Quantitative Analysis of the Hall of Fame Coevolutionary Archives

Paweł Liskowski
Institute of Computing Science, Poznan University of Technology
Piotrowo 2, 60965 Poznań, Poland
pliskowski@cs.put.poznan.pl

ABSTRACT

This paper provides an attempt to investigate the properties of the Hall of Fame archive in two-population competitive coevolution environment applied to the game of Othello. Using the measure of expected utility, a round-robin tournament and performance profiles, we show that coevolution can be biased towards playing better with stronger opponents if it is driven by interactions with the past champions kept in the archive, in addition to pure competition among coevolving individuals. Moreover, the Hall of Fame does not necessarily influence the overall perfromance in terms of expected utility, as it trades-off the ability to cope with opponents of various strength, so that the resulting players are more likely to win with a strong opponent than with a weak one.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning; I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search—*Heuristic methods*

General Terms

Algorithms, Experimentation

Keywords

Competitive Coevolution, Performance Profile, Maximization of Expected Utility, Strategy Learning, Othello

1. INTRODUCTION

Coevolutionary algorithms intend to model interactions between individuals as observed in nature to imitate the fitness evaluation function. Specifically, individual's fitness depends there on other individuals, and is thereby *subjective*. Of several genres of coevolutionary algorithms, competitive coevolution is particularly useful when the evaluation function is unknown or difficult to compute, fitting problems from the interactive domain such as game playing. By engaging the players in population in the mutual pressure to outperform each other, coevolutionary learning provides an adaptive gradient that might otherwise be hard to obtain [8].

However, the use of a subjective fitness measure in a selection process is known to contribute to the occurrence of certain *pathologies*, which prevent coevolution from producing high-quality solutions. Common variants of those pathological behaviors include disengagement, cyclic dynamics, overspecialization and forgetting [7, 1, 11]. Jointly, those harmful phenomena constitute a major challenge in the design of coevolutionary algorithms.

One of the methods that aims at counteracting the negative impact of coevolutionary pathologies are coevolutionary archives. An archive is typically a collection of individuals collected from multiple generations of an evolutionary run, and serves as a means of ensuring a stable basis for evaluation and selection. Thus, it can be considered as a variant of memory aimed at sustaining the continuous search progress by enforcing interactions with the archive members.

Even though the archive mechanisms such as the Hall of Fame (HoF) [9, 6, 5] have been studied in the past and are known to maintain progress in an evolutionary arms race, they do not provide any specific guarantees in terms of convergence to the optimal solution. Moreover, the characteristics of archive's influence on evolving individuals is little known, which constitutes the main motivation for this study.

In this paper we analyze the impact of coevolutionary archives, Hall of Fame in particular, on two-population competitive coevolution, applied to learning strategies for the game of Othello. To this end, we employ not only the overall players' performance analysis such as the measure of expected utility and the round-robin tournament, but also the performance profiles, a tool introduced in [4] that characterizes the strategies using a range of variably skilled opponents.

2. METHODS

2.1 Othello and the WPC representation

The interactive domain we consider in this study is the game of Othello, one of the most challenging board games played by two opponents on an 8×8 board.

We adopt the position-weighted piece counter (WPC) as strategy representation. WPC is a linear weighted board evaluation function that assigns weight w_i to board location i and uses the scalar product to calculate the value f of a

GECCO'13, July 6–10, 2013, Amsterdam, The Netherlands.

board state $\mathbf{b}$:

$$f(\mathbf{b}) = \sum_{i=1}^{8\times 8} w_i b_i,$$

where b_i is 0 in case of an empty location, -1 in case of a white piece or $+1$ if a black piece is present. The players interpret $f(\mathbf{b})$ in a complementary manner: the black player prefers moves leading towards states with a higher value, while lower values are favored by the white player.

All methods considered in this paper employ WPC as a state evaluator in 1-ply setup. WPC is used to represent the candidate solutions to the learning problem as well as the opponent strategies employed by a fitness function. Other details on Othello rules and WPC strategy representation can be found in [10].

2.2 Archives

Coevolutionary archives aim to improve reliability and sustain the overall progress during the search process. Typically, an archive is defined as a set of top-performing individuals encountered during the evolutionary process. The most common archives employed in the literature can be described as *best-of-generation* models as they consist of the fittest individuals of the m most recent generations. Archives which are explicitly used for evaluation purposes are also known as Hall of Fame archives [9]. Keeping a sample of individuals in the archive allows to achieve a sense of historical progress by testing current solutions against those already present in the archive [3]. More specifically, aside from explicit interactions between individuals for fitness assessment, a part of an individuals' fitness results from the outcome of interactions with some subset of individuals from the archive. In this way, individual's fitness is to some extent determined by confronting some of the past 'champions'.

3. THE EXPERIMENT

The primary objective of the experiment is to investigate the impact of using the Hall of Fame archive in the two-population coevolutionary setting on the performance of evolved Othello players. To this aim, we employ the expected utility performance measure and the round-robin tournament. Our second goal is to explain the characteristics of archives' influence by profiling the strategies using opponents of varying difficulty.

To ensure a fair comparison between the algorithms, we set them up in such a manner that the computational effort per one generation is equal for all methods. The computational effort is identified with the number of interactions between individuals. Each evolutionary run consists of 200 generations, and the total effort is $1,000,000$ games per run.

A single interaction between individuals is a double game, i.e., both individuals play two games, once as the black and once as the white player. After a playing a game the winner gets 1 point and the loser 0 points. In case of a draw the competitors get 0.5 point each.

All methods start with an initial population filled with individuals whose WPC weights are randomly drawn from $[-0.2, 0.2]$. They also share the same mutation operator, which is the only search operator used in this study. It perturbs the weight $w_i^{'}$ of the offspring by adding a small random value to the corresponding weight of the parent:

$$w_i^{'} = w_i + k \cdot U_i,$$

Table 1: **Subjective** best-of-run individuals.

Algorithm	Performance [%]
CEL-T-HOF	86.62 ± 0.54
CEL-T	86.55 ± 0.54
CEL-HOF	84.85 ± 0.82
CEL	81.11 ± 0.99

where $k = 0.1$ is a scaling constant and U_i is a real number drawn uniformly from $[-1, 1]$, independently for every i. Weights resulting from mutation are bounded by $[-10, 10]$.

Some of the setups and performance assessment methods employ *random WPC players*. Such players are obtained by drawing weights uniformly and independently from the interval $[-10, 10]$. In the following, by 'random player/opponent' we mean a WPC player obtained in this way.

3.1 CEL setups

In this study, we consider two variants of coevolutionary learning that vary only in the selection method.

CEL is a two-population competitive coevolutionary algorithm in which individuals (here: WPC strategies) are bred separately in two populations. One population contains candidate solutions, while the other one maintains tests, i.e., the opponent strategies that challenge the candidate solutions. Both populations use the same selection scheme of $(\mu + \lambda)$ evolutionary strategy, where $\mu = 25$ and $\lambda = 25$. In **CEL-T** however, the population of tests uses a variant of (μ, λ) evolution strategy where the $\mu = 40$ best performing strategies form the basis of a new population, while $\lambda = 10$ individuals are discarded and drawn anew from the search space. This is the only difference between these two algorithms.

The fitness of a candidate solution is the sum of points obtained by confronting all tests. Tests, in contrast, are rewarded for making *distinctions* between candidate solutions [2]. A test makes a distinction for a given pair of candidate solutions if the games it plays with them result in different outcomes. To promote diversity in the population of tests, we employ *competitive fitness sharing* [9]: each point for a specific distinction is weighted by the inverse of the number of tests that make that distinction. As a result, tests that make unique distinctions are unlikely to be lost.

CEL-HOF and **CEL-T-HOF** extend their corresponding base setups CEL and CEL-T with the Hall of Fame archive. The fitness of a candidate solution is the sum of points obtained in 25 double games with tests and 25 double games with opponents from the archive, so that its fitness is determined by the outcomes of 50 games, as in CEL. To maintain the same computational effort per generation, the population of tests is limited to 25 individuals.

4. RESULTS

We performed 30 runs for each method. In the following, the best-of-generation individual is the individual with the highest fitness in the population of candidate solutions. By the best-of-run player we mean the best-of-generation player of the last generation.

4.1 Performance Comparison

To objectively assess the individuals we use the approximated measure of expected utility. An individual to be

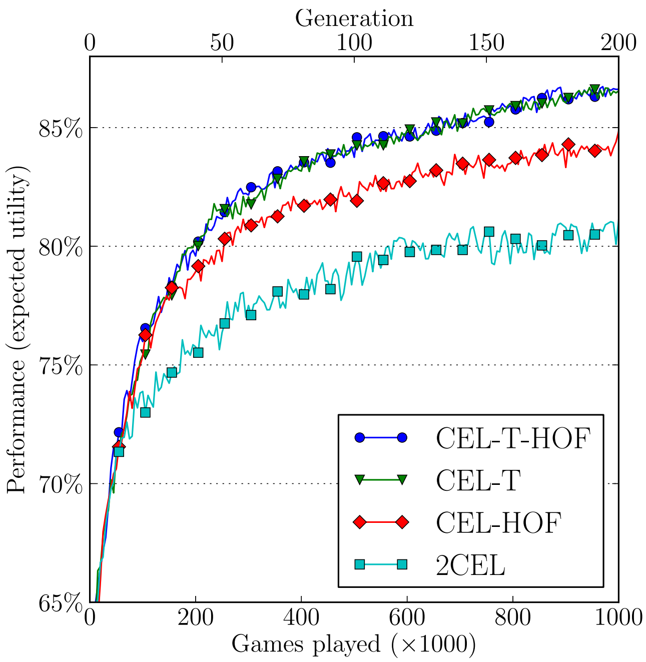

Figure 1: Performance of players over time

assessed plays 25,000 double games (50,000 games in total) against random WPC players, obtained by drawing weights uniformly from the interval $[-10, 10]$. With one point for winning the game, zero for losing, and 0.5 for a draw, the expected utility of a player ranges from 0 to 1, but for clearer presentation we report it in percent points. From now on, the term 'performance' refers to this measure.

Figure 1 shows how the performance of each method changes as a function of the computational effort. Each point on the plot is the performance of the best-of-generation player averaged over 30 runs. Table 1 compares the methods in terms of the average performance of the best-of-run individuals accompanied by 95% confidence intervals.

Forcing the candidate solutions to compete against the players in archive proved to be particularly beneficial for CEL. This is particularly evident in the course of curves in Fig. 1. However, this does not hold when comparing the CEL-T setups. The difference is much less apparent than in the case of CEL, and it is clear that HoF archive does not help to achieve a higher level of players in this case.

4.2 Round-Robin Tournament Comparison

In order to measure the relative performance of methods, we employ a round-robin tournament involving teams of best-of-generation individuals obtained from the studied learning algorithms. The best-of-run strategies which were subject to the expected utility assessment in the previous experiment are now gathered into teams, each representing the method they originate from. In this way we hope to gain better insight into method performance, because the randomly generated opponents used for expected utility assessment in the previous experiment cannot be expected to represent a rich repertoire of behaviors. Regarding the tournament organization, when two teams are confronted, each team member plays against all members from the other team and the final score is the overall sum of points obtained by all members of the team.

Table 2 presents the results of the tournament. To correctly interpret these results, we must first realize that contrary to the previous performance comparison which involved random opponents, the round-robin tournament engages only players that evolved to be strong. In this light, it is not surprising that the ranking of methods in terms of the expected utility presented in Table 1 is different than the one resulting from the round-robin tournament. However, both quality measures collectively indicate the positive influence of the Hall of Fame archive. CEL-HOF is again clearly better than CEL. Nevertheless, CEL-T-HOF is by far the best algorithm in terms of the round-robin performance. It wins against every other method in a series of head-to-head matches. This stays in contrast with the results obtained in the previous experiment, where we found out that the Hall of Fame archive is not beneficial in terms of expected utility for CEL-T-HOF. This raises the question whether the Hall of Fame archive changes the characteristics of the players by biasing them towards playing better with stronger opponents. We will try to verify this hypothesis in the following section.

4.3 Analysis with the Performance Profiles

In order to better understand the influence of the Hall of Fame archive we employ the performance profiles, a tool proposed for the first time in [4], which provides insight into how a particular strategy copes with opponents of different strength. In order to prepare a performance profile, we randomly generate 500,000 opponents by sampling WPC weights uniformly and independently from the $[-10, 10]$ interval. Subsequently, the performance of each generated opponent is estimated by playing 1,000 double games with random WPC strategies (generated individually and independently for each double game). The range of the possible performance values, i.e., $[0, 1]$, is then divided into 100 bins of equal width, and each opponent is assigned to one of these bins based on its performance. To inspect the performance profile of a method, each of the best-of-run individuals plays double games with all opponents from each bin, and the average game outcome is plotted against the bins.

The profiling results shown in Fig. 2 demonstrate that CEL-HOF dominates CEL in each bin of the plot. This confirms the general observation from the previous experiments that the archive helps to improve the overall performance. However, a detailed analysis of Fig. 2 reveals some of the less obvious properties of the Hall of Fame archive. While the difference between CEL-T-HOF and CEL-T is negligible for the moderately skilled opponents (performance around 50%), it becomes noticeable towards the extremes of the plot. This discrepancy is particularly evident on the right side of the plot and increases with the strength of opponents. This explains the supreme results of the archive-based algorithms in the round-robin tournament, as it favors players who excel in competition with relatively strong opponents. Thus, guiding the search using games with not only the coevolving individuals but also with the past champions biases the candidate solutions towards playing better with stronger opponents. However, as the strong opponents are less frequent than the weak ones, the advantage of CEL-T-HOF over CEL-T in terms of the approximated expected utility

Table 2: Round-robin tournament

Method	CEL-T-HOF	CEL-HOF	CEL-T	CEL	Total
CEL-T-HOF	-	55.1%	54.7%	56.4%	55.4%
CEL-HOF	44.9%	-	49.8%	52.4%	49.1%
CEL-T	45.3%	50.2%	-	50.9%	48.8%
CEL	43.6%	47.6%	49.1%	-	46.7%

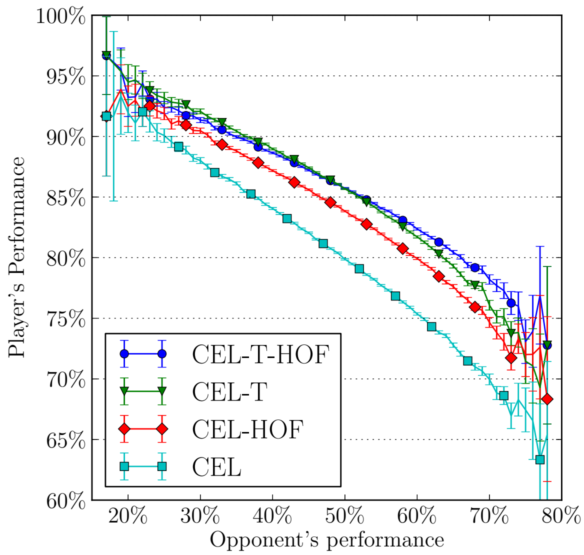

Figure 2: Performance profiles of considered methods

was less prominent (cf. Table 1). Interestingly, it is worth noting that in case of CEL-T-HOF, the Hall of Fame archive slightly impairs player's ability to cope with weaker opponents. This can be observed on the left side of the plot, as its curve is marginally below CEL-T for opponents whose performance is less than approx. 38%.

Overall, the experimental results suggest that the Hall of Fame archive biases the learning towards competition with stronger opponents in case of an algorithm that generalizes better in terms of expected utility. On the other hand, for weaker algorithms such as CEL it essentially helps to improve the expected utility of evolved players.

5. CONCLUSION

In this paper we studied the Hall of Fame archive and it's properties in the two-population coevolutionary environment. The archive-based algorithms proved to be superior to their archive-free counterparts in terms of the round-robin tournament, where individuals face the well-performing strategies. We used the performance profile tool to explain the differences between the studied algorithms in terms of the round-robin tournament and the expected utility. Our final conclusion is that if coevolution is driven, in addition to pure competition among coevolving individuals, by interactions with well-performing historic individuals (kept in the archive), it may produce strategies that are biased towards playing well with the relatively stronger opponents. However, one can also suspect that such a bias could have a negative impact on the overall performance of the evolved players, but there is no evidence for this at the moment.

6. REFERENCES

[1] S. G. Ficici and J. B. Pollack. Challenges in coevolutionary learning: Arms-race dynamics, open-endedness, and mediocre stable states. In *Proceedings of the Sixth International Conference on Artificial Life*, pages 238–247. MIT Press, 1998.

[2] S. G. Ficici and J. B. Pollack. Pareto optimality in coevolutionary learning. In J. Kelemen and P. Sosík, editors, *Advances in Artificial Life, 6th European Conference, ECAL 2001*, volume 2159 of *Lecture Notes in Computer Science*, pages 316–325, Prague, Czech Republic, 2001. Springer.

[3] S. G. Ficici and J. B. Pollack. A game-theoretic memory mechanism for coevolution. In E. C.-P. et al., editor, *Genetic and Evolutionary Computation - GECCO 2003*, volume 2723 of *Lecture Notes in Computer Science*, pages 286–297, Chicago, IL, 2003. Springer.

[4] W. Jaśkowski, P. Liskowski, M. Szubert, and K. Krawiec. Improving Coevolution by Random Sampling. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2013)*, 2013.

[5] K. Krawiec, W. Jaśkowski, and M. Szubert. Evolving small-board go players using coevolutionary temporal difference learning with archive. *International Journal of Applied Mathematics and Computer Science*, 21(4):717–731, 2011.

[6] M. Nogueira, C. Cotta, and A. J. Fernández-Leiva. An analysis of hall-of-fame strategies in competitive coevolutionary algorithms for self-learning in rts games. *Learning and Intelligent OptimizatioN Conference (LION 7)*, 2013.

[7] J. Paredis. Coevolving cellular automata: Be aware of the red queen. In *Proceedings of the Seventh International Conference on Genetic Algorithms*, pages 393–400, 1997.

[8] J. B. Pollack and A. D. Blair. Co-evolution in the successful learning of backgammon strategy. *Machine Learning*, 32(3):225–240, 1998.

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

[10] M. Szubert, W. Jaśkowski, and K. Krawiec. Coevolutionary temporal difference learning for othello. In *IEEE Symposium on Computational Intelligence and Games*, pages 104–111, Milano, Italy, 2009.

[11] R. A. Watson and J. B. Pollack. Coevolutionary dynamics in a minimal substrate. In L. S. et al., editor, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, pages 702–709, San Francisco, California, USA, 7-11 July 2001. Morgan Kaufmann.