



Commentary on “Jaws 30”, by W. B. Langdon

Alberto Bartoli¹ · Luca Manzoni² · Eric Medvet¹

Accepted: 13 October 2023 / Published online: 22 November 2023
© The Author(s) 2023

Abstract

While genetic programming has had a huge impact on the research community, it is fair to say that its impact on industry and practitioners has been much smaller. In this commentary we elaborate on this claim and suggest some broad research goals aimed at greatly increasing such impact.

The essay by Langdon provides an insightful summary of thirty years of research since the publication of the first book on genetic programming (GP) [1] and suggests some avenues for the next thirty years. We would like to offer a complementary perspective on these issues. There is no doubt that GP has had an enormous impact on the research community, even outside the research circles focused on evolutionary computation [2]. However, it is fair to say that the impact of GP on practitioners and industry has been much smaller. From this point of view, other frameworks aimed at addressing broadly similar issues to those relevant to GP have instead had a huge impact over a comparable period of time.

For example, Kaggle¹ is a mainstream online platform that hosts data science competitions where machine learning practitioners can share cloud-based

¹ <https://kaggle.com>.

This comment refers to the article available at <https://doi.org/10.1007/s10710-023-09467-x>.

Special Issue: Thirtieth Anniversary of Genetic Programming: On the Programming of Computers by Means of Natural Selection.

✉ Alberto Bartoli
bartoli.alberto@units.it

Luca Manzoni
lmanzoni@units.it

Eric Medvet
emedvet@units.it

¹ Department of Engineering and Architecture, University of Trieste, Via Valerio 6/1, 34127 Trieste, Italy

² Department of Mathematics and Geosciences, University of Trieste, Via Weiss 2, 34128 Trieste, Italy

notebooks, enabling reproducible and collaborative analysis. As of June 6-th 2023, Kaggle hosts 169 notebooks and 1 competition tagged “genetic programming”; almost 27 000 notebooks and 154 competitions tagged “neural network”; almost 30 000 notebooks and 1700 competitions tagged “deep learning” (more than 1700 notebooks and 22 competitions in the last 90 days). Kaggle users can also participate in discussions centered on a specific topic: those with the previously mentioned tags are, respectively, 95, more than 4000, almost 8000. As another example, Stack Overflow² is the leading online forum where developers engage in question and answer discussions. Of a similar order of magnitude is the difference in the number of discussions on the above topics: almost 350, almost 20 000, almost 28 000 for tags “genetic-programming”, “neural-network” and “deep-learning”, respectively. The related data on software tools are also quite relevant: more than 21 000 and 81 000 questions tagged “pytorch” and “tensorflow” (software packages specialized in deep learning and machine learning³); 10 and 172 questions tagged “gplearn” and “deap”, respectively (software packages for GP and distributed evolutionary computation). Major cloud platforms offer services for running machine learning workloads in the form addressed by “pytorch” and “tensorflow”,⁴ but we are not aware of any similar offering focused on GP.

Clearly, the value of a research area cannot be judged solely on the basis of its impact on practitioners and industry. Furthermore, it could be argued that thirty years may be too short a time frame to fully realize the potential of a new research area. In fact, the current rise in popularity of neural approaches builds on decades of previous research. However, we believe that the community should try to understand the reasons why GP has had such a limited impact on practitioners. Why there is no pytorch or tensorflow equivalent for GP, i.e., no library that can be applied to *many different application domains* with the expectation of achieving reasonable performance even without painful optimizations? Why there is no cloud-based service platform offering a GP framework?

More generally, several frameworks have emerged in recent years as a kind of go-to solution in specific application domains, e.g., convolutional neural networks (CNN) for image processing, transformers for natural language processing (NLP), XGBoost or decision trees for classifying tabular data. Why is there no application domain where GP plays such a role?

GP could perhaps become the go-to solution for *symbolic regression* problems: it delivers very good performance even with respect to other state-of-the-art methods [3]; and, it can be used easily without any application-specific modelling and implementation effort. On the other hand, symbolic regression is hardly a hot topic for practitioners—the Stack Overflow forum does not even have a tag for this term, nor

² <https://stackoverflow.com/>.

³ Their modules for exploiting GPU hardware can also be used by GP frameworks, but such uses are definitely not very common.

⁴ For example, Microsoft Azure and Google Cloud.

for widely used software packages in this area.⁵ Furthermore, the potential benefits of GP in this area may not be sufficient to motivate the use of a framework that is different from other frameworks that are already mainstream.

We believe this sentence in the essay by Langdon is crucial: “The aim of the book was: 1) a single technique could solve many diverse problems”. Indeed, a single term (GP) is used as an umbrella to denote a broad and diverse set of different frameworks. On the other hand, almost all practical applications of GP involve a significant modeling and implementation effort to find a *solution representation* suitable for the specific flavor of GP being used and for the specific application being considered. Other successful frameworks in the broad field of machine learning can be applied to a wide variety of different application domains much more easily than GP: all that is needed is to organize the training data in a tabular form and then define some constraints on the candidate solutions used by the framework—e.g., the size and number of layers in a neural network.

We, as a community, have maybe failed to make it more explicit that the many GP frameworks are indeed radically diverse from each other, much as CNN and transformers are radically different from each other despite both being neural networks. Perhaps most importantly, we have failed to provide a practically relevant mapping between these frameworks and application domains, because we do not have any tenet similar to “use CNN for image processing (or use transformers for natural language processing): you will most likely obtain very good results even if you cannot afford any fine-tuning or changes in the framework internals”. As a result, practitioners have no incentive to use GP as a black box, unlike what has happened with neural networks in machine learning. And, in a perverse feedback loop, there is no incentive to develop any GP equivalent of general-purpose libraries as pytorch.

Similarly, we have failed to promote GP-based frameworks that allow *reuse of solutions*. Whenever GP is applied to a particular problem, the only approach is to start modeling *and searching* from scratch. We have neither frameworks, nor software tools, nor public repositories that allow the construction of initial populations based on the results of previous searches on similar problems. Furthermore, solutions obtained with GP are rarely, if ever, applied unchanged in settings other than the one in which such solutions were found. Again, these facts represent a significant difference from practices that have become common in other fields, where one can download a neural network trained for image classification or text generation—many such models are publicly available, e.g., in Hugging Face.⁶ This clearly adds additional friction for practitioners interested in using GP in the industry.

While we fully agree with Langdon that “GP is doing well in its mission to help the world”, we also believe that as a community we could be even more ambitious in terms of impact on practitioners. Perhaps our next set of “impossible” goals should include some sort of *speciation* of GP resulting in frameworks tailored to specific application domains of practical interest; whose use requires very little modeling effort; that can be used as a black box while still providing robust and satisfactory results.

⁵ For example, <https://github.com/heal-research/operon> and <https://github.com/MilesCranmer/PySR>.

⁶ <https://huggingface.co/models>.

Funding Open access funding provided by Università degli Studi di Trieste within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. J.R. Koza, *Genetic programming II: automatic discovery of reusable programs* (MIT press, Cambridge, 1994)
2. A. De Lorenzo, A. Bartoli, M. Castelli, E. Medvet, B. Xue, Genetic programming in the twenty-first century: a bibliometric and content-based analysis from both sides of the fence. *Genet. Program Evolvable Mach.* **21**, 181–204 (2020)
3. Cava, W. G. L., Orzechowski, P., Burlacu, B., de Franc, a, F. O., Virgolin, M., Jin, Y., Kommenda, M., and Moore, J. H. (2021). Contemporary symbolic regression methods and their relative performance. CoRR, abs/2107.14351.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.