Mauro Castelli

SHORT CV

Mauro Castelli is an *associate professor and a senior research scientist in the field of Computational Intelligence and Machine Learning in the NOVA Information Management School (NOVA IMS), Universidade NOVA de Lisboa (Portugal)*. He received his Ph.D. in Computer Science at the University of Milano Bicocca (Italy). His research interests are focused on the study of machine learning methods that can be used to take advantage of the vast amount of data that are produced nowadays. In particular, the research focuses on the development, implementation, and application of computational intelligence systems for addressing complex real-world problems in different domains.

Dr. Castelli *has a proven record of outstanding research, as evidenced by existing high-quality publications.* He has published his research in a variety of top-quality academic outlets, such as Expert Systems With Applications, IEEE Transactions on Cybernetics, IEEE Transactions on Evolutionary Computation, Applied Soft Computing, Swarm and Evolutionary Computation, among others. Mauro has published more than 150 contributions in the area of machine learning in international conferences and journals. In particular, his main research interests are in the area of Evolutionary Computation with a focus on Genetic Programming.

His *pedagogical work* includes *lectures and organization of activities at all levels of programs at the home institution and abroad*. He was responsible for lecturing courses on Machine Learning, Deep Learning, Computational Intelligence, and many others. He was the principal investigator in three research projects, where he proposed the use of evolutionary computation techniques for addressing complex real-world problems. Moreover, he is responsible for lecturing a PhD course on Evolutionary Computation. This course received the highest level of satisfaction among the existing PhD courses for two consecutive years. Mauro has supervised more than 80 master students and 5 PhD students. Last but not least, Dr. Castelli is the director of the bachelor program in Data Science

Dr. Castelli has proved *the ability to transfer theoretical insights into practical implications*, as demonstrated by *his participation in projects with ministers and public administrations* including the Directorate General for Education and the Directorate General for Economic Activities among the others.

Last but not least, Dr. Castelli is an active member of the Evolutionary Computation community. He is constantly involved in the organization of conferences/workshops, and he integrated the program committee of the main conferences on Evolutionary Computation, including Evo*, GECCO, and PPSN. Finally, he is a member of the editorial board of the journal Genetic Programming and Evolvable Machines.

Research Group

The research will be carried out at NOVA Information Management School (NOVA IMS). The research group in which the visiting Ph.D. student will be integrated is composed of different senior and junior researchers working in the area of Machine Learning. In particular, the group consists of, among the others, Leonardo Vanneschi (full professor), Mauro Castelli (Associate Professor), Illya Bakurov (a Ph.D. student working in the area of Genetic Programming), Fernando Peres (a Ph.D. student working on the definition of nature-inspired metaheuristics), and Victor Costa (a Ph.D. student working at the integration of evolutionary computation within deep learning). Moreover, we share space and resources with other colleagues working in the area of Machine Learning.

The research group members already worked together on different projects and research, and they published many papers in the area of evolutionary computation in recent years. The research group has also received the best paper awards in 2013 and 2014 for the best paper presented at EuroGP. Additionally, the research group beneficiate from the existing collaboration that exists with the faculty of science of the University of Lisbon. In particular, we are constantly working with Sara Silva, one of the leading experts in the area of genetic programming in Portugal.

The visiting Ph.D. student will find a vibrant environment, friendly professors, and other Ph.D. students with whom sharing and developing new ideas and concepts. Last but not least, the research group has all the necessary state-of-the-art hardware for pursuing an advanced investigation.

Finally, the research group is constantly looking for funding opportunities at the national and international levels. The participation of the visiting student in the preparation of a research grant can provide funding for continuing the research that the visiting student will start at NOVA IMS.

The two leaders of the research group have significant experience in the supervision of Ph.D. students, and they present the following bibliographic indicators:

Mauro Castelli: https://scholar.google.com/citations?hl=en&user=23HUHAIAAAAJ

Leonardo Vanneschi: https://scholar.google.com/citations?hl=en&user=uR5K07QAAAAJ

Research Project

This research project aims to improve the current combination of gradient descent/neural network learning algorithms with Geometric Semantic Genetic Programming (GSGP) presented in EuroGP 2022 [9] by *increasing the number of parameters of GSGP that can be learned with techniques coming from the area of artificial neural networks* (ANNs).

Background

While Genetic Programming (GP) was first developed by John Koza in the nineties [6], in 2012 Moraglio and coworkers introduced a new variant of GP called Geometric Semantic GP [8], or GSGP for short. The main idea was that the operations of crossover and mutation could be re-defined to have a very specific meaning in terms of the effect on the output of an individual. Since the operators work on the output of the individuals they are semantic operators and, since the effects on the output can be defined geometrically, they are geometrical operators. We showed how the operators could be implemented efficiently by moving from a tree-based representation to a directed acyclic graph representation [14, 15]. The resulting algorithm proved to be both faster than standard GP by more than an order of magnitude and able to produce better solutions with a lower risk of overfitting. The work on GSGP continued to this day, with improvements focused on refining the solutions by employing local search methods, like in [3, 2], or more explainable models [7].

Concerning neuroevolution [11], one of the classical methods to evolve ANNs is NEAT (NeuroEvolution of Augmenting Topologies), which iteratively "grows" a network starting from a small initial structure [12].

Unfortunately, the NEAT approach cannot directly scale to large and deep networks, thus leading to the definition of indirect methods based on the same idea, like HyperNEAT [4], where the object of the evolution is not directly the networks but a "generator" for the actual network.

Another approach is to evolve only the structure and not the weights of the network by using a language that is defined to describe only "standard" deep ANN structures, like the approach recently taken by DENSER [1]. More recently, researchers from Google Brain have shown AutoML-zero, in which GP is used to evolve at the same time the structure and the learning algorithm of ANNs [10], finding via evolution the same algorithms that are already in use in the field (e.g., multiplicative interaction, ReLU). However, the resulting network is shallow and the system requires extensive computational resources, limiting its scalability.

In all cases of Neuroevolution, the resulting model is an ANN where the evolutionary part is only in the training. The reverse approach, using ANN optimization techniques during the training of an evolutionary model is instead the topic of this proposal. In the recent Evostar conference [9], we showed how to successfully combine GSGP with learning techniques from the area of ANN.

Proposal

To describe how the combination of GSGP and gradient descent in a way similar to ANN it is necessary to describe the main structure of a GSGP individual, both in its tree form and in its acyclic graph representation, at least in general terms.

Let T_1, T_2 represent GP individuals that can be interpreted as functions $\mathbb{R}^n \to \mathbb{R}$. Suppose all considered GP individuals to be derivable. This can be enforced by selecting only derivable functions in the functional set. Then the semantic crossover of T_1 and T_2 is the function $T_1 \times r + (1 - r) \times T_2$, where r is a randomly generated number in $[0,1]$ (another possibility is to use a randomly generated GP individual and non-linearity can – and should – also be added). Similarly, the semantic mutation of T_1 (or T_2) is the function $T_1 + m \times R$, where m is a small constant called the mutation step, and R a randomly generated GP individual with expected output 0. If all the original functions are derivable, then the results of both crossover and mutations are derivable. Looking at those operations like operations on GP trees actually obscure the fact that, during the evolutionary process, the individuals used in semantic crossover are actually taken by the limited pool given by the previous GSGP generation and that their output can be reconstructed as a weighted combination of the outputs of their parents.

Currently [9], only $r, 1 - r$, and m are the parameters that are modified by gradient descent. However, there are multiple parameters that can be modified in a GP tree. In particular, if we employ, for example, the standard arithmetic operations $+, -, \times, \div$, we can actually see them as having two parameters each, so that,

for example, $+$ applies to two subtrees T_1 and T_2 is not $T_1 + T_2$ but $\alpha T_1 + \beta T_2$ where α, β are initially 1 but are parameters that can be learned.

In addition to that, we remark that geometric crossover can be expanded to more than two parents in an obvious way [13], thus allowing even additional similarities with ANN structures. Thus, the proposed research aims at fostering a stronger combination of evolutionary search with gradient-based methods, intending to combine the advantages of the two approaches in the different stages of the evolutionary process.

References

[1] F. Assunçao, N. Lourenço, P. Machado, and B. Ribeiro. Denser: deep evolutionary network structured representation. Genetic Programming and Evolvable Machines, 20(1):5–35, 2019.

[2] M. Castelli, L. Manzoni, L. Mariot, and M. Saletta. Extending local search in geometric semantic genetic programming. In EPIA Conference on Artificial Intelligence, pages 775–787. Springer, 2019.

[3] M. Castelli, L. Trujillo, L. Vanneschi, S. Silva, et al. Geometric semantic genetic programming with local search. In Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, pages 999–1006. ACM, 2015.

[4] J. Gauci and K. O. Stanley. Autonomous evolution of topographic regularities in artificial neural networks. Neural computation, 22(7):1860–1898, 2010.

[5] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In International Conference on Learning Representations, pages 1–13, 2015.

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

[7] J. F. Martins, L. O. V. Oliveira, L. F. Miranda, F. Casadei, and G. L. Pappa. Solving the exponential growth of symbolic regression trees in geometric semantic genetic programming. In Proceedings of the Genetic and Evolutionary Computation Conference, pages 1151–1158. ACM, 2018.

[8] A. Moraglio, K. Krawiec, and C. G. Johnson. Geometric semantic genetic programming. In International Conference on Parallel Problem Solving from Nature, pages 21–31. Springer, 2012.

[9] G. Pietropolli, L. Manzoni, A. Paoletti, M. Castelli. Combining Geometric Semantic GP with Gradient-Descent Optimization. In European Conference on Genetic Programming (Part of EvoStar) (pp. 19-33). Springer, 2022.

[10] E. Real, C. Liang, D. R. So, and Q. V. Le. Automl-zero: Evolving machine learning algorithms from scratch. arXiv preprint arXiv:2003.03384, 2020.

[11] K. O. Stanley, J. Clune, J. Lehman, and R. Miikkulainen. Designing neural networks through neuroevolution. Nature Machine Intelligence, 1(1):24–35, 2019.

[12] K. O. Stanley and R. Miikkulainen. Evolving neural networks through augmenting topologies. Evolutionary computation, 10(2):99–127, 2002.

[13] L. Vanneschi, M. Castelli, L. Manzoni, K. Krawiec, A. Moraglio, S. Silva, and I. Gonçalves. Psxo: population-wide semantic crossover. In Proceedings of the Genetic and Evolutionary Computation Conference Companion, pages 257–258. ACM, 2017.

[14] L. Vanneschi, M. Castelli, L. Manzoni, and S. Silva. A new implementation of geometric semantic gp and its application to problems in pharmacokinetics. In European Conference on Genetic Programming, pages 205–216. Springer, 2013.

[15] L. Vanneschi, S. Silva, M. Castelli, and L. Manzoni. Geometric semantic genetic programming for real life applications. In Genetic programming theory and practice XI, pages 191–209. Springer, 2014.

[16] M. D. Zeiler. Adadelta: an adaptive learning rate method. arXiv preprint arXiv:1212.5701, 2012.

Other Information

The research will be carried out at NOVA Information Management School (NOVA IMS).

NOVA IMS is an academic unit of the Universidade Nova de Lisboa created in 1989 as a response to the shortage of experts in the areas of data science, data analytics, and information management and also to the growing demand for the use of information systems and technologies. NOVA IMS has currently about 1.500 students enrolled in 10 awarding degree programs (2 bachelor programs, 7 master programs, and 1 doctoral program) plus a large number of post-graduation programs. The degree programs of NOVA IMS are accredited by A3ES, the national agency for the evaluation and accreditation of higher education in Portugal.

NOVA IMS has managed to achieve significant levels of internationalization, not only regarding education, with a major demand from foreign students (in 2020 NOVA IMS had 25% foreign students coming from more than 70 different nationalities) and with the participation of internationally renowned teachers in its teaching staff, but also in research and development activities that are largely supported by international partnerships.

The following services will be available for the candidates that want to spend three (or more) months at NOVA IMS:

• Mediating housing facilities through personalized assistance, which includes contact with the Universities campus students' residences

• Possibility of attending a language course

• Coaching the students on administrative issues, e.g., software licenses, opening bank account, public transportation facilities, etc.

• Providing dedicated contact persons for any type of problems at NOVA IMS.

In addition, there are very active students' representations, which provide personal support for problems with studying and everyday life.

Last but not least, visiting students are invited to attend social and cultural events organized by the university's International Office, where they can mix up with other local and international students. This will be an excellent opportunity for discovering Lisbon and its surroundings.