Age-Fitness Pareto Optimization

Michael D. Schmidt
Computational Synthesis Lab
Cornell University
Ithaca, NY 14853
mds47@cornell.edu

Hod Lipson
Computational Synthesis Lab
Cornell University
Ithaca, NY 14853
hod.lipson@cornell.edu

ABSTRACT

We propose a multi-objective method for avoiding premature convergence in evolutionary algorithms, and demonstrate a three-fold performance improvement over comparable methods. Previous research has shown that partitioning an evolving population into age groups can greatly improve the ability to identify global optima and avoid converging to local optima. Here, we propose that treating age as an explicit optimization criterion can increase performance even further, with fewer algorithm implementation parameters. The proposed method evolves a population on the two-dimensional Pareto front comprising (a) how long the genotype has been in the population (age); and (b) its performance (fitness). We compare this approach with previous approaches on the Symbolic Regression problem, sweeping the problem difficulty over a range of solution complexities and number of variables. Our results indicate that the multi-objective approach identifies the exact target solution more often that the age-layered population and standard population methods. The multi-objective method also performs better on higher complexity problems and higher dimensional datasets – finding global optima with less computational effort.

Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search

General Terms

Algorithms, Design, Performance, Reliability

Keywords

Age, Pareto, Evolutionary Algorithms, Symbolic Regression

1. INTRODUCTION

A common problem in many applications of evolutionary algorithms is when the progress of the algorithm stagnates and solutions stop improving. Expending additional computational effort in the evolution often fails to make any substantial progress. This problem is known as *premature convergence* [1, 2].

A common method for dealing with premature convergence is to perform many evolutionary searches, randomizing and restarting the search multiple times [3]. This approach can be wasteful however, as the entire population is repeatedly thrown out. There is also the difficulty of deciding when to restart.

GECCO'10, July 7–11, 2010, Portland, Oregon, USA.
ACM 978-1-4503-0072-8/10/07.

One of the best performing methods in the genetic programming literature for addressing premature convergence is the Age-Layered Population Structure (ALPS) method [4]. ALPS uses a special notion of *age* – how long genotypic material has existed in the population – in order to partition the evolving population into age layers (see Figure 1). Random individuals are inserted into the youngest population layer.

Here, we consider using the ALPS concept of age as a fundamental property in the evolutionary optimization. Rather than using age to partition the population into layers, we use age as an independent dimension in a multi-objective Pareto front optimization. In this context, a solution is selected for if it has both higher fitness and lower genotypic age than other solutions.

As in the ALPS method, random individuals are added into the population at each generation. Rather than flowing up the age layers, they flow through a two-dimensional space of fitness and age (see Figure 1). Young solutions exist in the same population as the oldest and most fit, but persist because they are non-dominated on the age dimension of the Pareto space.

The following sections describe the proposed method and our primary results. See [5] for our complete description and analysis.

2. ALGORITHM

The age of a solution is measured in generations. All randomly initialized individuals start with age of one. With each generation an individual exists in the population, its age is incremented by one. During crossover and mutation events, the age is inherited as the maximum age of the parents [6].

The Age-Fitness Pareto Population method uses a single population, in contrast to the population layers in the ALPS algorithm. The algorithm tracks the fitness of each individual as in a normal evolutionary algorithm, and also the genotypic age.

The individuals in the population can be thought of lying on a two-dimensional plane of age and fitness, as in Figure 1. The multi-objective optimization task is to identify the non-dominated Pareto front of the problem domain [7]; here, the objectives are to maximize the fitness with minimum age.

3. EXPERIMENTAL SETUP

We perform identical experiments on three algorithms: (1) the ALPS algorithm [4], (2) the proposed Age-Fitness Pareto algorithm, and (3) the Deterministic Crowding algorithm [8], a well established diversity-maintenance method.

We experimented on the Symbolic Regression problem. Symbolic regression [9] is the problem of identifying the simplest equation that most accurately fits a given set of data. We used the symbolic regression algorithm described in [10] as the basis for our implementation. We simply swap out the population representation and selection for the three compared algorithms.

We tested each algorithm on 1000 randomly-generated symbolic regression problems. Each evolutionary search was performed on a single quad-core computer. Evolution was stopped if the algorithm identified a zero error solution on the validation data set (i.e. less than 10^{-3} normalized mean absolute error), or when the algorithm reached one million generations.

4. EXPERIMENTAL RESULTS

Figure 2 also shows the rate that each algorithm identifies the exact target solution. All algorithms show the standard *s*-shaped convergence rates where computational effort increases greatly for the hardest of the test problems. Late in the searches, the algorithms begin to diverge at different rates of finding the exact solution. The Age-Fitness Pareto algorithm performed the best, finding the exact solution approximately 5% more often than the ALPS algorithm.

Importantly, Figure 2 further demonstrates that the hardest problems solved by ALPS were solved by the Age-Fitness Pareto algorithm using a third of the computational effort.

The deterministic crowding algorithm, with the added randomized individual per generation, performed worst of the three algorithms. Here, deterministic crowding identified the exact target solution approximately 5% less often than the ALPS algorithm, and approximately 10% less often than the Age-Fitness Pareto algorithm.

5. CONCLUSION

Results on randomly generated symbolic regression problems indicate that the age-fitness multi-objective approach finds the exact target solution more often than previous methods over a range of target problem complexities and dataset dimensions. This approach can be readily incorporated into other evolutionary algorithms, as it makes no assumptions about the problem or solution representations.

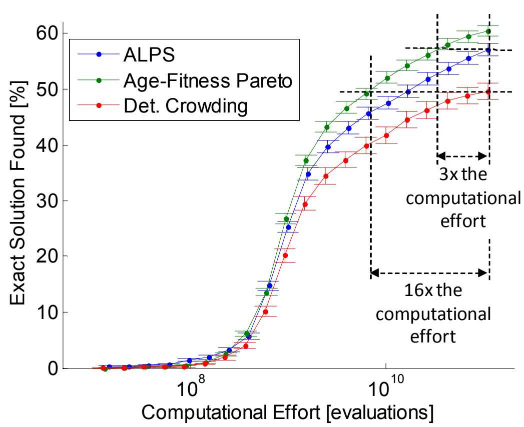

Figure 2. The convergence rate to the exact solution of the compared algorithms versus the total computational effort of the evolutionary search. Convergence to the exact solution is percent of the trials which reach epsilon error on the validation data set. The error bars indicate the standard error.

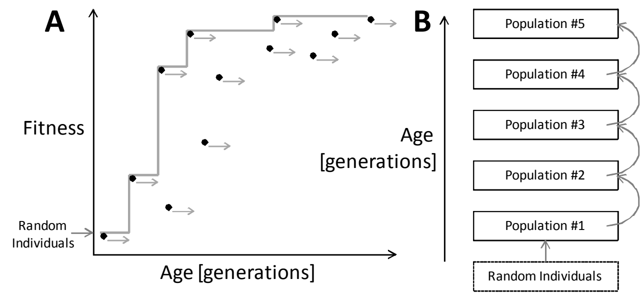

Figure 1. The two optimization methods compared. (A) The Age-Fitness Pareto Population algorithm has a single population of individuals moving in a two-dimensional Age-Fitness Pareto space. (B) The Age-Layered Population Structure (ALPS) algorithm maintains several layers of populations for each age group.

6. ACKNOWLEDGMENTS

This research is supported by the U.S. National Science Foundation (NSF) Graduate Research Fellowship Program, NSF Grant ECCS 0941561 on Cyber-enabled Discovery and Innovation (CDI), and the U.S. Defense Threat Reduction Agency (DTRA) grant HDTRA 1-09-1-0013.

7. REFERENCES

[1] R. Conor, "Reducing Premature Convergence in Evolutionary Algorithms," Ireland: University College, Cork, 1996.

[2] S. J. Louis and G. J. E. Rawlins, "Syntactic Analysis of Convergence in Genetic Algorithms," in *Foundations of Genetic Algorithms 2*: Morgan Kaufmann, 1992, pp. 141--151.

[3] T. Jansen, "On the Analysis of Dynamic Restart Strategies for Evolutionary Algorithms," in *Parallel Problem Solving from Nature — PPSN VII*, 2002, pp. 33-43.

[4] G. S. Hornby, "ALPS: the age-layered population structure for reducing the problem of premature convergence," in *Proceedings of the 8th annual conference on Genetic and evolutionary computation* Seattle, Washington, USA: ACM, 2006.

[5] M. D. Schmidt and H. Lipson, "Age-Fitness Pareto Optimization," in *Genetic Programming Theory and Practice VIII*: Springer, 2010, p. in press.

[6] G. S. Hornby, "A Steady-State Version of the Age-Layered Population Structure EA," in *Genetic Programming Theory & Practice VII*, R. L. Riolo, U.-M. O'Reilly, and T. McConaghy, Eds. Ann Arbor: Springer, 2009, pp. 87-102.

[7] D. Kalyanmoy and K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*: John Wiley \& Sons, Inc., 2001.

[8] S. W. Mahfoud, "Niching methods for genetic algorithms," University of Illinois at Urbana-Champaign, 1995.

[9] J. R. Koza, *Genetic programming: on the programming of computers by means of natural selection*: MIT Press, 1992.

[10] M. D. Schmidt and H. Lipson, "Coevolution of Fitness Predictors," *IEEE Transactions on Evolutionary Computation,* vol. 12, pp. 736-749, Dec 2008.