ERRATUM

Open Access

CrossMark



Daniel S. Himmelstein^{1,2}, Casey S. Greene² and Jason H. Moore^{1*}

* Correspondence: jason.h.moore@dartmouth.edu ¹Department of Genetics, Dartmouth Medical School, One Medical Center Drive, Lebanon, NH 03756, USA Full list of author information is available at the end of the article

Erratum to

After publication of this article [1], it has been noticed that Figs. 1 and 3 (Figs. 1 and 2 respectively here) had been incorrectly reverted in the original article [1].

The correct presentation of Figs. 1 and 3 (Figs. 1 and 2 respectively here) are included in this erratum.



© 2016 Himmelstein et al. **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/ publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.





figure maps the progress of one run of the three-way algorithm across the 2000 generations of the evolution strategy. Instead of a single three-dimensional graph, we decomposed the illustration into three pairwise plots in which each solution dataset drawn appears once on each plot. Each dot represents a dataset from a Pareto front and shows how that dataset scored on the x and y-axis attributes. The axis are drawn so points closer to the bottom-left corners of the plots represent more optimized solutions. The black dots represent the non-dominated solutions from the original random initialization of 1000 datasets. The Pareto fronts from every subsequent two-hundredth generation are drawn and assigned a color based on their generation. The chronological generation progression follows the colors of a rainbow and can be most easily discerned from the bottommost plot. The star indicates the dataset that was chosen from the final Pareto front to represent the run. These datasets are taken from each run, according to the euclidean distance strategy discussed in the Model Free Dataset Generation Method section, and used to calculate the summary statistics in Table 1. This figure provides insight into the difficulty of the problem. Minimizing the one and two-way accuracies cocturs relatively quickly (within the first few hundred generations). Maximizing the higher order accuracies continues throughout the entire run with progress continuing into the two-thousandth generation

Author details

¹Department of Genetics, Dartmouth Medical School, One Medical Center Drive, Lebanon, NH 03756, USA. ²LewisSigler Institute for Integrative Genomics, Princeton University, Carl Icahn Laboratory, Princeton, NJ 08544, USA.

Received: 18 January 2016 Accepted: 18 January 2016 Published: 3 February 2016

References

 Himmelstein DS, Greene CS, Moore JH. Evolving hard problems: generating human genetics datasets with a complex etiology. BioData Mining. 2011;4:21. doi:10.1186/1756-0381-4-21.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at www.biomedcentral.com/submit

