Machine learning-driven approach to discover novel tungsten multi-principal-element alloys

Samuel Price, Jonah Erlebacher[a], Ian McCue[b]

[a]Department of Materials Science & Engineering, Johns Hopkins University, Baltimore, MD;

[b]Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Since the Bronze Age, humans have discovered and improved upon numerous metal alloys. Traditionally, these alloys consisted of a single primary element with a few secondary elements (generally <15 at.% total). Recently, however, a different approach has arisen, one which seeks to design alloys containing several elements, each present in large fractions. This multi-principal-element (MPE) approach opens the door for alloys with more favorable microstructures and superior properties[1]. Unfortunately, the prized flexibility of the MPE method also presents the considerable challenge of navigating the vast realm of possible alloy combinations.

Previous efforts to synthesize multi-principal-element alloys (MPEAs) have often used an Edisonian approach, which is bound to become mired in failed experiments and local minima. Thus, it is imperative to use computational methods to guide the discovery of MPEAs. Calculation of PHAse Diagrams (CALPHAD) software is a powerful tool, which can predict the equilibrium phases for a given alloy composition at a specified temperature. These phase predictions can be used to find precipitation-hardenable alloys and screen out alloys that contain deleterious intermetallic or topologically close-packed phases. Although this high-throughput screening is much more efficient than blind experimentation, the sheer number of possible alloys makes brute force CALPHAD searching untenable. For example, searching a single five-element alloy system in increments of 1 at.% (on a modern laptop capable of evaluating 3 compositions per second) would take over 2 weeks to complete. Since there are tens of thousands of quinary alloy systems, a more targeted approach is needed.

To overcome this difficulty, I will use machine learning (ML) algorithms to guide the use of CALPHAD calculations and accelerate the discovery of MPEAs with favorable phases. I will demonstrate this approach on tungsten alloys; however, this approach is agnostic to the specific elements involved and can be generalized to any class of MPEAs. To begin, phase stability data will be gathered on over 1 million tungsten ternary alloys. For each ternary alloy, several thermochemical features that correlate well with phase stability[2,3] (e.g., average atomic radius, average valence electron concentration, etc.) will be calculated directly from its composition. These features, along with the phase results, will be used to train ML algorithms, such as k-nearest neighbors and random forest classifiers, in predicting phase stability from alloy composition. These algorithms can effectively "metamodel" CALPHAD since CALPHAD predictions are derived from continuous free energy functions of specific phases. These machine learning "metamodels" can predict an alloy's crystal structure over 1000 times faster than CALPHAD can, making systematic searches of tungsten quaternary and quinary alloys feasible.

References:

[1] 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*, 213-218.

[2] Ward, L., Agrawal, A., Choudhary, A., & Wolverton, C. (2016). A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials*, *2*(1), 1-7.

[3] Islam, N., Huang, W., & Zhuang, H. L. (2018). Machine learning for phase selection in multi-principal element alloys. *Computational Materials Science*, *150*, 230-235.

Characters (no spaces): 2496