

GP+EM 20th Anniversary Editorial

Nicholas Freitag McPhee and
William B. Langdon

29 January 2020

1 Introduction

Genetic Programming and Evolvable Machines (GPEM) published its inaugural volume in 2000, with four issues covering a range of topics, from applications in electronics to theoretical analyses. GPEM's birth was one of several important events in a time of change and new beginnings in the field. The first GECCO was only a year earlier (GPEM's issue 3 contained a short summary [12]), and the first EvoStar event [13,14], had been just the year before that. The last few years of the 1990's saw a dozen or so PhD dissertations in genetic programming and related fields, considerably more than in all of the previous 2/3rds of the decade.¹

The articles in the first volume have the feel of a new field finding its way; researchers trying out new ideas and engaging in self-discovery. This was consequential work, however, with those first dozen articles collectively receiving over 1 500 citations in the following two decades.² In the subsequent 20 years the field has grown and the ideas have spread (see, e.g., Figure 1).

The work in this 20th anniversary issue, on the other hand, feels less like a precocious teenager and more broadly developed, rich in both its breadth and depth. There are thoughtful reflections on our history, and mature considerations of the future. The application domains covered here are wonderfully varied, a realization of the promise in those early issues, using techniques and ideas not yet available when the first copies of GPEM arrived in offices, labs,

N. F. McPhee
Division of Science and Mathematics, University of Minnesota Morris
E-mail: mcphee@morris.umn.edu

W. B. Langdon
Department of Computer Science, University College London
E-mail: W.Langdon@cs.ucl.ac.uk

¹ See [9] for more detail.

² Data from Google Scholar, January 2020

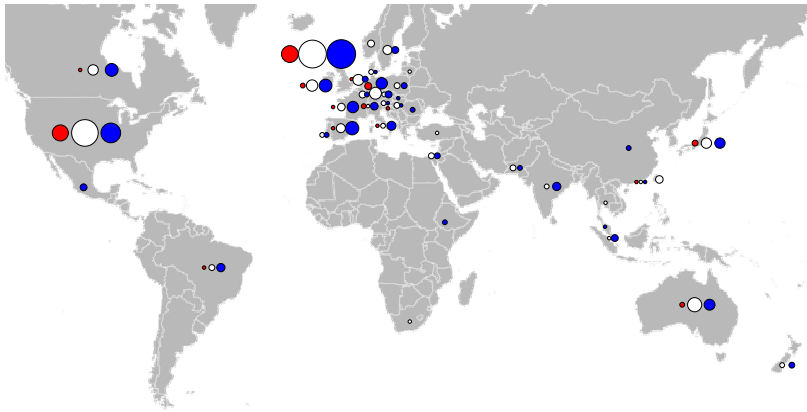


Fig. 1 Distribution of 586 PhD theses over time, as listed in the Genetic Programming Bibliography. The size of circles indicates the number of theses; red indicates the number before 2000, white the number between 2001 and 2009, and blue the number from 2010 onward. Note the substantial increase after 2000 in numerous countries. The bias towards English speaking countries is in part likely a bias of the data in the GP Bibliography, which focuses on material published in English.

and libraries. Perhaps one of the strongest signs of the field’s health is the number of people using things like genetic programming as “standard” tools to aid them in getting their work done, as illustrated in several of articles in this special issue.

This special issue is the result of an open call for review articles, articles highlighting major challenges, and cutting edge ideas. All of the submitted articles went through the journal’s standard, rigorous peer review process. The submissions reflected the breadth of the field, with a particular emphasis on genetic programming. Given this range of topics, we were particularly pleased with the high quality of the reviewing the articles received, and we hope you will agree with us that this is reflected in the high quality of the included articles.

Through this review process we accepted 11 articles ([1]–[11]) which we’ve loosely organized as (a) five articles reviewing uses of genetic programming in various application domains, (b) one review of a specific technique, (c) three surveys of publication venues and trends, and (d) two challenge articles. This is an obviously imperfect categorization because articles of this caliber play multiple roles; many, for example, propose challenges for the future. Hopefully, however, it will serve as a useful first approximation.

2 Review of application domains

This special issue contains five surveys that highlight the wide array of active application domains where GP is finding use. These range from “traditional”

AI applications like natural language processing [1], through finance [2], music [3], and education [4], to applications in heavy industry [5].

- In “Genetic Programming for Natural Language Processing” (NLP) [1] Lourdes Araujo reviews NLP applications of GP, with an emphasis on natural language generation and understanding, as well as information extraction. The review covers several specific applications such as evolving spam filters, mining Twitter content, and medical decision support.
- In “Applications of Genetic Programming to Finance and Economics: Past, Present, and Future” [2] Tony Brabazon et al. consider a consistently popular GP domain, but argue that, when considered objectively from outside the GP community, “there has been surprisingly little work published using GP in the mainstream finance academic community or indeed publicly disclosed by industry practitioners” and focuses on what needs to be done to resolve this.
- “Evolutionary music: Applying evolutionary computation to the art of creating music” by Roisin Loughran and Michael O’Neill [3] notes that computer music arguably started with Ada Lovelace in 1842, long before the invention of the electronic computer. Their review covers not just GP and other EC techniques but includes NEAT and artificial neural networks, including CTRNNs. The breadth of their review is illustrated by an impressive table (Table 1) which highlights the wide range of objective functions used to assess the quality of evolved music. Some of these use interactive techniques, where human users play a role in assessing the quality, but there are also a broad array of approaches that avoid or reduce the bottleneck of human interaction. They conclude by reviewing how we stand against Jon McCormack’s five open problems in evolutionary music and art from 2005 [15], and considering the complex question of how we define and understand creativity.
- Nelishia Pillay considers “The impact of Genetic Programming in Education” [4], with an emphasis on intelligent tutoring systems, predicting student performance, and designing learning environments. She suggests GP libraries will be needed for future automated teachers, allowing intelligent tutoring to be tailored for individual students.
- Our last application review deals literally with a white hot topic and solid financial saving from applying GP. In “Genetic programming in the steel-making industry” [5], Miha Kovacic and Uros Zuperl first review applications of GP to the steel industry world wide before focusing on its use in their company, Store Steel, Figure 2.

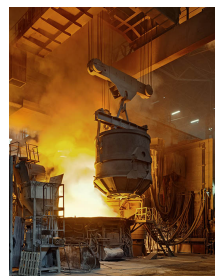


Fig. 2 Store Steel©

3 Review of techniques: Cartesian Genetic Programming

Cartesian Genetic Programming (CGP) was first published under that name in the very first GECCO conference in 1999 by Julian Miller [16]. In the following 20+ years it has been taken up by many people and has become one of the most widely used dialects of GP. In “Cartesian Genetic Programming: Its status and future” [6] Miller provides an easy to follow description of CGP, along with its many variants and applications. He also lists CGP tools in a variety languages from C to Python to Matlab, and finishes with several open questions for future research.

4 Reviews of publication venues and trends

A definite strength of this 20th anniversary issue is a collection of three articles providing a high-level review of publications in the field, exploring features such as content, venues, downloads, and citations, and looking at how these have varied across time and space.

- The annual Genetic Programming Theory and Practice (GPTP) workshop was started in 2003 on the premise that “Theory must inform practice and practice must test theory”. [17] It was coordinated for many years by Rick Riolo at the University of Michigan’s Center for the Study of Complex Systems and has recently moved an hour down the road to BEACON Center for the Study of Evolution in Action at Michigan State University. In “Genetic Programming Theory and Practice: A Fifteen-Year Trajectory”, Moshe Sipper and Jason Moore [7] highlight trends and discuss the trajectory of GP research as revealed by the first 15 GPTP workshops, and make some valuable observations about the development of the field, and the ongoing challenges of connecting theory to a complex and fluid area of active research.
- In “Genetic Programming in the 21st century: A bibliometric and content-based analysis from both sides of the fence” [8] Andrea De Lorenzo, Alberto Bartoli, Mauro Castelli, Eric Medvet, and Bing Xue use unsupervised automated analysis of the GP literature from 2000 to 2018. They compare the results on publications appearing in traditional evolutionary computational venue (such as GPEM) with the large number in non-EC venues (i.e., “both sides of the fence”). One observation, which we take as a sign of health in the field, is that as the absolute number of GP publications has gone up, the proportion in non-EC venues has risen, reflecting broadening interest in and applications of these techniques. They also observe a growing focus on optimization and control, which is reflected in several of the reviews of application domains.
- While De Lorenzo et al. use Scopus and automated analysis, in “Genetic programming and evolvable machines at 20” [9], W. B. Langdon addresses many of the same questions using the hand curated Genetic Programming

Bibliography³. While the GP Bibliography lacks information such as citations, it provides other types of information such as both quantities and sources of downloads. Also, where De Lorenzo et al. based their work on all publications returned by Google Scholar given the search term “genetic programming”, the GP Bibliography is maintained and managed by Langdon, so the two data sets no doubt differ, although their shared observations are predominantly consistent. An observation in this review that is also likely a sign of the health of the field is the substantial decrease in single author papers, and increase in the number of papers with three or more authors.

5 Challenge articles

While many of the preceding articles contain ideas for future work and challenges for the future, these last two articles are more explicitly focused on what they see as major opportunities and challenges for the field.

- “Adversarial genetic programming for cyber security: A rising application domain where GP matters” by Una-May O’Reilly et al. [10] argues that (co)evolutionary ideas and GP are together particularly suited to assist with the substantial cyber security challenges we face, both as individuals and as societies. They present an adversarial GP framework using co-evolutionary populations of (cyber) defenders and attackers as a means of evolving robust and creative solutions, and outline the current problems and future challenges for this kind of work.
- The promise and potential for automatic programming has been part of the GP conversation and culture since the very first publications in the field, and Michael O’Neill and Lee Spector argue in “Automatic Programming: The Open Issue?” [11] that we have largely failed to make substantive progress towards that goal. Little of the applications work reviewed in this special issue could be seen as evolving general purpose programs, and the content analyses in this issue highlight the prevalence of applications such as symbolic regression and optimization, where we’re typically evolving *expressions* instead of *programs*. While they review a number of promising examples in areas such as software synthesis and genetic improvement, they remind us that we remain well short of GP’s promise of automatic programming, and suggest that we’ll likely need multi-component hybrid approaches to reach that goal. They also highlight the complex issues of trust, ethics and regulation that this sort of work inevitably raises.

³ After [9] was copy edited, the GP bibliography was moved to <http://gpbib.cs.ucl.ac.uk> and a mirror site was established at <http://gpbib.pmacs.upenn.edu/>

6 Conclusions

We hope that you find this special issue both reflective and inspiring, helping us better understand how far we've come, and challenging us to push even farther. GP has come a long way from the "toy" problems of its early days to making valuable contributions in a wide array of fields, and exploring a variety of representations and approaches. That said, however, we are still well short of a general purpose automatic programming engine, and there's clearly much more still to be done.

We are very grateful to all the authors that contributed their work and ideas to this special issue, and the many reviewers who did so much to establish and improve its overall quality.

We also greatly appreciate all of you who take the time to explore this issue's contents. Hopefully this will help motivate you to go do the amazing work that will define the next twenty years of GPEM and genetic programming!

References

1. Lourdes Araujo. Genetic programming for natural language processing. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
2. Anthony Brabazon, Michael Kampouridis, and Michael O'Neill. Applications of genetic programming to finance and economics: past, present, future. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
3. Roisin Loughran and Michael O'Neill. Evolutionary music applying evolutionary computation to the art of creating music. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
4. Nelishia Pillay. The impact of genetic programming in education. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
5. Miha Kovacic and Uros Zuperl. Genetic programming in the steelmaking industry. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
6. Julian Francis Miller. Cartesian Genetic Programming: its status and future. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
7. Moshe Sipper and Jason H. Moore. Genetic Programming Theory and Practice: A fifteen-year trajectory. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
8. Andrea De Lorenzo, Alberto Bartoli, Mauro Castelli, Eric Medvet, and Bing Xue. Genetic programming in the twenty-first century: a bibliometric and content-based analysis from both sides of the fence. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
9. W. B. Langdon. Genetic programming and evolvable machines at 20. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
10. Una-May O'Reilly, Jamal Toutouh, Marcos Perterra, Daniel Prado Sanchez, Dennis Garcia, Anthony Erb Luogo, Jonathan Kelly, and Erik Hemberg. Adversarial genetic programming for cyber security: A rising application domain where GP matters. *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
11. Michael O'Neill and Lee Spector. Automatic programming: The open issue? *Genetic Programming and Evolvable Machines*. Twentieth Anniversary Issue.
12. Justinian Rosca. Remembering GECCO-99 – a short summary. *Genetic Programming and Evolvable Machines*, 1(3):297–300, July 2000.
13. Wolfgang Banzhaf, Riccardo Poli, Marc Schoenauer, and Terence C. Fogarty, editors. *Genetic Programming*, volume 1391 of *LNCS*, Paris, 14-15 April 1998. Springer-Verlag.

14. Phil Husbands and Jean-Arcady Meyer, editors. *Proceedings of the First European Workshop on Evolutionary Robotics*, volume 1468 of *LNCS*, Paris, 16-17 April 1998. Springer-Verlag.
15. Jon McCormack. Open problems in evolutionary music and art. In Franz Rothlauf, Juer-gen Branke, Stefano Cagnoni, David W. Corne, Rolf Drechsler, Yaochu Jin, Penousal Machado, Elena Marchiori, Juan Romero, George D. Smith, and Giovanni Squillero, editors, *Applications of Evolutionary Computing, EvoWorkshops2005: EvoBIO, Evo-COMNET, EvoHOT, EvoIASP, EvoMUSART, EvoSTOC*, volume 3449 of *LNCS*, Lau-sanne, Switzerland, 30 March-1 April 2005. Springer Verlag.
16. Julian F. Miller. An empirical study of the efficiency of learning Boolean functions using a cartesian genetic programming approach. In Wolfgang Banzhaf, Jason Daida, Agoston E. Eiben, Max H. Garzon, Vasant Honavar, Mark Jakiela, and Robert E. Smith, editors, *Proceedings of the Genetic and Evolutionary Computation Conference*, volume 2, pages 1135–1142, Orlando, Florida, USA, 13-17 July 1999. Morgan Kaufmann.
17. Rick Riolo and Bill Worzel. Genetic programming theory and practice. In Rick L. Riolo and Bill Worzel, editors, *Genetic programming theory and practice*, volume 6 of *Genetic Programming*, chapter 1, pages 1–10. Kluwer, Boston, MA, USA, 2003. Series Editor - John Koza.