The Journal of Open Source Software

nimCSO: A Nim package for Compositional Space Optimization

Adam M. Krajewski [1¶], **Arindam Debnath** [1], **Wesley F. Reinhart** [1,2], **Allison M. Beese** [1], **and Zi-Kui Liu** [1]

1 Department of Materials Science and Engineering, The Pennsylvania State University, USA 2 Institute for Computational and Data Sciences, The Pennsylvania State University, USA ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Summary

nimCSO is a high-performance, low-level tool for selecting components (dimensions) in compositional spaces which optimize the data availability for applications such as machine learning, which is a combinatorically hard problem for complex compositions existing in highly dimensional spaces due to the interdependency of components being present. Such spaces are often encountered in materials science, where datasets on Compositionally Complex Materials (CCMs) often span 20-40 chemical elements, while each data point contains several of them.

This tool employs a set of methods, ranging from (1) brute-force search through (2) genetic algorithms to (3) a newly designed search method. They use custom data structures and procedures written in Nim language, which are compile-time optimized for the specific problem statement and dataset pair, which allows nimCSO to run faster and use 1-2 orders of magnitude less memory than general-purpose data structures. All configuration is done with a simple human-readable config file, allowing easy modification of the search method and its parameters.

Statement of Need

The Compositionally Complex Materials (CCMs), and their matal-focused subset of High Entropy Alloys (HEAs), belong to a rapidly emerging class of materials, first proposed by (Cantor et al., 2004) and (Yeh et al., 2004). Contrary to more traditional materials, they contain a large number of chemical elements, typically 4-9 in similar proportions, in hope to thermodynamically stabilize the material by increasing its configurational entropy, by up to $\Delta S_{conf} = \Sigma_i^N x_i \ln x_i$ for ideally random mixing of N elements with fractions x_i.

DRAFT

The Journal of Open Source Software

Methods and Performance

Overview

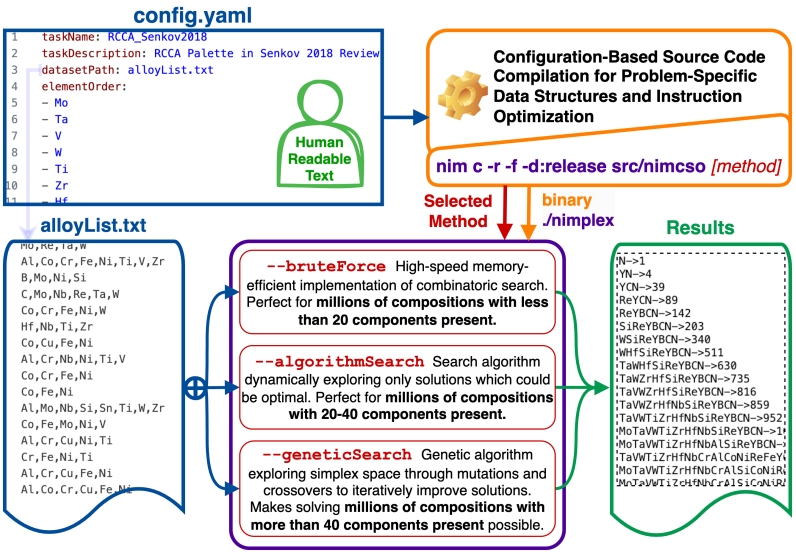

Figure 1: Schematic of core nimCSO data flow with a description of key methods. Metaprogramming is used to recompile the software optimized to the human-readable data and configuration files at hand.

DRAFT

Acknowledgements

This work has been funded through grants: **NSF-POSE FAIN-2229690**, **ONR N00014-23-2721**, and **DOE-ARPA-E DE-AR0001435**.

We would also like to acknowledge Dr. Jonathan Siegel at Texas A&M University for valuable discussions and feedback on the project.

References

Cantor, B., Chang, I. T. H., Knight, P., & Vincent, A. J. B. (2004). Microstructural development in equiatomic multicomponent alloys. *Materials Science and Engineering A*, *375-377*, 213–218. https://doi.org/10.1016/j.msea.2003.10.257

Yeh, J. W., Chen, S. K., Lin, S. J., Gan, J. Y., Chin, T. S., Shun, T. T., Tsau, C. H., & Chang, S. Y. (2004). Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. *Advanced Engineering Materials*, *6*, 299–303. https://doi.org/10.1002/adem.200300567