



Benchmarking: state-of-the-art and beyond

Anne Auger and Nikolaus Hansen
Inria and CMAP, Ecole Polytechnique, IP Paris

Full set of slides: <http://www.cmap.polytechnique.fr/~nikolaus.hansen/gecco-2021-benchmarking-state.pdf>

Permission to make digital or hard copies of all or part 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 components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

GECCO '21 Companion, July 10–14, 2021, Lille, France

© 2021 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-8351-6/21/07...\$15.00

<https://doi.org/10.1145/3449726.3461424>

ABSTRACT

Benchmarking is a compulsory task to assess the performance of a (new) optimization algorithm. While this appears as a mainly technical task, there are surprisingly many methodological problems that arise when benchmarking algorithms. Over the past decade, there was a great effort towards improving the benchmarking methodology for (gradient-free) optimization. It was started for continuous optimization problems and then extended to multi-objective and mix-integer problems.

In this tutorial, we will present and discuss these key methodological ideas emphasizing the importance of quantitative measurement, the use of instances of problems as well as choosing well the testbed to not bias results towards too easy problems. We will particularly review the advantages of presenting data using the empirical cumulative distribution of runtimes, a tool that everyone assessing the performance of an algorithm should know.

We will then review how this methodology is implemented within the COCO software and show how COCO can and should be used to benchmark an algorithm and write a scientific paper.

This tutorial is intended for young researchers starting in the field who need benchmarking for their research as well as for researchers that wish to get up-to-date with the latest methodological developments in benchmarking methodology.