

Low or no cost distributed evolutionary computation

Juan-Julián Merelo-Guervós
University of Granada
<http://ugr.university>
jmerelo@ugr.es

<http://www.sigevo.org/gecco-2015/>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).
GECCO'15 Companion, July 11–15, 2015, Madrid, Spain.
ACM 978-1-4503-3488-4/15/07.
<http://dx.doi.org/10.1145/2739482.2756560>



1

Summary



From the era of big science we are back to the "do it yourself", where you do not have any money to buy clusters or subscribe to grids but still have algorithms that crave many computing nodes and need them to measure scalability. Fortunately, this coincides with the era of big data, cloud computing, and browsers that include JavaScript virtual machines. Those are the reasons why this talk will focus on two different aspects of volunteer or freeriding computing: first, the pragmatic: where to find those resources, which ones can be used, what kind of support you have to give them; and then, the theoretical: how evolutionary algorithms can be adapted to a environment in which nodes come and go, have different computing capabilities and operate in complete asynchrony of each other.

2

What is the point?



The world has computational resources in spades. Most of them do not belong to you or your lab. That does not mean you cannot use it. The problem is how.

Most theory in parallel computing has been devoted to predict and optimize the performance in systems where the number of nodes, their connections, and the time every one is devoting to the computation is known in advance. However, even if Big Science is not over, the era of Citizen science has started (with SETI@home and then BOINC) and it offers a vast amount of computational resources to attract, if only you know how. But there is a challenge: knowing, or at least having a ballpark, of how your algorithm is going to perform in this uncertain environment, where none of the factors is known: neither the number of nodes, through how they are connected, to how long are they going to be focused on doing what you want them to.

3

What is out there?



Besides, since Amazon started selling EC2 several years ago, reliable and scalable computing resources are also available for a low price and on demand. Recently, Google has also refurbished its offering lowering their prices. This means that the conjunction of free or low-cost cloud computing engines, volunteer computing systems and and untapped capability of desktop systems can be used for creating massive, or at least potentially massive, distributed computing experiments.

4

What is the point?



In this talk we will offer our experience on using browser-based computing since 1999 and other emerging paradigms, such as peer to peer based computing, mainly applied to the design of evolutionary algorithms.

There are many issues involved in using these resources: from adapting current algorithms so that they match this environment to check which EA configuration works the best in it, or creating a framework that can use it easily. But the main challenge is that asking people to contribute resources implies that you are opening your science to society and you have to give something in return: you have to adopt a set of best practices that have come to be known as Open Science, because “Give, and it shall be given unto you”, you will get as much back from society as you give to it opening your science and explaining it to the public. This, among other things, means that popularity will become directly performance of the metacomputer you create by attracting more users. And you can do it, for instance, giving talks about distributed evolutionary computation.

5

The real deal?



This talk will be available as interactive slides in GitHub.

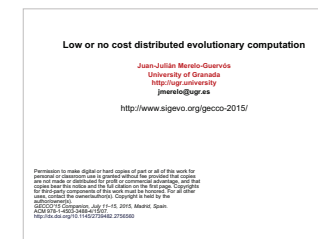
<http://jj.github.io/low-or-no-cost-ppsn2014/gecco15.html>

6

References

- ❖ J. González, Juan-Julián Merelo-Guervós, P. A. Castillo, V. Rivas, G. Romero, and A. Prieto. Optimized web newspaper layout using simulated annealing. In José Mira and Juan Vincente Sánchez-Andrés, editors, *Engineering Applications of Bio-Inspired Artificial Neural Networks, International Work-Conference on Artificial and Natural Neural Networks, IWANN '99, Alicante, Spain, June 2-4, 1999, Proceedings, Volume II*, volume 1607 of *Lecture Notes in Computer Science*, pages 759–768. Springer, 1999.
- ❖ Juan-Luis Jiménez-Laredo, A. E. Eiben, Maarten van Steen, and Juan-Julián Merelo-Guervós. EvAg: a scalable peer-to-peer evolutionary algorithm. *Genetic Programming and Evolvable Machines*, 11(2):227–246, 2010.
- ❖ J. J. Merelo, P.A. Castillo, J.L.J. Laredo, A. Mora, and A. Prieto. Asynchronous distributed genetic algorithms with JavaScript and JSON. In *WCCI 2008 Proceedings*, pages 1372–1379. IEEE Press, 2008.
- ❖ Juan Julián Merelo Guervós. NodeEO, a evolutionary algorithm library in Node. Technical report, GeNeura group, March 2014. Available at <http://figshare.com/articles/nodeeo/972892>.
- ❖ Juan Julián Merelo-Guervós and Pablo García-Sánchez. Modeling browser-based distributed evolutionary computation systems. *ArXiv e-prints*, March 2015. <http://arxiv.org/abs/1503.06424>

7



Summary

From the era of big science we are back to the "you yourself", where you do not have any money to rent or subscribe to grids but still have algorithms that need computing nodes and need them to measure scalability. Fortunately, this coincides with the era of big data, cloud computing, and browsers that include JavaScript virtual machines. Those are the reasons why this talk will focus on two different aspects of volunteer or freeriding computing: first, the pragmatic: where to find those resources, which ones can be used, what kind of support you have to give them; and then, the theoretical: how evolutionary algorithms can be adapted to a environment in which nodes come and go, have different computing capabilities and operate in complete asynchrony of each other.



What is the point?

The world has computational resources in spades, but they do not belong to you or your lab. That does not mean you can use it. The problem is how.

Most theory in parallel computing has been devoted to predict and optimize the performance in systems where the number of nodes, their connections, and the time every one is devoting to the computation is known in advance. However, even if Big Science is not over, the era of Citizen science has started (with SETI@home and then BOINC) and it offers a vast amount of computational resources to attract, if only you know how. But there is a challenge: knowing, or at least having a ballpark, of how your algorithm is going to perform in this uncertain environment, where none of the factors is known: neither the number of nodes, through how they are connected, to how long are they going to be focused on doing what you want them to.



What is out there?

Besides, since Amazon started selling EC2 services a few years ago, reliable and scalable computing resources are available for a low price and on demand. Recently, Google has also refurbished its offering lowering their prices. This means that the conjunction of free or low-cost cloud computing engines, volunteer computing systems and and untapped capability of desktop systems can be used for creating massive, or at least potentially massive, distributed computing experiments.



What is the point?

In this talk we will offer our experience on using browser-based computing since 1999 and other emerging paradigms, such as peer based computing, mainly applied to the design of evolutionary algorithms.

There are many issues involved in using these resources: from adapting current algorithms so that they match this environment to check which EA configuration works the best in it, or creating a framework that can use it easily. But the main challenge is that asking people to contribute resources implies that you are opening your science to society and you have to give something in return: you have to adopt a set of best practices that have come to be known as Open Science, because "Give, and it shall be given unto you": you will get as much back from society as you give to it opening your science and explaining it to the public. This, among other things, means that popularity will become directly performance of the metacomputer you create by attracting more users. And you can do it, for instance, giving talks about distributed evolutionary computation.



The real deal?

This talk will be available as interactive slides in GitHub.

<http://fj.github.io/low-or-no-cost-ppan2014/gecco15.html>

6



References

- ♦ J. González, Juan-Julian Merelo-Guervío, P. A. Castillo, V. Rivas, G. Ramon, and A. Prieto. Optimized web newspaper layout using simulated annealing. In José Mira and Juan Vincente Sánchez-Andrés, editors, *Engineering Applications of Bio-Inspired Artificial Neural Networks, International Work-Conference on Artificial and Natural Neural Networks, IWANN '99, Alicante, Spain, June 2-4, 1999, Proceedings, Volume II*, volume 1607 of *Lecture Notes in Computer Science*, pages 759–768. Springer, 1999.
- ♦ Juan-Luis Jiménez-Laredo, A. E. Eiben, Maarten van Steen, and Juan-Julian Merelo-Guervío. Evkag: a scalable peer-to-peer evolutionary algorithm. *Genetic Programming and Evolvable Machines*, 11(2):227–246, 2010.
- ♦ J. J. Merelo, P. A. Castillo, J. J. Laredo, A. Mora, and A. Prieto. Asynchronous distributed genetic algorithms with JavaScript and PHP. In *IFCT 2008 Proceedings*, pages 1372–1379. IEEE Press, 2008.
- ♦ Juan-Julian Merelo-Guervío. NoFit, a evolutionary algorithm library in Node. Technical report, GeNeura group, March 2014. Available at <https://github.com/articles/fitness/972832>.
- ♦ Juan-Julian Merelo-Guervío and Pablo García-Sánchez. Modeling browser based distributed evolutionary computation systems. *ArXiv e-prints*, March 2015. <http://arxiv.org/abs/1503.06424>

7