1 Speciation

One of the early implicit diversity maintenance techniques is *speciation* (or *niching*). Speciation is usually implemented with *fitness sharing* algorithms. The main idea of fitness sharing is to change fitness evaluations so that fitnesses of solutions are reduced to lower values if the solutions are more similar. Fitness sharing of solutions can be implemented by computing the shared fitness f_i^s from the original, unshared fitness f_i for each solution i in the population [39]. This is to emphasize the search for more dissimilar solutions and possibly leading to a population forming a diverse set of niches with unique solutions. *Competitive fitness sharing* [26] and *implicit fitness sharing* [28] are two fitness sharing approaches that have been studied for co-evolutionary learning.

In the competitive fitness sharing approach, each evolved strategy i in the population would interact with a random set of opponent strategies P_x sampled from the same population. Strategy i plays against an opponent j in P_x and receives a fitness value of $1/N_j$ if i wins. N_j is the number of opponents in the population that j has lost to. N_j can be precomputed for each strategy j in the population through a tournament where every strategy j plays against all strategies in the population. Obviously, no strategy will win against a copy of itself, which means that the maximum value of $1/N_j$ is $1/1 = 1$. The minimum value for the worst strategy that loses to all strategies in the population would be $1/N_j = 1/\mathrm{POPSIZE}$ (POPSIZE is the total number of parent and offspring strategies). The shared fitness for a strategy f_i^s is calculated as follows:

$$f_i^s = \sum_{j \in P_x} \frac{1}{N_j} \cdot O(i,j), \tag{1}$$

where $O(i,j)$ is 1 if i wins against j, and is 0 if i loses. Essentially, the competitive fitness sharing approach encourages the selection of strategies that are unique in terms of defeating specific opponents that few other competing strategies could defeat (having higher fitness value of $1/N_j$ as N_j is smaller). These strategies are rewarded with a higher total sum of fitness values even though they may defeat a smaller number of opponents.

In the implicit fitness sharing approach, the shared fitness for each strategy i in the population is obtained using the algorithmic procedure outlined in Figure 1. During one typical run of the procedure for a strategy i (there are C number of runs), a random sample σ of opponent strategies (other than strategy i itself) would be drawn from the current population. The strategy i would play against each opponent strategy in σ. The opponent that achieves the highest game payoff against strategy i would receive a fitness value (we use the game payoff itself). If $n \leq |\sigma|$ number of opponents tie in terms of their game payoffs against strategy i, they would share the fitness (the game payoff divided by n). The approach has been shown mathematically to approximate the original fitness sharing scheme [44] but with the advantage of not having to use a sharing function, which can be problem-specific [45].

Repeat C times

1. Select a random sample σ of opponent strategies from the current population.
2. Find the opponent strategy in σ that achieves the highest payoff (or the largest winning margin) against i.
3. The best opponent strategy in σ receives a fitness value (equivalent to the game payoff). In the case of a tie between opponent strategies, the fitness is shared equally.

Figure 1: The procedure for implicit fitness sharing to assign shared fitness to a strategy.

2.1.2 Pareto Co-evolution

Pareto co-evolution is a co-evolutionary learning approach with implicit diversity maintenance technique based on the framework of multi-objective optimization introduced earlier in EAs [32, 31, 23]. Pareto co-evolution is motivated to address the problem associated with the use of scalar fitness evaluations in classical co-evolutionary learning. In the context of games, it has been argued that a strategy's fitness based on a simple weighted sum of values given by game outcomes can be misleading. One may not know at the time fitness evaluations are made *a priori* whether all outcomes are of equal importance (uniform weightings) or that some outcomes are more important (nonuniform weightings).

Pareto co-evolution reformulates the fitness evaluation by explicitly using opponent strategies (test cases) to represent the different, underlying objectives of the problem that a strategy must address. The selection process in pareto co-evolution is based on the pareto-dominance relationship. For single population pareto co-evolution, to determine whether strategy i *pareto dominates* strategy j, we consider all $k \neq i, j$

strategies from the population as the N objectives. The objective value $O(i,k)$ refers to the game outcome of strategy i when it plays against strategy k. Strategy i pareto dominates j if i has at least the same game outcome as j for every one of the N objectives, and there is at least an objective for which i has a better game outcome than j [23]:

$$dom(i,j) \iff \forall k : O(i,k) \geq O(j,k) \wedge \exists k : O(i,k) > O(j,k), \tag{2}$$

with $1 \leq k \leq N$. Strategies i and j are mutually non-dominating if i is better than j on one objective and j is better than i on another objective, with both strategies being equivalent in the remaining objectives.

The selection process in pareto co-evolution can take the form of a ranking of strategies according to the pareto layers that are obtained from the pareto-dominance relationship. For example, mutually non-dominating strategies in the pareto front would be the most preferable strategies for selection. The second pareto layer can be obtained by first removing strategies in the pareto front from the population. The second pareto layer would contain mutually non-dominating strategies among the subpopulation that is obtained after removing the pareto front from the full population. The following pareto layers can be obtained systematically with respect to subpopulations that do not include strategies in higher pareto layers [31].

2.1.3 Reducing Virulence

Both speciation and pareto co-evolution approaches to diversity maintenance require sophisticated algorithmic implementations. In [24], a simple approach that is inspired from the natural interactions between hosts and parasites is proposed. In a biological co-evolutionary system, highly virulent parasites could potentially incapacitate the hosts, which in turn, reduce the chances for parasites to reproduce. In co-evolutionary learning, similar observations have been made whereby strong competing solutions can quickly take over the population, leading to *co-evolutionary disengagement* where the system is unable to differentiate between the solutions for continued evolutionary search [24]

The implementation of *reducing virulence* involves a form of rescaling of fitness values. Classical co-evolutionary learning operates at maximum virulence. The original fitness value f_i is assumed to be proportional to the solution i's ability to solve as many test cases as possible in the population (e.g., a simple weighted sum, which is normalized according to the fitness value of the best solution). In the context of game-playing, one form of reducing virulence is to lower the fitness of the top performing strategies in the co-evolving population that outperform the largest number of the opponents. Following [24], the virulence level v is defined as the tendency to maximize fitness and a reduced virulence level v changes an evolved strategy's original fitness value to:

$$f_i^v = \frac{2f_i}{v} - \frac{f_i^2}{v^2}, \tag{3}$$

where $0.5 \leq v \leq 1.0$.

A co-evolutionary learning system with Maximum Virulence level uses $v = 1$, while a system with reduced virulence level has v in the range of $[0.5, 1.0]$ (e.g., Moderate Virulence level $v = 0.75$). Figure 2 compares how fitness values are rescaled between the use of Maximum and Moderate Virulence level. Note that the lowest value for v is 0.5 as any lower value will indicate fitness is rescaled to encourage strategies to outperform smaller numbers of opponents [24].

2.2 Explicit Diversity Maintenance

Explicit diversity maintenance techniques involve increasing the level of variation in the population through variation operators in the co-evolutionary search process. A simple approach is to use a higher

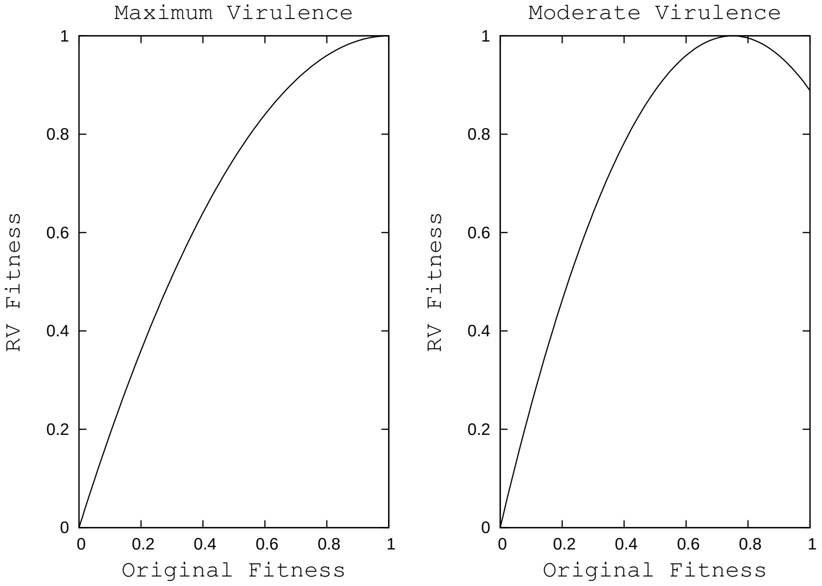


Figure 2: Comparison of virulence curves between Maximum Virulence ($v = 1.0$) and Moderate Virulence ($v = 0.75$) [24].

value for the parameter used in the variation operator that can increase the mutation probability. This simple approach is investigated in our case studies. However, it is noted that there are other more sophisticated approaches where the parameter controlling mutation probability can be varied according to some algorithmic procedures, e.g., setting higher parameter values to introduce and maintain more variations in the population if measurements on the population indicate low diversity levels [42].

3 Investigating the Relationship Between Diversity and Generalization Performance in Co-evolutionary Learning

In this section, we present a methodology to investigate whether there is a relationship between diversity and generalization performance in co-evolutionary learning. We use case studies to systematically investigate the impact of various diversity maintenance techniques on the generalization performance of co-evolutionary learning quantitatively. The problem of the IPD game is considered. Two categories of diversity maintenance techniques that we investigate are: (1) implicit diversity maintenance techniques, which include speciation (competitive and implicit fitness sharing), pareto co-evolution, and reduced virulence, and (2) explicit diversity maintenance techniques with different and higher parameter settings for mutation probability.

3.1 The IPD Game

In the classical IPD game setting, two players engage in repeated interactions are given two choices to make, cooperate and defect [46, 47, 48]. The classical IPD game is formulated using a predefined payoff matrix that specifies the payoff that a player receives for the joint moves of the player and the opponent. Figure 3 illustrates the payoff matrix for the classical IPD game. Both players receive R (*reward*) units of payoff if both cooperate. They both receive P (*punishment*) units of payoff if they both defect. However, when one player cooperates while the other defects, the cooperator will receive S (*sucker*) units of payoff while the defector receives T (*temptation*) units of payoff.

	Cooperate	Defect
Cooperate	R R	T S
Defect	S T	P P

Figure 3: The payoff matrix framework of a two-player, two-choice game. The payoff given in the lower left-hand corner is assigned to the player (row) choosing the move, while that of the upper right-hand corner is assigned to the opponent (column).

In the classical IPD, the values R, S, T, and P must satisfy the constraints: $T > R > P > S$ and $R > (S+T)/2$. Any set of values can be used as long as they satisfy the IPD constraints (we use $T = 5$, $R = 4$, $P = 1$, and $S = 0$ in all our experiments). The game is played when both players choose between the two alternative choices over a series of moves. The dilemma in the game is reflected by the tension between rationality of individuals who are tempted to defect and rationality of the group where every individual is rewarded by mutual cooperation since both players who are better off mutually cooperating than mutually defecting are vulnerable to exploitation by one of the party who defects [48].

The classical IPD game has been extended to more complex versions to better model real-world interactions. One example is to extend the classical IPD with multiple, discrete levels of cooperation [25, 49, 50, 51, 52]. The n-choice IPD game can be defined by the payoffs obtained through the following linear interpolation:

$$p_{\mathrm{A}} = 2.5 - 0.5c_{\mathrm{A}} + 2c_{\mathrm{B}}, \quad -1 \leq c_{\mathrm{A}}, c_{\mathrm{B}} \leq 1, \tag{4}$$

where p_{A} is the payoff to player A, given that c_{A} and c_{B} are the cooperation levels of the choices that players A and B make, respectively. Figure 4 [49] illustrates the payoff matrix for the *four*-choice IPD game we use in the case studies.

		PLAYER B			
		$+1$	$+\frac{1}{3}$	$-\frac{1}{3}$	-1
PLAYER A	$+1$	4	$2\frac{2}{3}$	$1\frac{1}{3}$	0
	$+\frac{1}{3}$	$4\frac{1}{3}$	3	$1\frac{2}{3}$	$\frac{1}{3}$
	$-\frac{1}{3}$	$4\frac{2}{3}$	$3\frac{1}{3}$	2	$\frac{2}{3}$
	-1	5	$3\frac{2}{3}$	$2\frac{1}{3}$	1

Figure 4: The payoff matrix for the two-player *four*-choice IPD game [49]. Each element of the matrix gives the payoff for Player A.

It is noted that when generating the payoff matrix for an n-choice IPD game, the following conditions must be satisfied [49]:

1. For $c_{\mathrm{A}} < c'_{\mathrm{A}}$ and constant c_{B}: $p_{\mathrm{A}}(c_{\mathrm{A}}, c_{\mathrm{B}}) > p_{\mathrm{A}}(c'_{\mathrm{A}}, c_{\mathrm{B}})$,

2. For $c_\mathrm{A} \leq c'_\mathrm{A}$ and $c_\mathrm{B} < c'_\mathrm{B}$: $p_\mathrm{A}(c_\mathrm{A}, c_\mathrm{B}) < p_\mathrm{A}(c'_\mathrm{A}, c'_\mathrm{B})$, and
3. For $c_\mathrm{A} < c'_\mathrm{A}$ and $c_\mathrm{B} < c'_\mathrm{B}$: $p_\mathrm{A}(c'_\mathrm{A}, c'_\mathrm{B}) > (1/2)(p_\mathrm{A}(c_\mathrm{A}, c'_\mathrm{B}) + p_\mathrm{A}(c'_\mathrm{A}, c_\mathrm{B}))$.

These conditions are analogous to those for the classical IPD's. The first condition ensures that defection always pays more. The second condition ensures that mutual cooperation results in higher payoff than that of mutual defection. The third condition ensures that alternating between cooperation and defection pay less in comparison to just playing cooperation.

Many past studies using IPD games are motivated to investigate specific conditions that can lead to the evolution of certain behaviors, for example, how and why cooperative behaviors can be learned through a co-evolutionary process guided by interactions between competing strategies [53, 54, 55, 56, 57, 58]. In this paper, our use of IPD games is motivated to determine whether diversity can have an impact on the co-evolutionary learning of strategies that are "robust".

The notion of robustness has been defined in the context of strategy performance against some expert-designed strategies [12] and in the context of generalization of strategies obtained from a learning process [59, 30, 43]. Other studies such as in [**?**] have investigated in the context of non-local adaptation, performance between populations of strategies. In this paper, we consider the notion of generalization and apply the theoretical framework of generalization performance in co-evolutionary learning that we have formulated earlier in [29]. We consider the *four*-choice IPD game with 150 rounds for deterministic and reactive, memory-one strategies as it provides a difficult search problem (a large search space that cannot be exhaustively searched) for co-evolutionary learning but one which we can analyze quantitatively.

3.2 Strategy Representation

For the IPD games involving deterministic and reactive strategies, several representations have been investigated in the past, which include a look-up table with bit-string encoding [12], finite state machines [54, 55, 56], and neural networks [49, 52, 60, 61]. In this paper, we use the direct look-up table strategy representation that we have introduced and studied in [49], which directly represents IPD strategy behaviors through a one-to-one mapping between the genotype space (strategy representation) and the phenotype space (behaviors).

Figure 5 illustrates the direct look-up table representation for the IPD strategies with four choices and memory-one that we consider in the case studies. The elements of the direct look-up table, $m_{ij}, i, j = 1, ..., n$, specify choices that are to be made given the inputs i (player's own previous choice) and j (opponent's previous choice). Instead of using pre-game inputs (two for memory-one strategies), the first move is specified independently, m_fm. For the *four*-choice IPD (Fig. 5), each element can take values of $+1$, $+1/3$, $-1/3$, and -1. These values are the same as those that are used to generate the payoffs in the payoff matrix (Fig. 4) through a linear interpolation (Equation 4).

A simple mutation-based variation operator is used to generate an offspring from a parent strategy when using the direct look-up table for strategy representation [49]. For the n-choice IPD, mutation replaces the original choice of a particular element in the direct look-up table with one of the other remaining $n-1$ possible choices with an equal probability of $1/(n-1)$. Each table element has a fixed probability, p_m, of being replaced by one of the remaining $n-1$ choices. The mutation-based variation operator can provide sufficient variations on strategy behaviors directly with the use of the direct look-up table representation (even for the more complex IPD game with more choices) [49].

The important advantage to using the direct look-up table representation is that the search space (given by the strategy representation) is the same as the strategy space (assuming uniform strategy distribution in the strategy space) [29]. This simplifies and allows direct investigation to be made on the conditions in a co-evolutionary learning search process that can have impact on the generalization performance. Here, variation and selection operators can directly affect the learning of strategy behaviors from the use of the direct look-up table representation. This in turn allows us to investigate the impact of both explicit

		Opponent's Previous Move			
		$+1$	$+\frac{1}{3}$	$-\frac{1}{3}$	-1
	$+1$	m_{11}	m_{12}	m_{13}	m_{14}
Player's	$+\frac{1}{3}$	m_{21}	m_{22}	m_{23}	m_{24}
Previous Move	$-\frac{1}{3}$	m_{31}	m_{32}	m_{33}	m_{34}
	-1	m_{41}	m_{42}	m_{43}	m_{44}

Figure 5: The direct look-up table representation for the deterministic and reactive memory-one IPD strategy that considers four choices (also includes m_{fm} for the first move, which is not shown in the figure).

and implicit diversity maintenance techniques on the co-evolutionary learning of strategy behaviors. Furthermore, we can measure (estimate) the generalization performance of evolved strategies using a random sample of test strategies [29].

3.3 Co-evolutionary Learning

All the co-evolutionary learning systems that we investigate share the same generate-and-test framework guided by strategic interactions through IPD game-play. The basic co-evolutionary learning system without any specific application of diversity maintenance (our baseline) is the classical co-evolutionary learning (CCL), which is implemented through the following procedure [49]:

1. Generation step, $t = 1$:
   Initialize POPSIZE/2 parent strategies, $P_i, i = 1, 2, ...,$ POPSIZE/2, randomly.
2. Generate POPSIZE/2 offspring, $O_i, i = 1, 2, ...,$ POPSIZE/2, from POPSIZE/2 parents using a variation operator.
3. All pairs of strategies compete, including the pair where a strategy plays itself (round-robin tournament). For POPSIZE strategies in a population, every strategy competes a total of POPSIZE games.
4. Select the best POPSIZE/2 strategies based on total payoffs of all games played. Increment generation step, $t \leftarrow t + 1$.
5. Step 2 to 4 are repeated until a termination criterion (a fixed number of generation) is met.

The settings for the classical co-evolutionary learning system (CCL-PM05) that we use as baseline for comparison with the co-evolutionary learning systems with specific application of diversity maintenance are summarized in Table 1.

In implementing co-evolutionary learning systems with diversity maintenance, changes are made only to the affected variation or selection operators. This is to facilitate a more direct comparison of the impact of different diversity maintenance techniques on co-evolutionary learning, whether explicitly through the variation process, or implicitly through the selection process. Note that different implicit diversity maintenance techniques would rank strategies in the population differently, and that the best-half of the ranked population are selected as parents for the following generation. The implementation details have been described earlier in Section 2, although we summarize further details such as the choices of parameters and differentiate between the various implicit diversity maintenance techniques in Table 2.

Table 1: Summary of settings used in the classical co-evolutionary learning (baseline).

Strategy Representation	Direct Look-up Table
Variation Operator	Simple mutation of direct look-up table with probability of mutation, $p_m = 0.05$
Population Structure	Single-population
POPSIZE	30
Fitness Evaluation	Round-robin tournament
Selection	Best-half of the population after ranking, (μ, λ) EP Selection Scheme
Generation	600

Table 2: Summary of implicit diversity maintenance techniques used in the experiments.

CFSD	*Competitive fitness sharing with disassortative sampling* Fitness values obtained using Equation 1 Opponent sample size of 5
CFSR	*Competitive fitness sharing with random sampling* Fitness values obtained using Equation 1 Opponent sample size of 5
IFSD	*Implicit fitness sharing with disassortative sampling* Fitness values obtained using procedure outlined in Figure 1 Opponent sample size of 3 and procedure repeated 3 times
IFSR	*Implicit fitness sharing with random sampling* Fitness values obtained using procedure outlined in Figure 1 Opponent sample size of 3 and procedure repeated 3 times
PAR	*Pareto co-evolution* Ranking based on pareto dominance relation given by Equation 2
RV	*Reduced virulence* Fitness values obtained using Equation 3 with $v = 0.75$

For case studies involving the co-evolutionary learning with speciation, we investigate both competitive and implicit fitness sharing techniques. Furthermore, we investigate two sampling procedures to obtain the set of opponents that is used for fitness evaluation. The first procedure is the *disassortative* sampling that first selects a random opponent from the current population before choosing successive opponents that are the most different (in terms of edit distance for the direct look-up table representation that we describe in more details later) from the previous opponents [28]. The second procedure is the simple random sampling [28].

For pareto co-evolution, the single-population approach that is proposed earlier in [32] is used. This is to facilitate a more direct comparison with other co-evolutionary learning systems we investigate in this study that use only one population. For the co-evolutionary learning with reduced virulence, we consider the Moderate Virulence level that has been studied in detail earlier in [24].

For case studies involving the application of explicit diversity maintenance to co-evolutionary learning, we consider the classical co-evolutionary learning approach and investigate systems with higher parameter value settings for the probability of mutation p_m. We consider the CCL with $p_m = 0.05$ (CCL-PM05) as the basic co-evolutionary learning system (baseline) for comparison with other co-evolutionary learning systems having diversity maintenance. We use other p_m settings, i.e., $p_m = 0.1$ (CCL-PM10), $p_m = 0.15$ (CCL-PM15), $p_m = 0.2$ (CCL-PM20), and $p_m = 0.25$ (CCL-PM25), to investigate the impact of explicit diversity maintenance on the generalization performance of co-evolutionary learning. In the case studies, all experiments involving different co-evolutionary learning systems are repeated in 30 independent runs to allow for statistical analysis.

3.4 Measuring Generalization Performance

We have recently formulated a theoretical framework for measuring generalization performance in co-evolutionary learning, and have applied the framework in the context of the co-evolutionary learning of IPD games. The generalization performance of a strategy is defined as its *average performance against test (opponent) strategies* [29].

In formulating the generalization performance measure for co-evolutionary learning, consider a game with the strategy space given by a set $\mathcal{S}$ of M distinct strategies, $\mathcal{S} = \{1, 2, ..., M\}$. Let the selection of individual test strategies from $\mathcal{S}$ be represented by a random variable J taking on values $j \in \mathcal{S}$ with probability $P_{\mathcal{S}}(j)$. Denote the game outcome of strategy i playing against the test strategy j by $G_i(j)$ (conversely, $G_j(i)$ is the game outcome of strategy j against strategy i).

Different definitions of $G_i(j)$ (i.e., different notions of game outcomes) for the generalization performance indicate different measures of quality. The win-lose and average-payoff functions for $G_i(j)$ are two examples. Let $g(i, j)$ and $g(j, i)$ be the payoffs for strategies i and j at the end of the IPD game, respectively. Strategy i wins the game if $g(i, j) > g(j, i)$, else i loses the game. The win-lose function for $G_i(j)$ for the IPD game is defined as:

$$G_{\text{W}}(i,j) = \begin{cases} C_{\text{WIN}} & for \quad g(i,j) > g(j,i), \\ C_{\text{LOSE}} & for \quad otherwise, \end{cases} \tag{5}$$

where C_{WIN} and C_{LOSE} are arbitrary constants with $C_{\text{WIN}} > C_{\text{LOSE}}$ (we use $C_{\text{WIN}} = 100$ and $C_{\text{LOSE}} = 0$).

In this paper, we consider the win-lose function as it simplifies the analysis for the experiments. In particular, we know the “all defect” strategy that plays full defection only to be the strategy with the maximum generalization performance when win-lose function is used irrespective of how test strategies are distributed in $\mathcal{S}$ (this is not necessarily true for other definitions such as the average-payoff function). We arbitrarily choose a stricter form of a win-lose function to measure performance (since a loss is awarded to both sides in the case of a tie).

The true generalization performance of a strategy i is defined as the mean performance against all test strategies,

$$G_i = \sum_{j=1}^{M} P_{\mathcal{S}}(j)\ G_i(j). \tag{6}$$

That is, G_i is the mean of the random variable $G_i(J)$,

$$G_i = E_{P_{\mathcal{S}}}[G_i(J)].$$

Some test strategies may be favoured over the others, or all test strategies can be considered with equal probability *a priori*. In particular, when all strategies are equally likely to be selected, i.e., when $P_{\mathcal{S}}$ is uniform, we have

$$G_i = \frac{1}{M} \sum_{j=1}^{M} G_i(j).$$

The size M of the strategy space $\mathcal{S}$ can be vary large, making direct estimation (6) of the generalization performance G_i computationally infeasible. For example, in the case of the *four*-choice IPD game involving deterministic and reactive, memory-one strategies, $M = 4^{4^2+1} = 4^{17} \approx 1.7 \times 10^{10}$.

In practice, one can estimate G_i using a random sample S_N of N test strategies drawn i.i.d. from $\mathcal{S}$ with probability $P_{\mathcal{S}}$. The estimated generalization performance of strategy i is then

$$\hat{G}_i(S_N) = \frac{1}{N} \sum_{j \in S_N} G_i(j). \tag{7}$$

If the game outcomes $G_i(J)$ vary within a finite interval $[G_{\text{MIN}}, G_{\text{MAX}}]$ of size $R = G_{\text{MAX}} - G_{\text{MIN}}$, the variance of $G_i(J)$ is upper-bounded by $\sigma^2_{\text{MAX}} = (G_{\text{MAX}} - G_{\text{MIN}})^2/4 = R^2/4$. Using Chebyshev's Theorem [62], we obtain

$$P(|\hat{G}_i - G_i| \geq \epsilon) \leq \frac{R^2}{4N \cdot \epsilon^2} \tag{8}$$

for any positive number $\epsilon > 0$. Note that the Chebyshev's bound is distribution-free, e.g., no particular form of distribution of $G_i(J)$ is assumed. Equation 8 allows one to make a statistical claim about how confident one is about the accuracy of an estimate given a random test sample of a known size N.

Let $|D_N| = |\hat{G}_i - G_i|$. The inequality (8) can be rewritten as follows:

$$P(|D_N|' \geq \epsilon') \leq \frac{1}{4N \cdot {\epsilon'}^2},$$

where $|D_N|' = |D_N|/R$ and $\epsilon' = \epsilon/R$. For the case studies, we use a sample of random test strategies to estimate the generalization performance of strategy i obtained through co-evolutionary learning, $\hat{G}_W(i)$, based on IPD game-outcomes defined by $G_{\text{W}}(i,j)$. Given the high computation cost that is involved in estimating the generalization performance through a large number of IPD game-plays, we consider a sample size of $N = 100000$. Assuming $\epsilon' = 0.01$ (error at 1% of $R = 100$), the Chebyshev's bound gives $P(|D_N|' \geq \epsilon') \leq 1/(4 \times 100000 \times 0.01^2) = 0.025$. We can claim with a probability at least $1 - 0.025 = 0.975 > 0.95$ that the error for the generalization performance estimate satisfies $|D_N|' \leq 0.01$.

In game-playing, one may be more interested in estimating the generalization performance with respect to a biased test sample, e.g., biased towards test strategies that are more likely to be encountered in a real-world scenario such as game tournaments, rather than an unbiased sample where every test strategy is equally likely to be encountered. A biased test sample may consist of "good" strategies based on some performance criteria, e.g., outperforming a large number of test strategies. This biased sample of "good" test strategies can be obtained using a procedure we have introduced and studied earlier in [29]:

1. Partial enumerative search run, $r = 1$:
   Sample PS strategies, $Q_i, i = 1, 2, ..., \mathrm{PS}$, randomly.
2. Every strategy competes with all other strategies in the sample, including its twin (i.e., each strategy competes a total of PS games).
3. Detect the strategy index $s \in \{1, 2, ..., \mathrm{PS}\}$ so that Q_s is yielding the highest total payoff. Let $Q^{(r)} := Q_s$. Increment run, $r = r + 1$.
4. Repeat steps one to three PE times to obtain PE-sized biased sample of test strategies, $Q^{(r)}, r = 1, 2, ..., \mathrm{PE}$.

The population (sample) size for each partial enumerative search is $\mathrm{PS} = 10^6$, which is significantly larger than the maximum number of unique strategies in an evolutionary run ($\mathrm{PS} > (\mathrm{generation} \times \mathrm{POPSIZE})$). This is so that the test strategy is (1) likely not to have been encountered before by evolved strategies and (2) more challenging (having outperformed a significantly larger number of opponents compared to evolved strategies). The procedure is independently repeated $\mathrm{PE} = 20$ times so that a more diverse sample of test strategies is obtained. A sample of PE test strategies obtained from a single partial enumerative search may be less diverse (they are behaviorally similar with the same weakness that can be exploited) and lower performing due to a poor population sample.

For the case studies, we estimate the generalization performance with respect to unbiased and biased samples of test strategies, which we denote as $\hat{G}_W$ and $\hat{G}^B_W$, respectively. Both estimates $\hat{G}_W(i)$ and $\hat{G}^B_W(i)$ can be computed for the current top performing evolved strategy i in a particular generation. However, analysis on the generalization performance of the co-evolutionary learning system can be made by taking generalization performance estimates on the best U evolved strategies from the population, $\mathrm{SPOP}_U = (\mathrm{spop}_1, ..., \mathrm{spop}_U)$, which is an ordered set (ordered according to their internal fitness values):

- Best,
$$\mathrm{Best}(G_{\mathrm{SPOP}_U}) = \hat{G}_{\mathrm{spop}_1}. \tag{9}$$
- Average,
$$\mathrm{Avg}(G_{\mathrm{SPOP}_U}) = \frac{1}{U}\sum_{l=1}^{U} \hat{G}_{\mathrm{spop}_l}. \tag{10}$$
- Ensemble,
$$\mathrm{Ens}(G_{\mathrm{SPOP}_U}) = \frac{1}{N}\sum_{j \in S_N} \min\Big(\sum_{l=1}^{U} G_{\mathrm{spop}_l}(j), G_{\mathrm{MAX}}\Big). \tag{11}$$

The "Ensemble" measurement allows us to estimate the generalization performance of the co-evolving population as a whole. For simplicity, we assume a form of ensemble system having a perfect gating mechanism [59, 28]: it perfectly chooses the evolved strategy in SPOP_U that best performs against each of the test strategies. For example, given five test strategies, if spop_1 outperforms the first two test strategies, and spop_2 outperforms the last three test strategies, the ensemble of $\{\mathrm{spop}_1, \mathrm{spop}_2\}$ outperforms all test strategies (it will choose spop_1 for the first two test strategies, and spop_2 for the others). As such, the "Ensemble" measurement would obtain an upper-bound on the generalization performance of the population, which provides a basis to compare the generalization performance of the population with that of an individual strategy. This analysis allows us to investigate if diversity maintenance can have an impact on the co-evolving population and subsequently be exploited for improvements in generalization performance.

For the case studies, $U = 5$ top performing individual strategies are selected from the co-evolving population. Estimates $\hat{G}_W$ and $\hat{G}_W^B$ are made for the top performing individual strategies in the population SPOP_U ($\mathrm{Best}(G_{\mathrm{SPOP}_U})$ and $\mathrm{Avg}(G_{\mathrm{SPOP}_U})$) and for the population ($\mathrm{Ens}(G_{\mathrm{SPOP}_U})$).

3.5 Measuring Diversity

Different diversity measurements obtain different information on the levels and type of variety from the population of an evolutionary system. Some diversity measures may be more important in comparison to other diversity measures for specific problems. Here, the main motivation is to obtain some general information on the diversity levels for comparison between co-evolutionary learning systems with and without diversity maintenance. For the case studies, we consider two different diversity measurements, i.e., genotypic and phenotypic diversity measurements. In an earlier study [33] that had conducted a systematic empirical investigation on various diversity measures, results suggested strong correlations between the diversity levels based on the genotypic and phenotypic diversity measurements with the fitness of the top performing solutions obtained from evolutionary search processes. The study concluded that these two diversity measurements may be useful in obtaining information on diversity levels that reflects the evolutionary search process with respect to its performance.

Genotypic diversity relates to the level of variety in the genotypic space (the search space that is specified by the solution representation). The genotypic diversity measurement that we use is based on the *edit distance* diversity measurement [63, 64, 33], which we have adapted to measure the edit distance of deterministic and reactive, memory-one IPD strategies with a direct look-up table representation[1]. The edit distance between two strategies i and j is related to the minimum number of transformations that is required to change one strategy's direct look-up table to the other strategy's:

$$\mathrm{dist}_{\mathrm{ed}}(i,j) = \frac{1}{n_e}\sum_{k=1}^{n_e} d(i_k, j_k), \tag{12}$$

where i_k and j_k are the direct look-up table elements of strategies i and j, respectively, and $d(i_k, j_k)$ is a distance metric on i_k and j_k. n_e is the total number of elements of the direct look-up table. For the *four*-choice IPD strategy, $n_e = 17$ (1 element for the first move and 16 elements for decisions based on joint past moves as shown in Fig. 5).

For the distance metric $d(i_k, j_k)$, we use the Levenshtein distance as in [65]:

$$d(i_k, j_k) = \begin{cases} 0, & i_k = j_k \\ 1, & i_k \neq j_k \end{cases} .$$

Note that a transformation for a particular element is made only as a substitution from one choice to another. For example, if two strategies i and j differ only in the choice they make for the first move, e.g., $+1$ for strategy i and $-1/3$ for strategy j, only one transformation is required to change one strategy's direct look-up table to the other strategy's (substituting $+1$ with $-1/3$ to transform the direct look-up table representation of strategy i to that of strategy j). The edit distance between these two strategies is 1/17.

---

[1] Note that since the strategy representation has a fixed length data structure, we are measuring the *Hamming distance*. However, we use the term *edit distance* in this paper based on the references of past studies [63, 64, 33] and also to differentiate with the more common usage of the term Hamming distance in EAs using binary string representation.

The genotypic diversity measurement based on edit distance on a co-evolving population can be obtained by taking the average value of pairwise distances of strategies in the population [66]:

$$\frac{2}{\text{POPSIZE}(\text{POPSIZE}-1)} \sum_{1 \leq i < j \leq \text{POPSIZE}} \text{dist}_{\text{ed}}(i, j). \tag{13}$$

Phenotypic diversity relates to the level of variety in the phenotypic space (e.g., fitness values used in the selection process of co-evolutionary learning) [33, 67]. Studies such as in [?] have investigated fingerprints of strategies based on their behavioral responses against opponents. Here, we use the simple phenotypic diversity measurement based on the *entropy* diversity measurement [33, 34]:

$$-\sum_{k} p_k \cdot \log p_k, \tag{14}$$

where p_k refers to the proportion of a particular fitness value in the population. In our case studies, the fitness value is the average IPD game payoff obtained from a round robin tournament with the entire population of co-evolving strategies.

4 Results and Discussion

In this section, we present an empirical study to investigate whether there is a relationship between generalization performance and diversity in co-evolutionary learning. Using case studies, we systematically investigate the impact of different diversity maintenance techniques to the co-evolutionary learning of the *four*-choice IPD on: (1) the generalization performance and (2) diversity. We use the results of the classical co-evolutionary learning without diversity maintenance as a baseline for comparison with results of the co-evolutionary learning systems with diversity maintenance.

We first present the results of the case studies that investigate the impact of implicit diversity maintenance on the generalization performance of co-evolutionary learning. This is followed by the results of case studies that investigate the impact of explicit diversity maintenance on the generalization performance of co-evolutionary learning. After that, we present results on the case studies that investigate the impact of implicit and explicit diversity maintenance on the diversity levels of the co-evolving population.

In general, we have made observations that are consistent with those from past studies on the impact of individual diversity maintenance techniques on the generalization performance or diversity (as some of the past studies have made actual measurements for only one quantity) for the co-evolutionary learning of the IPD game. More importantly, through the use of the case studies with common quantitative measurements for generalization performance and diversity, our results indicate that the introduction and maintenance of diversity do not necessarily lead to a significant increase in the generalization performance of co-evolutionary learning. However, our observations and subsequent analysis of results have suggested the potential that diversity can play an important role in improving the generalization performance of co-evolutionary learning if strategies can be combined in the form of an ensemble.

4.1 The Generalization Performance of Co-evolutionary Learning with Implicit Diversity Maintenance

We first present the results of the generalization performance measurements with respect to an unbiased sample, $\hat{G}_W$. We compare the generalization performance of different co-evolutionary learning with implicit diversity maintenance with the classical co-evolutionary learning without diversity maintenance (CCL-PM05) for the *four*-choice IPD. Three measurements for $\hat{G}_W$, Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}), are obtained. Table 3 summarizes the results of the measurements at the end of the co-evolutionary learning process.

Table 3: Summary of results for the generalization performance with respect to an unbiased sample given by Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) for different co-evolutionary learning systems with implicit diversity maintenance for the *four*-choice IPD taken at the final generation. For each system, the average value and the error at 95% confidence interval over 30 runs are noted. Values in bold indicate that the particular system has a significantly higher diversity level compared to that of CCL-PM05 based on the paired Wilcoxon Signed Rank test for statistical significance difference at the 0.05 level of significance.

Experiment	Best(G_{SPOP_U})	Avg(G_{SPOP_U})	Ens(G_{SPOP_U})
CCL-PM05	61.56 ± 5.91	61.17 ± 5.00	73.67 ± 4.56
CFSD	**89.98 ± 2.28**	**90.23 ± 1.27**	**95.02 ± 0.67**
CFSR	**88.64 ± 2.62**	**86.49 ± 2.14**	**95.01 ± 0.85**
IFSD	55.26 ± 7.42	53.63 ± 3.39	**91.68 ± 1.61**
IFSR	49.63 ± 6.89	50.27 ± 5.27	**79.01 ± 4.01**
PAR	69.43 ± 6.70	66.96 ± 5.11	**91.62 ± 3.37**
RV	44.89 ± 7.43	47.72 ± 7.31	68.22 ± 7.20

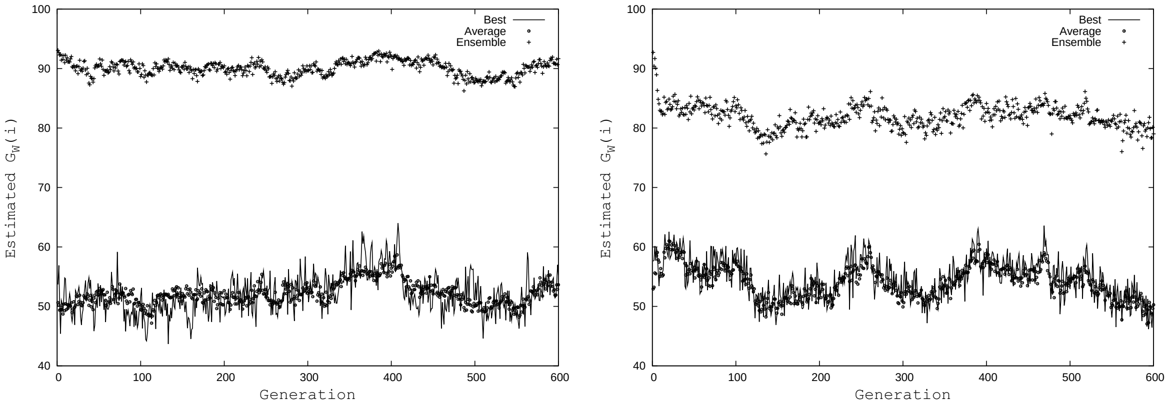
In general, mixed results are obtained for the comparison of $\hat{G}_W$ measurements of co-evolutionary learning systems with different implicit diversity maintenance techniques and the co-evolutionary learning without diversity maintenance (CCL-PM05). For example, the generalization performance of the co-evolutionary learning with reduced virulence (RV) is actually lower than that of CCL-PM05. Table 3 shows that all three measurements Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) for $\hat{G}_W$ are lower for the case of RV in comparison to the case of CCL-PM05.

However, the co-evolutionary learning with competitive fitness sharing has resulted in significantly higher generalization performance compared to CCL-PM05. For example, Table 3 shows that both CFSD and CFSR have higher generalization performance in all three measurements Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) for $\hat{G}_W$ compared to the case of CCL-PM05 that are statistically significant. In comparing between the two sampling procedures the impact on the generalization performance of co-evolutionary learning, there is little difference in the results for using disassortative sampling in CFSD and for using simple random sampling in CFSR.

For the co-evolutionary learning with implicit fitness sharing (IFSD and IFSR) and pareto co-evolution, Table 3 shows that only the comparisons for Ens(G_{SPOP_U}) measurement for $\hat{G}_W$ with that of CCL-PM05 result in statistical significant differences. For both Best(G_{SPOP_U}) and Avg(G_{SPOP_U}) measurements for $\hat{G}_W$, the comparisons between co-evolutionary learning with and without the particular implicit diversity maintenance do not reveal any statistical significant difference. The Best(G_{SPOP_U}) and Avg(G_{SPOP_U}) measurements for both IFSD and IFSR are actually lower compared to that of CCL-PM05. For PAR, Best(G_{SPOP_U}) and Avg(G_{SPOP_U}) measurements are slightly higher in comparison to that of CCL-PM05.

The results for the co-evolutionary learning with implicit fitness sharing (IFSD and IFSR) are in agreement with previous claims made in [59, 28]. Speciation can have an effect on the population of co-evolving strategies that leads to significantly higher generalization performance if the population can be combined in the form of an ensemble. The generalization performance of each individual evolved strategy is not necessarily improved, e.g., each specializes against smaller but different groups of opponents (as indicated by lower Best(G_{SPOP_U}) and Avg(G_{SPOP_U}) measurement values in Table 3). However, an ensemble of these specialist strategies has a significantly higher generalization performance if the ensemble can sufficiently often and consistently choose the best strategy from the available pool against the opponents, e.g., through a gating mechanism (as indicated by significantly higher Ens(G_{SPOP_U}) measurement values in Table 3). In addition, the effect of speciation on the co-evolving population is maintained throughout the evolutionary process for both IFSD and IFSR. Figure 6 shows that the difference between the generalization performance measurements for individual evolved strategies and the ensemble of the population is maintained on average across the generations.

By taking plots of the average results of 30 independent runs for the generalization performance measure-



(a) IFSD (b) IFSR

Figure 6: Generalization performance given by Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) measurements for co-evolutionary learning systems with implicit fitness sharing across the generations for the *four*-choice IPD. Figures (a) and (b) show the results of generalization performance during co-evolutionary learning for IFSD and IFSR, respectively. Each graph is obtained by averaging from 30 independent runs.

ments, we can observe the overall generalization performance and how it changes as a result of applying a particular diversity maintenance technique in co-evolution. For implicit fitness sharing, we observe in Figure 6 a significant difference between the generalization performance of the ensemble and individual strategies throughout co-evolution. However, we also note some small degree of local fluctuations in the generalization performance particularly for the measurement of individual strategies. This is not surprising since the implicit fitness sharing technique aims to increase diversity directly (to promote speciation in the population), which would indirectly result in an increase in the generalization performance from combining the individual strategies as an ensemble. Although diversity is maintained in the co-evolving population, it does not necessarily result in a well-speciated population of strategies from the point of view of generalization performance (as one observes from the local fluctuations in Figure 6).

Similar observations are also made for PAR in Table 3 where the generalization performance for the ensemble of the population (Ens(G_{SPOP_U})) is significantly higher compared to the generalization performance of individual evolved strategies (Best(G_{SPOP_U}) and Avg(G_{SPOP_U})). These results for PAR are consistent with the emphasis of pareto co-evolution in selecting non-dominated strategies in the co-evolving population where each would maintain performance against different opponents that are increasingly higher in quality. The pareto co-evolution would result in the evolutionary search of strategies that have more degree of specialization, in which the combination of these individual strategies in the form of an ensemble would result in higher generalization performance.

In addition, we have taken measurements of the generalization performance with respect to a biased sample, $\hat{G}_W^B$, for the various co-evolutionary learning systems with implicit diversity maintenance and the classical co-evolutionary learning without diversity maintenance. Table 4 summarizes the results of the Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) measurements for $\hat{G}_W^B$. In general, observations made from comparisons between co-evolutionary learning systems with and without implicit diversity maintenance for $\hat{G}_W^B$ are similar to observations made for the case of $\hat{G}_W$.

For example, both CFSD and CFSR resulted with statistically significant and higher generalization performance for all measurements compared to that of CCL-PM05, while all measurements of generalization performance for RV are lower compared to CCL-PM05's (Table 4). The same observation is made for PAR where the Ens(G_{SPOP_U}) measurement for $\hat{G}_W^B$ is statistically significantly higher compared to that of CCL-PM05. The only exception are the results for the co-evolutionary learning with implicit fitness sharing (IFSD and IFSR) where the Ens(G_{SPOP_U}) measurement for $\hat{G}_W^B$ is no longer statistically significantly different even though the values are still higher compared to that of CCL-PM05.

Table 4: Summary of results for the generalization performance with respect to a biased sample given by Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) for various co-evolutionary learning systems with implicit diversity maintenance for the *four*-choice IPD taken at the final generation. For each system, the average value and the error at 95% confidence interval over 30 runs are noted. Values in bold indicate that the particular system has a significantly higher diversity level compared to that of CCL-PM05 based on the paired Wilcoxon Signed Rank test for statistical significance difference at the 0.05 level of significance.

Experiment	Best(G_{SPOP_U})	Avg(G_{SPOP_U})	Ens(G_{SPOP_U})
CCL-PM05	31.37 ± 10.15	29.37 ± 8.13	40.17 ± 10.19
CFSD	**89.50 ± 6.47**	**91.43 ± 3.58**	**100.00 ± 0.00**
CFSR	**87.17 ± 8.29**	**80.77 ± 6.14**	**98.00 ± 3.05**
IFSD	22.50 ± 9.06	16.20 ± 4.10	50.00 ± 9.38
IFSR	19.83 ± 10.22	20.00 ± 8.18	43.17 ± 10.97
PAR	36.00 ± 14.15	33.80 ± 10.88	**63.83 ± 12.96**
RV	12.83 ± 6.10	15.67 ± 6.17	26.33 ± 10.35

In addition, it is noted that although the relative generalization performance of the different co-evolutionary learning systems with implicit diversity maintenance remains unchanged (i.e., they can be ranked in the ascending order of RV, IFSR, IFSD, PAR, CFSR, and CFSD), the overall generalization performance measurements are lower with respect to a biased sample of "good" test strategies compared to the measurements made with respect to an unbiased sample of test strategies. These results are consistent with our earlier conclusion in [29] where it is more difficult for evolved strategies to perform well against strong game-playing test strategies obtained from the multiple partial enumerative search.

4.2 The Generalization Performance of Co-evolutionary Learning with Explicit Diversity Maintenance

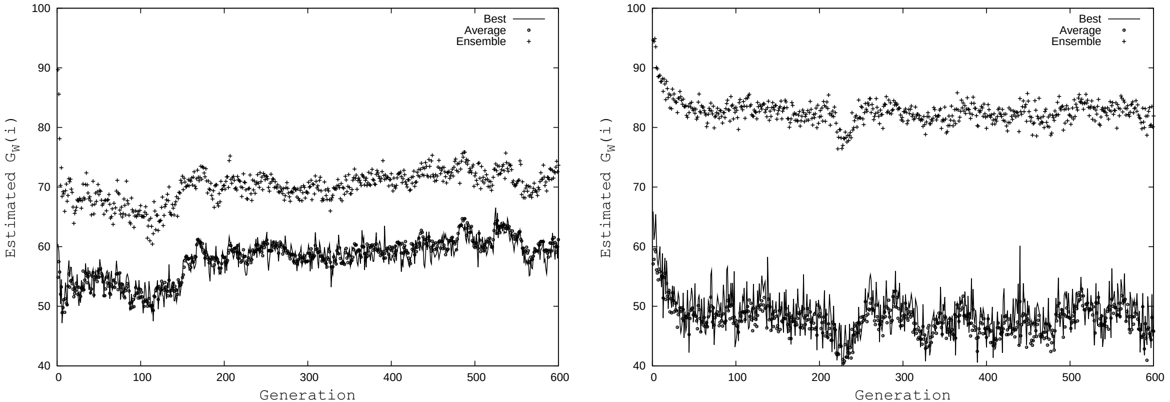
Here, we present the results of experiments that have investigated the impact of explicit diversity maintenance on the generalization performance of co-evolutionary learning. Table 5 summarizes the results of generalization performance measurements for $\hat{G}_W$ that have been made for different co-evolutionary learning systems with explicit diversity maintenance in comparison to that of CCL-PM05. In general, results show that the application of explicit diversity maintenance (by increasing the p_m setting) does not lead to an increase in the generalization performance of the top performing individual evolved strategies. Table 5 shows that Best(G_{SPOP_U}) and Avg(G_{SPOP_U}) measurements for $\hat{G}_W$ actually decrease. However, the generalization performance given by the ensemble of the population increases significantly if p_m is sufficiently high. Table 5 shows that the Ens(G_{SPOP_U}) measurements for $\hat{G}_W$ in the case of CCL-PM20 and CCL-PM25 are statistically significantly higher compared to the case of CCL-PM05.

The results show that increasing the level of variation in the population does not lead to the co-evolutionary learning of individual strategies with higher generalization performance. However, the generalization performance of the population as a whole, which can be combined in the form of an ensemble, is increased. This observation suggests that the population is effectively speciating, with the individual specialist strategies outperforming smaller and different groups of opponents, and which can be combined as an ensemble that effectively outperforms a larger number of opponents. Figure 7 illustrates this observation, where the large differences between the generalization performance for individual evolved strategies (Best(G_{SPOP_U}) and Avg(G_{SPOP_U})) and the ensemble of the population (Ens(G_{SPOP_U})) are maintained on average throughout co-evolution when p_m is significantly increased (CCL-PM25). The same observation has been made earlier for IFSD and IFSR that also speciate the population (Fig. 6).

Table 6 summarizes results for the generalization performance measurements with respect to a biased sample, $\hat{G}_W^B$, for the various co-evolutionary learning systems with explicit diversity maintenance. We observe that the application of explicit diversity maintenance to co-evolutionary learning by increasing

Table 5: Summary of results for the generalization performance with respect to an unbiased sample given by Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) for various co-evolutionary learning systems with explicit diversity maintenance for the *four*-choice IPD taken at the final generation. For each system, the average value and the error at 95% confidence interval over 30 runs are noted. Values in bold indicate that the particular system has a significantly higher diversity level compared to that of CCL-PM05 based on the paired Wilcoxon Signed Rank test for statistical significance difference at the 0.05 level of significance.

Experiment	Best(G_{SPOP_U})	Avg(G_{SPOP_U})	Ens(G_{SPOP_U})
CCL-PM05	61.56 ± 5.91	61.17 ± 5.00	73.67 ± 4.56
CCL-PM10	54.15 ± 6.37	55.74 ± 5.62	77.84 ± 4.14
CCL-PM15	51.41 ± 6.68	51.17 ± 5.81	75.09 ± 5.14
CCL-PM20	55.04 ± 5.90	53.63 ± 5.33	**82.61 ± 3.13**
CCL-PM25	46.78 ± 6.45	45.81 ± 4.07	**81.93 ± 2.45**



(a) CCL-PM05 (b) CCL-PM25

Figure 7: Comparison of generalization performance given by Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) measurements for co-evolutionary learning systems with and without explicit fitness sharing across the generations for the *four*-choice IPD. Figures (a) and (b) show the results of generalization performance during co-evolutionary learning for CCL-PM05 and CCL-PM25, respectively. Each graph is obtained by averaging from 30 independent runs.

the p_m setting can have a positive impact in increasing the generalization performance $\hat{G}_W^B$ if an ensemble can be constructed from the population. However, the increase is not statistically significant. This is not surprising as the evolved strategies are competing against a biased sample of strong game-playing strategies obtained from the multiple partial enumerative search.

Table 6: Summary of results for the generalization performance with respect to a biased sample given by Best(G_{SPOP_U}), Avg(G_{SPOP_U}), and Ens(G_{SPOP_U}) for various co-evolutionary learning systems with explicit diversity maintenance for the *four*-choice IPD taken at the final generation. For each system, the average value and the error at 95% confidence interval over 30 runs are noted. Values in bold indicate that the particular system has a significantly higher diversity level compared to that of CCL-PM05 based on the paired Wilcoxon Signed Rank test for statistical significance difference at the 0.05 level of significance.

Experiment	Best(G_{SPOP_U})	Avg(G_{SPOP_U})	Ens(G_{SPOP_U})
CCL-PM05	31.37 ± 10.15	29.37 ± 8.13	40.17 ± 10.19
CCL-PM10	16.17 ± 7.03	18.23 ± 5.83	33.33 ± 8.32
CCL-PM15	18.67 ± 8.02	17.83 ± 6.48	33.33 ± 9.20
CCL-PM20	21.83 ± 8.13	24.93 ± 5.39	48.67 ± 7.90
CCL-PM25	17.67 ± 6.12	17.97 ± 3.88	49.00 ± 8.07

4.3 The Impact of Diversity Maintenance on Diversity Levels in Co-evolutionary Learning

Results from the experiments show that not all applications of diversity maintenance techniques lead to positive and significant improvements in the generalization performance of co-evolutionary learning. However, before any conclusion can be drawn, measurements of diversity levels in the population for the different co-evolutionary learning systems with and without diversity maintenance would need to be performed. Table 7 summarizes the results of genotypic and phenotypic diversity measurements for the different co-evolutionary learning systems that are taken at the final generation. The results provide a means to determine the impact of diversity maintenance on the diversity levels in the population of co-evolutionary learning through the comparison of diversity levels between co-evolutionary learning systems with diversity maintenance and the classical co-evolutionary learning system (CCL-PM05) that is not designed specifically to introduce and maintain high diversity levels.

In general, the results on the measurements of diversity levels are consistent and in agreement with the design motivations of the respective co-evolutionary learning systems with diversity maintenance. For the co-evolutionary learning with explicit diversity maintenance, it is expected that the introduction of more variations to the evolved strategies in the population by increasing the p_m setting would lead to an increase in the genotypic diversity levels. Table 7 shows that the increase in genotypic diversity levels is statistically significant (e.g., comparing the genotypic diversity level of CCL-PM25 with that of CCL-PM05). Furthermore, the statistically significant increase that is shown in Table 7 in phenotypic diversity levels is consistent with the use of the direct look-up table strategy representation. More genotypic variation in the population of strategies would lead to more variation in the behavioral interactions, which are reflected by the higher phenotypic diversity levels.

For the co-evolutionary learning systems with implicit fitness sharing, the general increase in diversity levels (genotypic and phenotypic) is consistent with the original design motivation of speciating the population by creating niches of evolved strategies having different and unique behaviors [43, 28]. Table 7 shows that both genotypic and phenotypic diversity levels for IFSD and IFSR are statistically significantly higher compared to that of CCL-PM05. Similar observation is also made for the case of pareto co-evolution (PAR) in comparison with CCL-PM05. The higher diversity levels would be a result of the evolution and subsequent maintenance of a large set of mutually non-dominating strategies (e.g., the entire population) [32].

Table 7: Summary of genotypic and phenotypic diversities for various co-evolutionary learning systems with diversity maintenance for the *four*-choice IPD taken at the final generation. For each system, the average value and the error at 95% confidence interval over 30 runs are noted. Values in bold indicate that the particular system has a significantly higher diversity level compared to that of CCL-PM05 based on the paired Wilcoxon Signed Rank test for statistical significance difference at the 0.05 level of significance.

Experiment	Genotypic Diversity	Phenotypic Diversity
CCL-PM05	0.18 ± 0.02	1.23 ± 0.27
CCL-PM10	**0.33 ± 0.03**	**2.12 ± 0.34**
CCL-PM15	**0.41 ± 0.02**	**2.85 ± 0.17**
CCL-PM20	**0.50 ± 0.02**	**2.99 ± 0.19**
CCL-PM25	**0.59 ± 0.02**	**3.34 ± 0.02**
CFSD	**0.31 ± 0.02**	0.83 ± 0.21
CFSR	**0.36 ± 0.03**	1.20 ± 0.27
IFSD	**0.55 ± 0.03**	**2.82 ± 0.10**
IFSR	**0.37 ± 0.03**	**3.15 ± 0.07**
PAR	**0.39 ± 0.04**	**2.42 ± 0.24**
RV	0.23 ± 0.03	**1.72 ± 0.25**

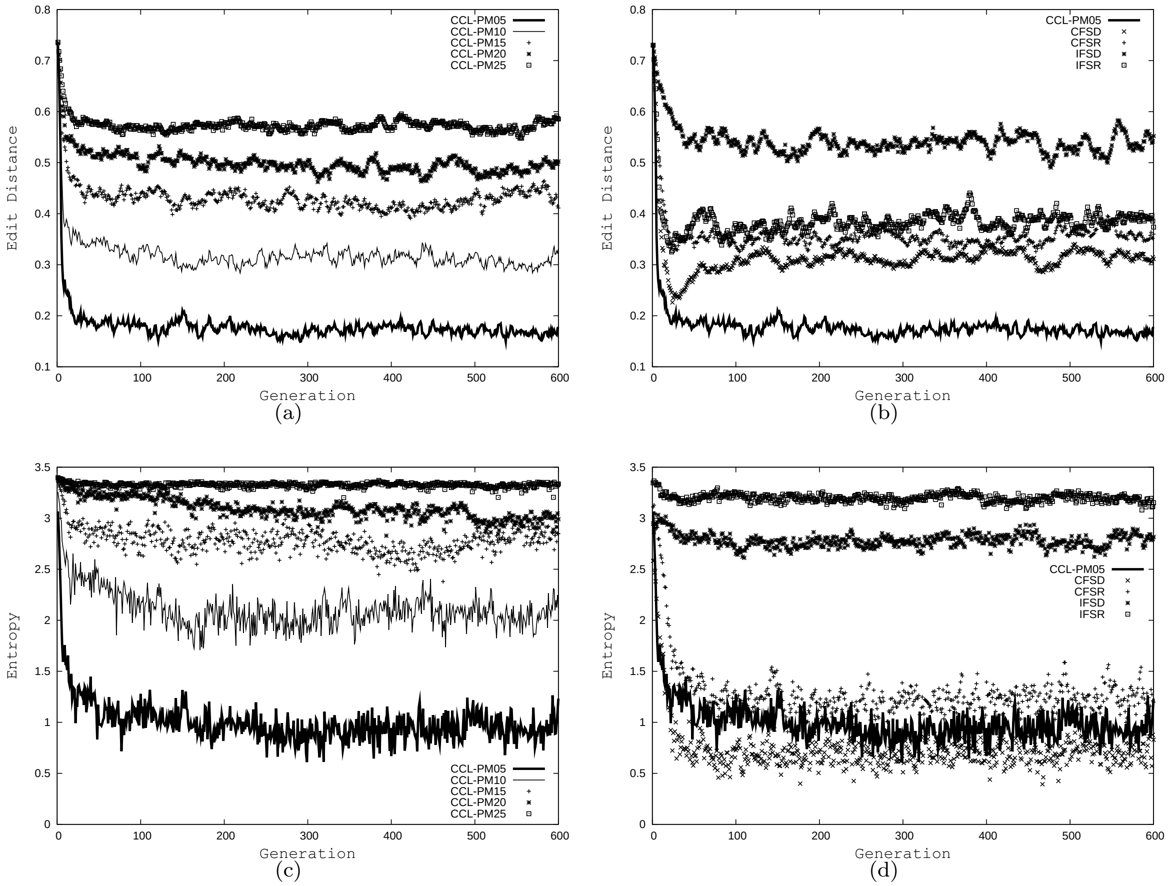
For the co-evolutionary learning with reduced virulence (RV), Table 7 shows that higher phenotypic diversity level is introduced and maintained in the population compared to that of CCL-PM05. However, RV has only slightly higher genotypic diversity level in the population compared to that of CCL-PM05. Despite this, the results are consistent with the design motivation of the technique. In particular, the technique does not seek to diversify a population into niches, but instead, aims to encourage the population to *engage* in interactions [24]. This would result in the co-evolution of a population of strategies responding differently to one another, leading to an increase in fitness variations that would be observed as an increase in the phenotypic diversity level.

The only exception is the co-evolutionary learning with competitive fitness sharing. In our case studies, the application of competitive fitness sharing does not lead to an increase in the phenotypic diversity level of the population in co-evolutionary learning. Table 7 shows that the phenotypic diversity levels for CFSD and CFSR are lower compared to that of CCL-PM05. This observation is different compared to the observation that has been reported earlier in [26]. In particular, the study uses a diversity measurement that counts the number of unique strategy responses above a *significant usage*, which is similar to the phenotypic diversity measurement. This discrepancy may be a result of applying co-evolutionary learning to different games.

Despite this, the use of competitive fitness sharing does lead to a positive and significant increase on the genotypic diversity level in the population that is shown in Table 7. These results are verified through a closer inspection of evolved strategies. It is observed that evolved strategies in the population have engaged mostly with "full defection" play, which would result in similar average payoffs that lead to low phenotypic diversity. However, observations on the representation of these evolved strategies show differences on the parts that are not accessed during game-play interactions, which nevertheless would result in higher genotypic diversity.

In comparing the relative levels of diversity that are introduced and maintained due to applications of different diversity maintenance techniques in co-evolutionary learning, both implicit fitness sharing and explicit diversity maintenance appear to be particularly effective. Figure 8 plots the average genotypic and phenotypic diversity levels of the different co-evolutionary learning systems across the generations. Both figures indicate that the implicit fitness sharing and explicit diversity maintenance (at sufficiently high p_m settings) can introduce and maintain higher diversity levels (genotypic and phenotypic) in the population throughout the evolutionary process compared to other diversity maintenance techniques.

Thus far, we have presented and discussed separately the impact of different diversity maintenance



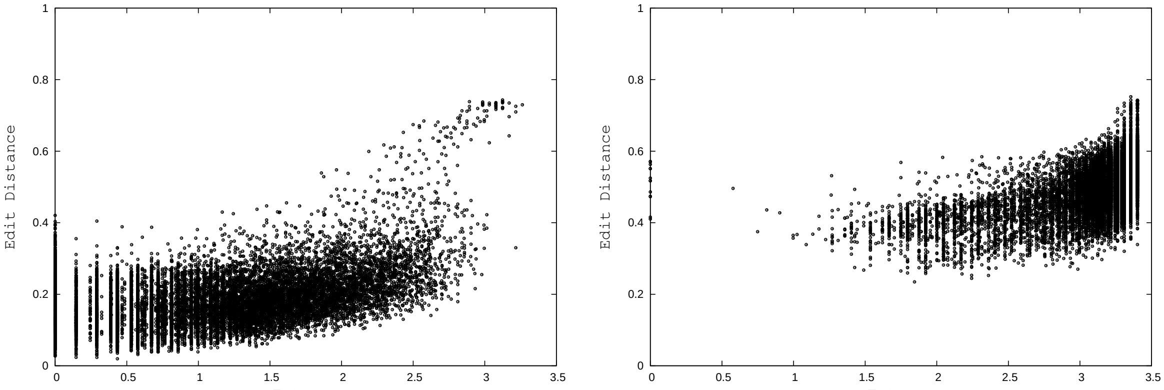
(a) (b)

(c) (d)

Figure 8: Comparison of genotypic and phenotypic diversity for various co-evolutionary learning systems across the generations for the *four*-choice IPD. Figures in (a) and (b) compare the genotypic diversity for the case of CCL-PM05 with that of co-evolutionary learning with explicit diversity maintenance (CCL-PM10, CCL-PM15, CCL-PM20, and CCL-PM25) and implicit diversity maintenance (CFSD, CFSR, IFSD, and IFSR), respectively. Figures in (c) and (d) compare the phenotypic diversity for the case of CCL-PM05 with that of co-evolutionary learning with explicit diversity maintenance (CCL-PM10, CCL-PM15, CCL-PM20, and CCL-PM25) and implicit diversity maintenance (CFSD, CFSR, IFSD, and IFSR), respectively. Each graph is obtained by averaging over 30 independent runs.

techniques in introducing and maintaining different diversity levels (genotypic and phenotypic) in co-evolutionary learning. Here, we investigate if there are relationships between genotypic and phenotypic diversity in co-evolutionary learning with diversity maintenance. Our analysis would be based on a series of scatter plots (which have been employed previously in [33] for analysis) of genotypic and phenotypic diversity. Each point in the scatter plot represents one population in a particular generation from each of the 30 independent runs. Since each co-evolutionary learning system is ran for 600 generations and repeated 30 times, the scatter plot would have a total of 600×30 points. The scatter plot may provide information on the interactions and possible relationships between genotypic and phenotypic diversity as a result of the application of specific diversity maintenance techniques in co-evolutionary learning.

We have compared the genotypic-phenotypic diversity scatter plots for co-evolutionary learning with and without explicit diversity maintenance. Figure 9 illustrates two general observations we have made. First, there is a trend of higher genotypic diversity occuring with higher phenotypic diversity for CCL-PMs at higher p_m settings (such as CCL-PM20). The higher p_m settings would introduce more genotypic variations, which in turn, would lead to more variations in the behavioral interactions (phenotypic diversity) in co-evolutionary learning. Second, CCL-PMs with higher p_m settings can better maintain both genotypic and phenotypic diversity, e.g., the points are distributed more towards the top-right region of the scatter plot for CCL-PM20 compared to that of CCL-PM05.



(a) CCL-PM05

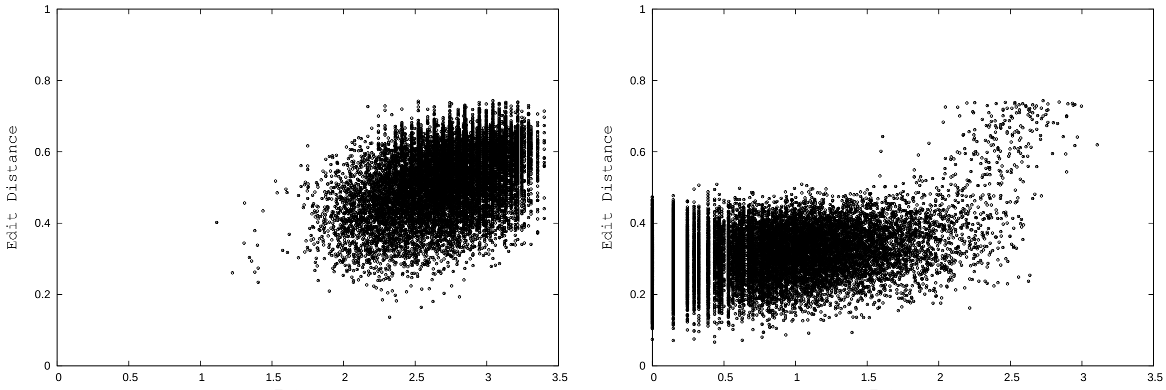
(b) CCL-PM20

Figure 9: Comparison of genotypic-phenotypic diversity scatter plots for various co-evolutionary learning systems with explicit diversity maintenance for the *four*-choice IPD. Each figure plots the population's genotypic diversity against the population's phenotypic diversity. Each point represents one population in a particular generation from each of the 30 independent runs. There are altogether 600×30 points.

We have also compared the genotypic-phenotypic diversity scatter plots for various co-evolutionary learning systems with implicit diversity maintenance. In general, we observe a trend of higher genotypic diversity occuring with higher phenotypic diversity for co-evolutionary learning with implicit fitness sharing (Fig. 10) and also pareto co-evolution. The observation for such a trend is less apparent for the case of co-evolutionary learning with reduced virulence and more so with competitive fitness sharing (Fig. 10). For our case studies, the positive relationship of genotypic and phenotypic diversity can be explained by considering the effect of speciation in the population to niches of evolved strategies having different and unique behaviors. A speciated population would have high genotypic and phenotypic diversity. The speciation effect would be very pronounced in IFSD.

4.4 Relationship between the Generalization Performance and Diversity in Co-evolutionary Learning

The results of generalization performance and diversity measurements from the case studies show that the introduction and maintenance of diversity through the applications of diversity maintenance techniques



(a) IFSD (b) CFSD

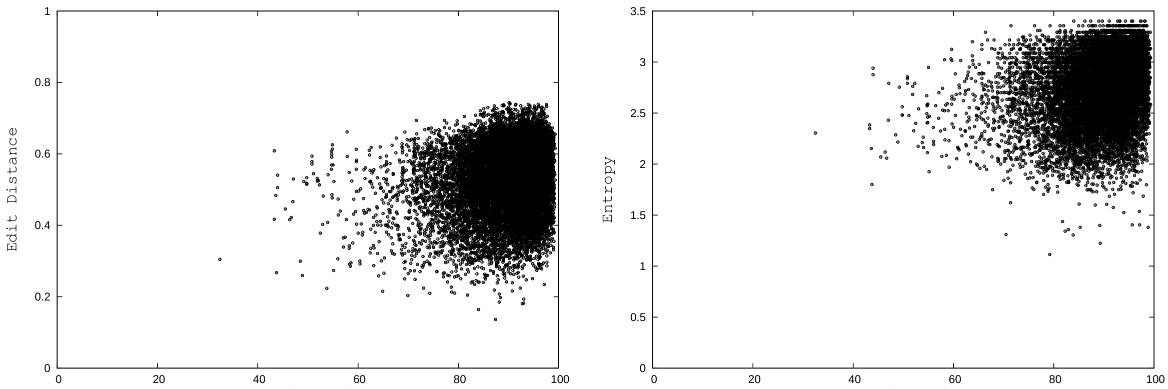
Figure 10: Comparison of genotypic-phenotypic diversity scatter plots for various co-evolutionary learning systems with implicit diversity maintenance for the *four*-choice IPD. Each figure plots the population's genotypic diversity against the population's phenotypic diversity. Each point represents one population in a particular generation from each of the 30 independent runs. There are altogether 600×30 points.

do not necessarily lead to improvements in the generalization performance of co-evolutionary learning. A case in point is the use of co-evolutionary learning with reduced virulence where the resulting generalization performance is lower compared to the classical co-evolutionary learning without diversity maintenance, regardless of whether the generalization performance measurements are made with respect to individual evolved strategies or with respect to the population where an ensemble can be constructed.

However, it is also not necessarily true that a diversity maintenance technique that does not introduce and maintain high diversity levels during co-evolution will lead to lower generalization performance. A case in point is the use of the competitive fitness sharing in co-evolutionary learning. Results have shown that the application of competitive fitness sharing significantly improves on the generalization performance in all measurements, i.e., regardless of whether an unbiased or biased sample of test strategies is used, or whether the measurements are made on individual evolved strategies or the population as a whole (the ensemble). This is despite the co-evolutionary learning systems with competitive fitness sharing not having higher diversity levels compared to systems using other diversity maintenance techniques.

The success of competitive fitness sharing in improving the generalization performance of co-evolutionary learning appears to depend on how the strategy fitness is determined using Equation 1. The emphasis of the selection process on strategies outperforming opponents that few others could, leads to the co-evolutionary learning of strategies that play "full defection" more often. Given that an "all defect" strategy would obtain the maximum possible generalization performance (defined by the win-lose outcome), a co-evolutionary learning system that evolves a population of strategies effectively playing "all defect" would have the highest generalization performance. In this case, it would appear that diversity plays little role in improving the generalization performance of co-evolutionary learning (we note the similarity in the genotypic-phenotypic diversity scatter plots for CFSD in Fig. 10(b) and for CCL-PM05 in Fig. 9(a)).

However, one general observation that we have made consistently from the case studies is that the introduction and maintenance of diversity that lead to the speciation of the population to unique niches of specialized strategies have a positive and significant impact on the generalization performance of co-evolutionary learning if an ensemble can be constructed from the population. We have conducted an analysis based on diversity-generalization scatter plots to investigate this issue. Figure 11 shows that the speciation effect of a co-evolving population is particularly effective in improving the generalization performance through ensembles of speciated strategies obtained from the application of implicit fitness sharing. One can observe a trend of high diversity (genotypic and phenotypic) occurring with high generalization performance of ensembles for co-evolutionary learning with implicit fitness sharing. Similar observations can be made to some extent for pareto co-evolution and co-evolutionary learning with explicit diversity maintenance, although it is apparent that implicit fitness sharing has a more pronounced speciation effect to co-evolutionary learning that can improve generalization performance.



(a) IFSD

(b) IFSD

Figure 11: Comparison of diversity-generalization scatter plots for various co-evolutionary learning systems with implicit fitness sharing for the *four*-choice IPD. Each figure plots the population's diversity against the population's generalization performance estimated by $\mathrm{Ens}(G_{\mathrm{SPOP}_U})$. Each point represents one population in a particular generation from each of the 30 independent runs. There are altogether 600 × 30 points.

Although overall our case studies have not been able to establish a clear positive relationship between the generalization performance (as we define it) and diversity in co-evolutionary learning, there are some diversity maintainace techniques, e.g. fittness sharing, that seem to be able to boost the generalization performance. Some past studies also suggest that diversity in the population can be exploited in co-evolutionary learning to improve "out-of-sample" performance **CITE**. In those approaches, diversity plays an important role in co-evolutionary learning in speciation to obtain a diverse set of solutions (each solution specializing with respect to different groups of test cases) that as a whole, has the potential for higher performance. The issues of quantification of population diversity and performance measurement remain important open research questions.

5 Conclusion

In this paper, we have conducted a detailed empirical study to investigate the issue of whether there is a relationship between generalization performance and diversity in co-evolutionary learning. Understanding this issue is important given that the lack of diversity in the population may have a negative impact on the generalization performance of co-evolutionary learning. More importantly, specific conditions may be identified with potential implications for design choices in co-evolutionary search algorithms whereby diversity can be exploited to improve the generalization performance in co-evolutionary learning.

We have used games as they provide a natural framework to study co-evolutionary learning. The interactions between the co-evolving solutions in the population can be framed as game-playing. This allows one to study the two main properties of a co-evolutionary learning system, i.e., decision-making and learning, in detail. With games, one is interested in the ability of the co-evolutionary learning system to search for strategies that can outperform a large number of different opponents. We have conducted case studies involving IPD games to systematically investigate the impact of various diversity maintenance techniques on the generalization performance of co-evolutionary learning. The case studies compare the co-evolutionary learning with and without diversity maintenance and address two shortcomings in previous studies through a series of quantitative measurements for both the generalization performance and diversity levels in the population of the various co-evolutionary learning systems.

We have been motivated to answer two related questions in our studies: (1) what is the right form of diversity that would be beneficial for generalization, and (2) what amount of the 'right' form of diversity is needed for good generalization. Here, the results from the case studies had shown that the introduction and maintenance of diversity in the population of co-evolutionary learning do not necessarily lead to a significant increase in the generalization performance. Although some diversity maintenance techniques, e.g. fittness sharing, have been shown to significantly improve the generalization performance of co-evolutionary learning, there are cases (e.g. reducing virulance) where a positive relationship between the generalization performance and diversity has not been observed and established in our experiments.

We also investigated the coverage of the whole strategy space by the population of co-evolved strategies. In particular, we generated a large number of test strategies and for each test strategy we checked whether the co-evolved population contains a strategy capable of winning against that test strategy. Lower values of the **???** measure would signify that the population is "homogeneous" and there is a large pool of possible test strategies for which we have not coevolved any strong opponent. On the other hand, larger values of **???** mean a good coverage of the strategy space by the co-evolved population. We observed that diversity maintainace that led to population speciation (fittness sharing and pareto co-evolution) did increase the coverage of the strategy space. For the other diversity maintainace techniques we did not observe a strong correlation between increased diversity and strategy space coverage.

However, the coverage measurements only indicate the *potential* of the co-evolved population to do well against arbitrary test strategies. Of course, given a test strategy, we would need a mechanism for choosing the "right" co-evolved strategy from the population. Many real-world problems are too complex to be tackled by any monolithic solution, e.g., automatic online trading, contract negotiation, conflict resolution, and others. In such cases, divide-and-conquer is an effective problem-solving methodology to provide a practical solution to such problems. Different individuals in the population will be specialists in dealing with different aspects of the entire problem. As a result, thepopulation as a whole will be a stronger and more complete solution. This is an important issue beyond the scope of this paper. It has received an increasing amount of attention **CITE** and it is a matter for our future research.

The study has considered different forms of fitness measurements (e.g., implicit diversity maintenance techniques) in co-evolutionary learning that are evaluated using a generalization performance measure based on the win-lose function. However, the motivation of our study is not constrained in the use of a specific choice of fitness measure in co-evolutionary learning and its evaluation by a specific generalization performance measure. We use the generalization performance measure based on the win-lose function to simplify analysis for the experiments since we know that "all defect" is the strategy with the maximum generalization performance. Other functions such as average payoffs can be used although the analysis will be much more difficult since we may not know *a priori* the strategy having maximum generalization performance. Further studies will be of interest although intuitively we would expect that some forms of diversity in co-evolutionary learning can be exploited to evolve population of strategies that can be combined in the form of ensembles having high generalization performance. It is very difficult to evolve a single strategy that performs against a large number of different opponents for any generalization performance measure.

The empirical study here is a first step towards further understanding of the relationship between the generalization performance and diversity in co-evolutionary learning. In the context of the IPD games, other frameworks for quantitative analysis involving non-local adaptation to compare performance between populations and fingerprints to measure diversity in the population can be employed to investigate further how diversity can be related to generalization performance in co-evolutionary learning. Future studies may also focus on addressing how diversity can be exploited for problem contexts other than games that allows the design of algorithmic implementations of co-evolutionary learning that can lead to improvements in generalization performance.

6 Acknowledgement

The authors would like to thank the anonymous reviewers for their comments that have helped to improve this paper. The authors are grateful to The Centre of Excellence for Research in Computational Intelligence and Applications (CERCIA), School of Computer Science, The University of Birmingham, UK, for providing computing support in running the experiments. This work is partially supported by an EPSRC grant (GR/T10671/01) to the third author.

References

[1] X. Yao. Evolutionary artificial neural networks. *Proceedings of the IEEE*, 87(9):1423–1447, 1999.

[2] X. Yao and Y. Xu. Recent advances in evolutionary computation. *Journal of Computer Science and Technology*, 21(1):1–18, 2006.

[3] X. Yao. Introduction. *Informatica (Special Issue on Evolutionary Computation)*, 18:375–376, 1994.

[4] D. B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press, Piscataway, NJ, 1995.

[5] T. Bäck, U. Hammel, and H. -P. Schwefel. Evolutionary computation: Comments on the history and current state. *IEEE Transactions on Evolutionary Computation*, 1(1):3–17, 1997.

[6] D. B. Fogel. An introduction to simulated evolutionary optimization. *IEEE Transactions on Neural Networks*, 5(1):3–14, 1994.

[7] C. A Coello Coello. A comprehensive survey of evolutionary-based multiobjective optimization techniques. *Knowledge and Information Systems*, 1(3):129–156, 1999.

[8] Y. Liu, X. Yao, and T. Higuchi. Evolutionary ensembles with negative correlation learning. *IEEE Transactions on Evolutionary Computation*, 4(4):380–387, 2000.

[9] N. García-Pedrajas, C. Hervás-Martínez, and D. Ortiz-Boyer. Cooperative coevolution of artificial neural network ensembles for pattern classification. *IEEE Transactions on Evolutionary Computation*, 9(3):271–302, 2005.

[10] D. B. Fogel, A. D. Blair, and R. Mikkulainen. Special issue: Evolutionary computation and games. *IEEE Transactions on Evolutionary Computation*, 9(6):537–539, 2005.

[11] K. Chellapilla and D. B. Fogel. Evolution, neural networks, games, and intelligence. *Proceedings of the IEEE*, 87(9):1471–1496, September 1999.

[12] R. Axelrod. The evolution of strategies in the iterated prisoner's dilemma. In L. D. Davis, editor, *Genetic Algorithms and Simulated Annealing*, chapter 3, pages 32–41. Morgan Kaufmann, New York, 1987.

[13] S. Luke and R. P. Wiegand. When coevolutionary algorithms exhibit evolutionary dynamics. In A. Barry, editor, *Workshop on Understanding Coevolution: Theory and Analysis of Coevolutionary Algorithms (GECCO'02)*, pages 236–241, 2002.

[14] E. Popovici and K. De Jong. Relationships between internal and external metrics in co-evolution. In *Proc. IEEE 2005 Congress on Evolutionary Computation (CEC'05)*, pages 2800–2807, Edinburgh, UK, September 2–5, 2005.

[15] D. Cliff and G. F. Miller. Coevolution of pursuit and evasion ii: Simulation methods and results. In *From Animals to Animats, Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior (SAB'96)*, volume 4, pages 506–515. MIT Press, 1996.

[16] S. Nolfi and D. Floreano. Co-evolving predator and prey robots: Do "arms races" arise in artificial evolution? *Artificial Life*, 4(4):311–335, 1998.

[17] S. G. Ficici and J. B. Pollack. Challenges in coevolutionary learning: Arms-race dynamics, open-endedness, and mediocre stable states. In C. Adami, R. K. Belew, H. Kitano, and C. Taylor, editors, *Artificial Life VI: Proc. of the Sixth Int. Conf. on Artificial Life*, pages 238–247, Cambridge, MA, 1998. The MIT Press.

[18] R. A. Watson and J. B. Pollack. Coevolutionary dynamics in a minimal substrate. In *Proc. of the 2001 Genetic and Evolutionary Computation Conference (GECCO'01)*, pages 702–709, California, USA, 2001.

[19] S. Y. Chong, M. K. Tan, and J. D. White. Observing the evolution of neural networks learning to play the game of othello. *IEEE Transactions on Evolutionary Computation*, 9(3):240–251, 2005.

[20] J. Noble. Finding robust texas hold'em poker strategies using pareto coevolution and deterministic crowding. In *Proc. of the 2002 International Conference on Machine Learning and Applications (ICMLA'02)*, pages 233–239, Nevada, USA, 2002.

[21] K. O. Stanley and R. Miikkulainen. Competitive coevolution through evolutionary complexification. *Journal of Artificial Intelligence Research*, 21:63–100, 2004.

[22] T. P. Runarsson and S. M. Lucas. Coevolution versus self-play temporal difference learning for acquiring position evaluation in small-board go. *IEEE Transactions on Evolutionary Computation*, 9(6):540–551, 2005.

[23] E. D. de Jong and J. B. Pollack. Ideal evaluation from coevolution. *Evolutionary Computation*, 12(2):159–192, 2004.

[24] J. Cartlidge and S. Bullock. Combating coevolutionary disengagement by reducing parasite virulence. *Evolutionary Computation*, 12(2):193–222, 2004.

[25] P. J. Darwen and X. Yao. Co-evolution in iterated prisoner's dilemma with intermediate levels of cooperation: Application to missile defense. *International Journal of Computational Intelligence and Applications*, 2(1):83–107, 2002.

[26] C. D. Rosin and R. K. Belew. New methods for competitive coevolution. *Evolutionary Computation*, 5(1):1–29, 1997.

[27] S. G. Ficici. *Solution Concepts in Coevolutionary Algorithms*. PhD thesis, Brandeis University, Massachusetts, USA, 2004.

[28] P. J. Darwen and X. Yao. Speciation as automatic categorical modularization. *IEEE Transactions on Evolutionary Computation*, 1(2):101–108, 1997.

[29] S. Y. Chong, P. Tino, and X. Yao. Measuring generalization performance in co-evolutionary learning. *IEEE Transactions on Evolutionary Computation*, 12(4):479–505, 2008.

[30] P. J. Darwen and X. Yao. On evolving robust strategies for iterated prisoner's dilemma. In *Progress in Evolutionary Computation*, volume 956 of *Lecture Notes in Artificial Intelligence*, pages 276–292. Springer, 1995.

[31] S. G. Ficici and J. B. Pollack. Pareto optimality in coevolutionary learning. In *Advances in Artificial Life: 6th European Conference (ECAL'01)*, volume 2159 of *Lecture Notes in Computer Science*, pages 316–325. Springer, Prague, Czech Republic, September 10–14, 2001.

[32] J. Noble and R. A. Watson. Pareto coevolution: Using performance against coevolved opponents in a game as dimensions for pareto selection. In *Proc. of the 2001 Genetic and Evolutionary Computation Conference (GECCO'01)*, pages 493–500, California, USA, 2001.

[33] E. K. Burke, S. Gustafson, and G. Kendall. Diversity in genetic programming: An analysis of measures and correlation with fitness. *IEEE Transactions on Evolutionary Computation*, 8(1):47–62, 2004.

[34] J. P. Rosca. Entropy-driven adaptive representation. In *Proc. Workshop Genetic Programming: From Theory to Real-world Applications*, pages 23–32, Tahoe City, CA, 1995.

[35] K. A. DeJong. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems.* PhD thesis, University of Michigan, Ann Arbor, MIchigan, USA, 1975.

[36] D. E. Goldberg and J. Richardson. Genetic algorithms with sharing for multimodal function optimization. In *Proc. Second Intl. Conf. on Genetic Algorithms and their Applications*, pages 41–49, 1987.

[37] K. Deb and D. E. Goldberg. An investigation of niche and species formation in genetic function optimization. In *Proc. Third Intl. Conf. on Genetic Algorithms*, pages 42–50, 1989.

[38] S. W. Mahfoud. *Niching Methods for Genetic Algorithms.* PhD thesis, University of Illinois, Urbana-Champaign, Illinois, USA, 1995.

[39] P. J. Darwen and X. Yao. Every niching method has its niches: Fitness sharing and implicit fitness sharing compared. In *The 4th International Conference on Parallel Problem Solving from Nature PPSN IV*, volume 1141 of *Lecture Notes in Computer Science*, pages 398–407. Springer, Berlin, Germany, September 22–26, 1996.

[40] A. Pétrowski. A new selection operator dedicated to speciation. In *Proc. Seventh Intl. Conf. on Genetic Algorithms (ICGA'97)*, pages 144–151, 1997.

[41] M. Bessaou, A. Pétrowski, and P. Siarry. Island model cooperating with speciation for multimodal optimization. In *The 6th International Conference on Parallel Problem Solving from Nature PPSN VI*, volume 1917 of *Lecture Notes in Computer Science*, pages 437–446. Springer, Paris, France, September 18–20, 2000.

[42] R. K. Ursem. Diversity-guided evolutionary algorithms. In *The 7th International Conference on Parallel Problem Solving from Nature PPSN VII*, volume 2439 of *Lecture Notes in Computer Science*, pages 462–474. Springer, Granada, Spain, September 7–11, 2002.

[43] P. J. Darwen. *Co-evolutionary Learning by Automatic Modularization with Speciation.* PhD thesis, University of New South Wales, Sydney, Australia, November 1996.

[44] R. E. Smith, S. Forrest, and A. S. Perelson. Searching for diverse, cooperative populations with genetic algorithms. *Evolutionary Computation*, 1(2):127–149, 1992.

[45] P. J. Darwen and X. Yao. A dilemma for fitness sharing with a scaling function. In *Proc. of the 1995 IEEE Conference of Evolutionary Computation (ICEC'95)*, pages 166–171, Perth, Australia, 1995.

[46] R. Axelrod. *The Evolution of Cooperation.* Basic Books, New York, 1984.

[47] R. Axelrod. Effective choice in the prisoner's dilemma. *The Journal of Conflict Resolution*, 24(1):3–25, March 1980.

[48] R. Axelrod. More effective choice in the prisoner's dilemma. *The Journal of Conflict Resolution*, 24(3):379–403, September 1980.

[49] S. Y. Chong and X. Yao. Behavioral diversity, choices, and noise in the iterated prisoner's dilemma. *IEEE Transactions on Evolutionary Computation*, 9(6):540–551, 2005.

[50] P. J. Darwen and X. Yao. Why more choices cause less cooperation in iterated prisoner's dilemma. In *Proc. IEEE 2001 Congress on Evolutionary Computation (CEC'01)*, pages 987–994, Seoul, Korea, May 27–30, 2002.

[51] X. Yao and P. J. Darwen. How important is your reputation in a multi-agent environment. In *Proc. IEEE 1999 Conference on Systems, Man, and Cybernetics (SMC'99)*, pages 575–580, Tokyo, Japan, October 12–15, 1999.

[52] P. J. Darwen and X. Yao. Does extra genetic diversity maintain escalation in a co-evolutionary arms race. *International Journal of Knowledge-Based Intelligent Engineering Systems*, 4(3):191–200, July 2000.

[53] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[54] D. B. Fogel. The evolution of intelligent decision making in gaming. *Cybernetics and Systems: An International Journal*, 22:223–236, 1991.

[55] D. B. Fogel. Evolving behaviors in the iterated prisoner's dilemma. *Evolutionary Computation*, 1(1):77–97, 1993.

[56] D. B. Fogel. On the relationship between the duration of an encouter and the evolution of cooperation in the iterated prisoner's dilemma. *Evolutionary Computation*, 3(3):349–363, 1996.

[57] B. A. Julstrom. Effects of contest length and noise on reciprocal altruism, cooperation, and payoffs in the iterated prisoner's dilemma. In *Proc. 7th International Conf. on Genetic Algorithms (ICGA'97)*, pages 386–392, San Francisco, CA, 1997. Morgan Kauffman.

[58] S. Y. Chong, J. Humble, G. Kendall, J. Li, and X. Yao. Learning ipd strategies through co-evolution. In G. Kendall, X. Yao, and S. Y. Chong, editors, *The Iterated Prisoners' Dilemma: 20 Years On*, chapter 3, pages 63–87. World Scientific, Singapore, 2007.

[59] X. Yao, Y. Liu, and P. J. Darwen. How to make best use of evolutionary learning. In R. Stocker, H. Jelinck, B. Burnota, and T. Bossomaier, editors, *Complex Systems - From Local Interactions to Global Phenomena*, pages 229–242. IOS Press, Amsterdam, 1996.

[60] P. G. Harrald and D. B. Fogel. Evolving continuous behaviors in the iterated prisoner's dilemma. *BioSystems: Special Issue on the Prisoner's Dilemma*, 37:135–145, 1996.

[61] N. Franken and A. P. Engelbrecht. Particle swarm optimization approaches to coevolve strategies for the iterated prisoner's dilemma. *IEEE Transactions on Evolutionary Computation*, 9(6):562–579, 2005.

[62] B. V. Gnedenko and G. V. Gnedenko. *Theory of Probability.* Taylor & Francis, 1998.

[63] A. Ekárt and S. Németh. A metric for genetic programs and fitness sharing. In *Proc. of the European Conf. on Genetic Programming EuroGP 00*, volume 1802 of *Lecture Notes in Computer Science*, pages 259–2701. Springer, Edinburg, UK, April 15–16, 2000.

[64] E. D de Jong, R. A. Watson, and J. B. Pollack. Reducing bloat and promoting diversity using multi-objective methods. In *Proc. Genetic Evolutionary Computation Conf.*, pages 11–18, 2001.

[65] U. -M. O'Reilly. Using a distance metric on genetic programs to understand genetic operators. In *Proc. IEEE 1997 Intl. Conf. on Systems, Man, and Cybernetics*, pages 4092–4097, Piscataway, NJ, 1997. IEEE Press.

[66] A. Ekárt and S. Németh. Maintaining the diversity of genetic programs. In *Proc. of the 5th European Conf. on Genetic Programming EuroGP 02*, volume 2278 of *Lecture Notes in Computer Science*, pages 162–171. Springer, Kinsale, Ireland, April 3–5, 2002.

[67] S. Gustafson. *An Analysis of Diversity in Genetic Programming.* PhD thesis, University of Nottingham, Nottingham, UK, 2004.