The Journal of Open Source Software

pyGOURGS - global optimization of n-ary tree representable problems using uniform random global search

Sohrab Towfighi[1]

1 Faculty of Medicine, University of Toronto

DOI: 10.21105/joss.02074

Software

- Review
- Repository
- Archive

Editor: Ariel Rokem

Reviewers:

- @iljah
- @nicoguaro

Submitted: 02 January 2020
Published: 19 March 2020

Summary

Global optimization problems are ubiquitous to engineering and the sciences. Many such problems are not amenable to analytical techniques and an examination of some potential solutions for these problems often suggests that hill-climbing algorithms would be unable to navigate the jagged and confusing terrain. Despite this, genetic programming is often applied to these problems in the hopes that it will be able to identify high-quality solutions. We suspect that genetic programming would perform no better than random search, in agreement with the No Free Lunch Theorems (Wolpert, Macready, & others, 1997), and we devised this software to allow us to perform uniform random global search, also known as pure random search, on these problems. The challenge lies in creating a system that enumerates all the possible solutions, such that we are then able to randomly select from this space of solutions, giving each solution the same probability of being selected.

We use a dense enumeration of full binary trees (Tychonievich, 2013) and modify it to allow for enumeration of n-ary trees. The enumeration algorithm we use is flexible, modifying its enumeration depending on the arity of the functions that the user supplies and the number of variables that the user supplies. Uniform random global search is proven to converge on the ideal solution as the number of iterations tends to infinity (Solis & Wets, 1981), and this is intuitive because with infinite repetitions the algorithm reduces to an exhaustive search. The software comes with three ready examples derived from the popular DEAP software (Fortin, Rainville, Gardner, Parizeau, & Gagné, 2012). These include the artificial ant problem, the even parity problem, and the multiplexer problem. The software is the successor to our earlier work (Towfighi, 2019), but uses a different enumeration algorithm that is much more generalizable whereas our previous algorithm was only suitable for symbolic regression problems.

In the seminal work of (Langdon & Poli, 1998), they enumerated the solution space using brute force and were able to determine that different types of random search can require different amounts of computational effort to reach a high-quality solution. They found that the random search method commonly used to generate the initial population of genetic programming solutions performs much worse than does uniform random search. We found one highly cited paper that claimed that genetic programming outperformed random search (Sipper, Fu, Ahuja, & Moore, 2018), when in fact they were comparing genetic programming to a biased type of random search which they then put on further unequal footing. This software has broad applicability in the examination of the solution space for global optimization problems and in the analysis of benchmark problems, as it permits brute force computations in addition to random search. This software will be of use to researchers looking to compare the performance of their algorithms with that of pure random search on a wide variety of global optimization problems.

Acknowledgements

Luther Tychonievich was the original author of the algorithm which mapped non-negative integers to unique binary trees.

References

Fortin, F.-A., Rainville, F.-M. D., Gardner, M.-A., Parizeau, M., & Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research*, *13*(Jul), 2171–2175.

Langdon, W. B., & Poli, R. (1998). Why ants are hard. In J. R. Koza, W. Banzhaf, K. Chellapilla, K. Deb, M. Dorigo, D. B. Fogel, M. H. Garzon, et al. (Eds.), *Genetic programming 1998: Proceedings of the third annual conference* (pp. 193–201). University of Wisconsin, Madison, Wisconsin, USA: Morgan Kaufmann. Retrieved from http://www.cs.ucl.ac.uk/staff/W.Langdon/ftp/papers/WBL.antspace/_gp98.pdf

Sipper, M., Fu, W., Ahuja, K., & Moore, J. H. (2018). Investigating the parameter space of evolutionary algorithms. *BioData Min.*, *11*(1), 2. doi:10.1186/s13040-018-0164-x

Solis, F. J., & Wets, R. J.-B. (1981). Minimization by random search techniques. *Mathematics of Operations Research*, *6*(1), 19–30. doi:10.1287/moor.6.1.19

Towfighi, S. (2019). PySRURGS-a python package for symbolic regression by uniform random global search. *The Journal of Open Source Software*, *4*. doi:10.21105/joss.01675

Tychonievich, L. (2013). Enumerating Trees. Retrieved from https://www.cs.virginia.edu/~lat7h/blog/posts/434.html

Wolpert, D. H., Macready, W. G., & others. (1997). No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, *1*(1), 67–82. doi:10.1109/4235.585893