Reconnoiter by Candle: Identifying Assumptions in Genetic Programming

Jason M. Daida
The University of Michigan, Artificial Intelligence Laboratory and Space Physics Research Laboratory
2455 Hayward Avenue, Ann Arbor, Michigan 48109-2143
daida@eecs.umich.edu
(734) 647-4581 wk. (734) 764-5137 FAX
http://www.sprl.umich.edu/acers

Abstract

This paper takes the position that GP theory could stand to take a temporary detour with the following presumption: that genetic programming and genetic algorithm phenomena are not even remotely related to each other even if they may appear to be similar.

1 INTRODUCTION

Aside from using a tree representation, genetic programming (GP) shares much in common with genetic algorithms (GA). Both are population-based algorithms that uses some form of fitness function. In terms of operators, both have crossover and mutation operators. Both presumably make trade-offs between exploration of solutions in a search space and exploitation of existing subsolutions.

Given these significant similarities, one would find it logical to presume that theories that have been formulated for GA have a significant bearing on explaining phenomena concerning GP. In a sense, GA theory should be able to serve as illumination for understanding processes in GP, if only because GP's fundamental operations are so similar to those employed in GA.

It would seem prudent to transfer theory and associated empirical methodologies from the GA research tradition to GP. The GA research tradition has over two decades of theoretical work; theoretical work in GP has much less than that. There should be analogous GP theories concerning schema and building blocks. There should be analogies for royal road and one-max test problems. It should seem possible to jump-start GP theory by using extant GA theory.

The leap in applying GA theory to GP would be justified if GP exhibited similar phenomena to that observed when using GA. However, that has not yet been established. Similar processes and operations do not and have not necessarily resulted in similar behaviors.

2 POSITION

On one hand, one could take the position that the conditions that have resulted in a type of behavior observed when using a GA were not appropriately transferred over to formulate the analogous conditions for GP. On the other hand, I have taken the following position: temporarily presume that the phenomena in GP have *no* analogies in GA. What are the consequences of this position in the formulation of theory?

This position comes with the following provisos. I make distinctions between belief, position, and theory. Belief is a conviction, a faith in things unseen. Belief provides one an intuition on goals and possible paths in which to make those goals. A position is a way-point along a path that presumably brings one to a goal. One can deliberately chose a *wrong* position or even a position that is antithetical to one's belief, if that position helps one to bracket closer to a goal. A theory represents a falsifiable explanation. (For this position paper, I use the terms *theory* and *hypothesis* interchangeably. In the traditional sense, however, a theory represents a verified and established explanation; a hypothesis represents a conjecture.)

3 RESPONSE

By taking this position, I and my group have revisited and reevaluated assumptions and practices that are common to the GA research tradition. We have started from scratch, since it is no longer a given that even commonplace notions are established. For example, we have presumed that we know little about the kinds of phenomena that can occur in GP. Consequently, we have needed to methodically explore, examine, and classify the kinds of phenomena that are possible for even simple problems. Likewise, we have discarded typical GA explanations of what makes a problem hard (e.g., deception, multimodal landscapes) and have started over by taking a uniquely GP approach. We have raised the possibility that what has made for satisfactory reporting practice in the GA community is not necessarily beneficial for the GP community. We have derived and examined a GP test problem that has no analogy in GA, but are tractable for independent mathematical treatment. We have questioned whether the historical philosophical framing of biological metaphors for GA is well suited for GP.

In claiming (temporary) ignorance of GA, we have forced ourselves to articulate and account for assumptions in seemingly obvious areas. In every case that we have examined this year, we have found that what may seem obvious to the GA

community is not a given for the GP community. It has become increasingly apparent that GP phenomena exhibits significant differences from that observed for GA. It has also become increasingly apparent that the task of formulating theory could benefit in having clear descriptions and taxonomies of GP phenomena.

4 ABOUT THE RESEARCH GROUP

My research group, The University of Michigan Adaptive Computation in Environmental and Remote Sensing Sciences (U-M ACERS), has been investigating theoretical aspects of genetic programming (GP) for several years. U-M ACERS consists of 14 to 20 researchers, consisting primarily of students and a few industrial affiliates.

References

Daida, J.M., D.S. Ampy, M. Ratanasavetavadhana, H. Li, and O. Chaudhri, "Challenges with Verification, Repeatability, and Meaningful Comparison in Genetic Programming: Gibson's Magic," in W. Banzhaf, J.M. Daida, A.E. Eiben, M.H. Garzon, V. Honavar, M. Jakiela, and R.E. Smith, *GECCO-99: Proceedings of the Genetic and Evolutionary Computation Conference, July 13-17, 1999, Orlando, Florida USA*, San Francisco: Morgan Kaufmann, 1999. In Press. ftp://ftp.eecs.umich.edu/people/daida/papers/GECCO99challenges.pdf

Daida, J.M., S.J. Ross, J.J. McClain, D.S. Ampy, and M.R. Holczer, "Challenges with Verification, Repeatability, and Meaningful Comparisons in Genetic Programming," in *Genetic Programming 1997: Proceedings of the Second Annual Conference, July 13–16, 1997, Stanford University*. J.R. Koza, D. Marco, D. Fogel, M. Garzon, H. Iba, and R. Riolo (eds.), San Francisco: Morgan Kaufmann Publishers, 1997. pp. 64–69. ftp://ftp.eecs.umich.edu/people/daida/papers/GP97challenges.pdf

Ross, S.J., J.M. Daida, C.M. Doan, T.F. Bersano-Begey, and J.J. McClain, "Variations in Evolution of Subsumption Architectures Using Genetic Programming: The Wall Following Robot Revisited," in *Genetic Programming 1996: Proceedings of the First Annual Conference*. J.R. Koza, D.E. Goldberg, D.B. Fogel, and R.L. Riolo (eds.), Cambridge: The MIT Press, 1996. pp. 191–99. ftp://ftp.eecs.umich.edu/people/daida/papers/GP96_robot/

Daida, J. M., R. R. Bertram, J. A. Polito 2, and S. A. Stanhope. "Analysis of Single-Node (Building) Blocks in Genetic Programming." In L. Spector, W. B. Langdon, U.-M. O'Reilly, and P. J. Angeline (Eds.), *Advances in Genetic Programming 3*. Cambridge: The MIT Press. In Press.

Daida, J.M., J.A. Polito 2, S.A. Stanhope, R.R. Bertram, J.C. Khoo, S.A. Chaudhary, "What Makes a Problem GP-Hard? Analysis of a Tunably Difficult Problem in Genetic Programming," in W. Banzhaf, J.M. Daida, A.E. Eiben, M.H. Garzon, V. Honavar, M. Jakiela, and R.E. Smith, *GECCO-99: Proceedings of the Genetic and Evolutionary Computation Conference, July 13-17, 1999, Orlando, Florida USA*, San Francisco: Morgan Kaufmann, 1999. In Press. ftp://ftp.eecs.umich.edu/people/daida/papers/GECCO99landscapes.pdf

Daida, J. M., S. P. Yalcin, P. M. Litvak, G. A. Eickhoff, and J. A. Polito 2. "Of Metaphors and Darwinism: Deconstructing Genetic Programming's Chimera." *Proceedings of the Congress on Evolutionary Computation*. In Press. ftp://ftp.eecs.umich.edu/people/daida/papers/CEC99metaphors.pdf