Improving Estimation of Distribution Genetic Programming with Novelty Initialization

Christian Olmscheid
Johannes Gutenberg University
Mainz, Germany
christian.olmscheid@gmx.de

David Wittenberg
Johannes Gutenberg University
Mainz, Germany
wittenberg@uni-mainz.de

Dominik Sobania
Johannes Gutenberg University
Mainz, Germany
dsobania@uni-mainz.de

Franz Rothlauf
Johannes Gutenberg University
Mainz, Germany
rothlauf@uni-mainz.de

ABSTRACT

Estimation of distribution genetic programming (EDA-GP) replaces the standard variation operations of genetic programming (GP) by learning and sampling from a probabilistic model. Unfortunately, many EDA-GP approaches suffer from a rapidly decreasing population diversity which often leads to premature convergence. However, novelty search, an approach that searches for novel solutions to cover sparse areas of the search space, can be used for generating diverse initial populations. In this work, we propose novelty initialization and test this new method on a generalization of the royal tree problem and compare its performance to ramped half-and-half (RHH) using a recent EDA-GP approach. We find that novelty initialization provides a higher diversity than RHH and the EDA-GP also achieves a better average fitness using novelty initialization.

CCS CONCEPTS

• **Theory of computation** → **Design and analysis of algorithms**;
• **Computing methodologies** → **Genetic programming**;

KEYWORDS

Estimation of distribution genetic programming, Novelty initialization

ACM Reference Format:
Christian Olmscheid, David Wittenberg, Dominik Sobania, and Franz Rothlauf. 2021. Improving Estimation of Distribution Genetic Programming with Novelty Initialization. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459410

1 INTRODUCTION

Estimation of distribution genetic programming (EDA-GP) replaces the standard mutation and recombination operators of genetic programming (GP) by sampling from a learned probabilistic model. In each generation, EDA-GP first captures relevant properties of the parent population in a probabilistic model and then uses the probabilistic model to transfer these properties to the offspring. The idea is to reduce the number of fitness evaluations and to increase the overall quality of the search.

Unfortunately, many EDA-GP approaches suffer from a rapidly decreasing population diversity which often leads to pre-mature convergence. This is true since sampling from a probabilistic model usually results in a strong exploitation of the search space. Especially in low locality problem domains (e.g, needle-in-a-haystack problems [5]), EDA-GP easily gets stuck in local optima. Therefore, diversity preserving mechanisms are needed that assure high diversity during an EDA-GP run.

A diverse initial population is the first step towards a high diversity search where ramped half-and-half (RHH) is a popular initialization method for GP. Here, we can use large population sizes to guarantee low sampling error and hopefully high diversity in the initial population [4]. However, RHH is also known for introducing a strong bias and often lacks in generating populations with high diversity [1]. Novelty search [2, 3] is an approach that searches for novel solutions that cover sparse areas of the search space. Therefore, novelty search seems suitable for finding diverse initial populations.

In this work, we propose novelty initialization. We test the new initialization technique on a generalization of the royal tree problem and compare its performance to RHH using denoising autoencoder genetic programming (DAE-GP) [6] as probabilistic model.

2 EXPERIMENTS AND RESULTS

In the following, we briefly introduce the experimental setting and present our results.

2.1 Experimental Setting

To evaluate novelty initialization and compare its performance to RHH, we use a generalization of the royal tree problem where we previously define a target individual x_{opt} and calculate the fitness of each candidate solution (fitness_x) as the normalized Levenshtein distance to x_{opt}. Thus, $\text{fitness}_x \in [0, 1]$, where $\text{fitness}_x = 0$ means that x is identical to x_{opt} (minimization problem) [6].

We search for 8 different target individuals initialized with RHH, minimum depth $d_{\min} = 2$, maximum depth $d_{\max} = 3$, and conduct 10 runs per target individual. As a function and terminal set, we

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07…$15.00
https://doi.org/10.1145/3449726.3459410

use F = {+, −, exp, log} and T = {x}. We set the population size to 750, the tournament selection size to 2, and use the same hyper-parameters as well as the same frameworks for the DAE-GP as in [6].

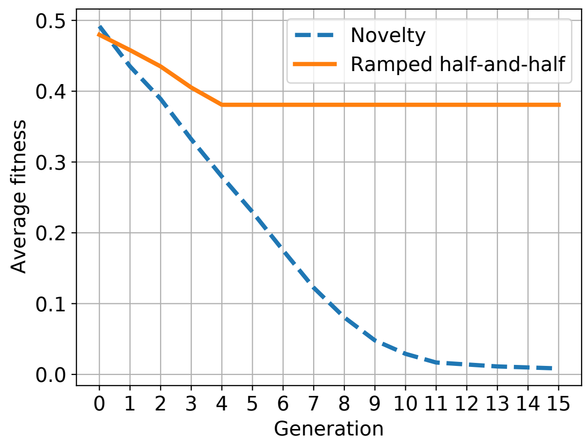

Figure 1: Average fitness over generations for novelty initialization and ramped half-and-half.

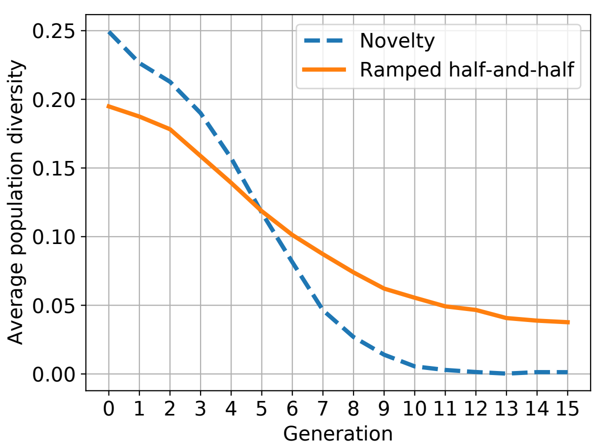

Figure 2: Population diversity over generations for novelty initialization and ramped half-and-half.

To generate a population with novelty initialization, we perform a GP run using novelty search. We define the novelty of a candidate solution by its average normalized Levenshtein distance to the k-nearest neighbors and set k to 10% of the population size, thus k = 75. At the end of the novelty search run, we select the individuals for our initial population from the archive used by novelty search. To increase diversity even for RHH, we remove eventually generated duplicate individuals.

2.2 Results

To analyze the influence of novelty initialization, we compare the development of the average fitness and the development of the average population diversity for aggregated DAE-GP runs. We define the population diversity as the average novelty over all candidate solutions in a population. Results are averaged over 80 runs.

Figure 1 shows the average fitness over generations for novelty initialization and RHH. For both initialization methods we observe a similar average fitness in the initial population. However, while the DAE-GP initialized with RHH gets stuck in generation four, fitness can be further improved with novelty initialization. The DAE-GP achieves a best average fitness of around 0.39 with RHH and a near zero best average fitness using novelty initialization. The results strongly indicate that novelty initialization provides an initial population that better covers the search space.

Figure 2 shows the average population diversity over generations for novelty initialization and RHH. In the initial population, novelty initialization achieves a higher population diversity compared to RHH. As expected, the population diversity decreases during a DAE-GP run. However, the decrease is stronger when using novelty initialization. From generation 5 onwards, population diversity even falls below the population diversity of the DAE-GP initialized with RHH, which can be explained by the strong decrease in the fitness (lower values are better) of the DAE-GP initialized with novelty initialization (Figure 1). Here, the DAE-GP converges strongly towards near optimal solutions resulting in a less diverse population.

3 CONCLUSIONS AND FUTURE WORK

We introduced novelty initialization as a new initialization method for GP. We showed that novelty initialization generates a more diverse initial population and better covers the search space compared to RHH. For a generalization of the royal tree problem, we demonstrated that a DAE-GP with novelty initialization outperforms DAE-GP initialized with RHH.

In future work, we will study the method's performance on additional benchmark problems, other EDA-GPs and further analyze the provided diversity.

REFERENCES

[1] Edmund Burke, Steven Gustafson, and Graham Kendall. 2003. Ramped half-n-half initialisation bias in GP. In *Genetic and Evolutionary Computation - GECCO 2003*. Springer, 1800–1801. https://doi.org/10.1007/3-540-45110-2_71

[2] Joel Lehman and Kenneth O Stanley. 2008. Exploiting open-endedness to solve problems through the search for novelty. In *Proceedings of the Eleventh International Conference on Artificial Life (ALIFE)*. 329–336.

[3] Joel Lehman and Kenneth O Stanley. 2010. Efficiently evolving programs through the search for novelty. In *Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation*. 837–844. https://doi.org/10.1145/1830483.1830638

[4] Dirk Schweim, David Wittenberg, and Franz Rothlauf. 2021. On sampling error in genetic programming. *Natural Computing* (2021). https://doi.org/10.1007/s11047-020-09828-w

[5] Dominik Sobania and Franz Rothlauf. 2020. Challenges of Program Synthesis with Grammatical Evolution. In *European Conference on Genetic Programming (Part of EvoStar)*. Springer, 211–227. https://doi.org/10.1007/978-3-030-44094-7_14

[6] David Wittenberg, Franz Rothlauf, and Dirk Schweim. 2020. DAE-GP: denoising autoencoder LSTM networks as probabilistic models in estimation of distribution genetic programming. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*. 1037–1045. https://doi.org/10.1145/3377930.3390180