JOM, Vol. 72, No. 12, 2020
https://doi.org/10.1007/s11837-020-04388-x

Designing a Periodic Table for Alloy Design: Harnessing Machine Learning to Navigate a Multiscale Information Space

SCOTT R. BRODERICK[1] and KRISHNA RAJAN[1,2]

1.—Department of Materials Design and Innovation, University at Buffalo, Buffalo, NY 14260, USA. 2.—e-mail: krajan3@buffalo.edu

We provide an overview of how to apply statistical learning methods to directly track the role of alloying additions in the multiscale properties of alloys. This leads to a mapping process analogous to the Periodic Table where the resulting visualization scheme exhibits the grouping and proximity of elements based on their impact on the properties of alloys. Unlike the conventional Periodic Table of elements, the distance between neighboring elements in our Alloy Periodic Table uncovers relationships in a complex high-dimensional information space that would not be easily seen otherwise. We embed this machine learning approach with an epistemic uncertainty assessment between data. We provide examples of how this data-driven exploratory platform appears to capture the alloy chemistry of known engineering alloys as well as to provide potential new directions for tuning chemistry for enhanced performance, consistent with accepted mechanistic paradigms governing alloy mechanical properties.

INTRODUCTION

The complexity of the role of elemental chemistry in the properties, processing, and performance of superalloys has been and continues to be an area of intense study. Of course, this is due to the well-known multidimensional/multiscale impact of alloying additions on the phase stability, microstructural evolution, dislocation dynamics, and chemical and thermal resistance, to mention just a few of the metrics that need to be met. At present, there are sophisticated models and experimental techniques that address specific segments of engineering design such as: developing new materials (e.g., first-principles calculations), refining legacy materials (e.g., through processing and microstructural modification), and engineering design and manufacturing. Linking together information across multiple scales requires accounting for the interaction of the myriad parameters that govern materials development and the complexity of the engineering performance.[1–4] Current approaches that utilize a data-driven approach do so in conjunction with physical-based and/or heuristically driven models. The computational design of alloys has largely focused on the issue of phase stability,[5–8] and this in itself is a massive combinatorial problem in selecting which combination of elements need to be added to the base alloy chemistry.

The search for elemental substitutions and/or additions needed to refine metal alloy compositions and enhance their properties is a classical problem in metallurgical alloy design. Finding appropriate alloy chemistries based on a systematic exploration using either computational and/or experimental approaches is often guided by prior heuristic knowledge that harnesses expected trends captured in the Periodic Table that can influence phase stability coupled to optimization schemes for linking chemistry to different properties/mechanisms.[9,10] Despite decades of work, we have, as of yet, no unified mathematical formalism for harnessing this heuristic knowledge and thus more rapidly targeting our next potential discovery of an alloy.

The magnitude of the combinatorial problem in alloy design was succinctly pointed out by Brewer[11] over 50 years ago when he noted that predictions of the phase behavior for multicomponent systems containing the 30 metals of the three transition series (from potassium through nickel, rubidium

(Received July 6, 2020; accepted September 14, 2020; published online October 9, 2020)

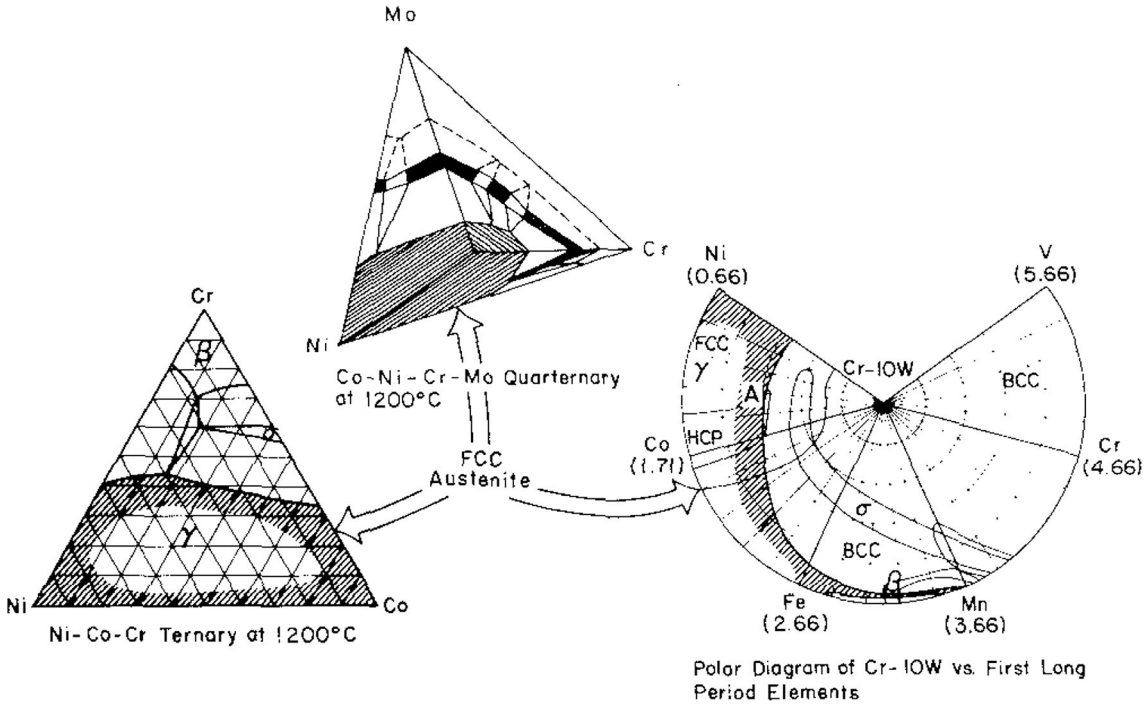

Fig. 1. Mapping the influence of elements on phase chemistry and structure in multicomponent alloys: (left) traditional two-dimensional isothermal projection, (center) three-dimensional isothermal projection, and (right) polar projection mapping the influence of alloying additions based on a string of elements in a period of the Periodic Table on the phase stability of a binary matrix composition. Adapted from Ref. 12.

through palladium, and cesium through platinum) require over two billion diagrams. For example, the number of 14-component diagrams is 582,000,000, there are 142,500 four-component diagrams, and even 435 diagrams are required for only the binary systems. Brewer and others have proposed different ways in which such high-dimensional data could be linked to elemental compositions (see, for example, Fig. 1); however, their complexity of interpretation has limited their application and extensive use.

Clearly, we are still left with the daunting task of exploring the parametric trajectory of systematic changes in chemistry, both computationally and experimentally; i.e., *in what directions will systematic changes in chemistry most likely lead to promising materials with the desired properties?* The challenge and the source of uncertainty here is that knowledge of both chemistry and properties exists for many material systems, but it is not clear how they all connect and influence each other and ultimately define the performance of the alloy. This therefore represents an epistemic uncertainty problem, because in principle the connection is represented in some form in the data, but it is just not known how. It is not obvious from simple inspection of known data or the examination of property trends of elements from the Periodic Table, despite the decades of theoretical and empirical research in the field of alloy optimization and design.

To address this challenge, we describe a data-driven methodology for tracking the collective influence of the multiple attributes of minor alloying elements on both thermodynamic and mechanical properties of engineering alloys. By mapping the high-dimensional nature of the systematics of elemental and alloy data embedded in the Periodic Table into the form of a network graph, one can identify the influence of specific elements on targeted properties. This provides a fundamentally new means to rapidly identify new stable alloy chemistries with enhanced high-temperature properties. The resulting visualization scheme exhibits the grouping and proximity of elements based on their impact on the properties of the alloy in the limited parameter space. We fuse information from multiple sources (legacy data, first-principles calculations, etc.) in an iterative fashion, provide insights into the physics of the materials behavior, and rapidly establish the connections and pathways through the complex information network that links fundamental materials chemistry to materials design.

MAPPING ELEMENTAL CHEMISTRY OF ALLOYS

Manifold Learning

The IsoMap approach,[13] which is briefly described here, gives results which vary depending on two different inputs. The first is the selection of data, and the second is the number of nearest-neighbor connections defined for each node. Beyond the tracking of known information in the final network,

we have developed an automated approach for tracking the sensitivity of the networks as a function of input data. The approach inputs a subset of the data of a given dimensionality and develops the network, defining the nearest neighbors and the geodesic distance contained by each dimension. This is repeated multiple times for the same dimensionality but with different descriptors chosen in each iteration for that given number of descriptors. The dimensionality is then increased, and the process repeated with multiple descriptor combinations for a given dimensionality. This then defines how much the result is dependent on the descriptors selected and provides a measure of uncertainty associated with the data selection.

The selection of data (or "descriptors") was organized into broad classes of information: discrete scalar parameters that relate to solid-state properties of single elements and static thermodynamic and physical properties of potential alloy chemistries using Miedema's model, coupled to alloy design rules from the classical theories on phase stability of Villars, Mooser–Pearson, Pettifor, and Hume-Rothery.[14–16] Beyond these descriptor sets built around elemental characteristics and from the Miedema model, two additional descriptors were developed which account for more microstructural aspects of the alloy: lattice mismatch and phase volume change, which are based on the differences between the matrix of binary Me-Al (where Me refers to the dominant transition metal, and in our examples either Ti or Ni) and for Me-Al-X. The list of descriptors corresponding with Fig. 2 are listed in the Electronic Supplementary Material.

The IsoMap algorithm was used to discover the optimal low-dimensional graph embedding of elements in terms of their role as alloying additions, such that the geodesic distance between the elements in the higher-dimensional manifold is preserved when it is mapped onto the lower-dimensional graph (Fig. 3).

The objective of the IsoMap algorithm is to map the distribution of elements in the high-dimensional space, represented by the set of data points $\{x_i\} \epsilon R^n$, onto a convex nonlinear manifold M^d of lower dimension $d < n$. Then through dimensionality reduction, we obtain a two- or three-dimensional embedding of the elements onto a weighted graph. The mapping is carried out such that the geodesic distances between the elements in the higher-dimensional manifold are preserved when it is mapped onto the lower-dimensional graph, such that the edges of the graph are weighted in their length according to the original geodesic distances. The similarity/dissimilarity between alloying elements with respect to their effect on the alloy as compared with targeted alloy properties is captured by the distances between the vertices along the edges that connect them to their nearest neighbors. The mathematics and interpretation of the IsoMap approach have been detailed in our prior work.[17–20] As a brief summary, the IsoMap procedure involves mapping x_i (points within the larger data manifold) to y_i (calculated points within the lower-dimensional sub-manifold), and sequentially comprises the following:

1. Computation of the Euclidean distance matrix $[E]$ between each pair of points (x_i, x_j) in R^n.
2. Retention of only those entries of $[E]$ which connect every point to its k nearest neighbors to obtain the reduced matrix $[E_k]$.
3. Calculation of the geodesic distances matrix $[G]$ by approximating with the graph distances

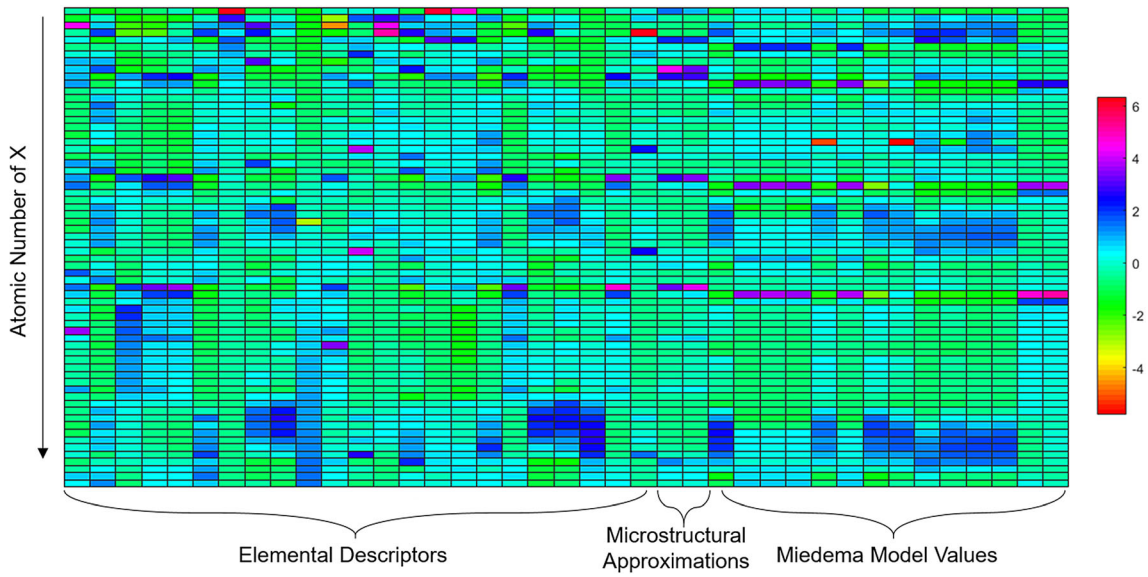

Fig. 2. Heat map representing the descriptor space screened in this paper. This figure contains descriptors derived from various well-established models. The distribution of correlation demonstrates that our analysis will not be overly sensitive to outliers, while containing sufficient physical diversity.

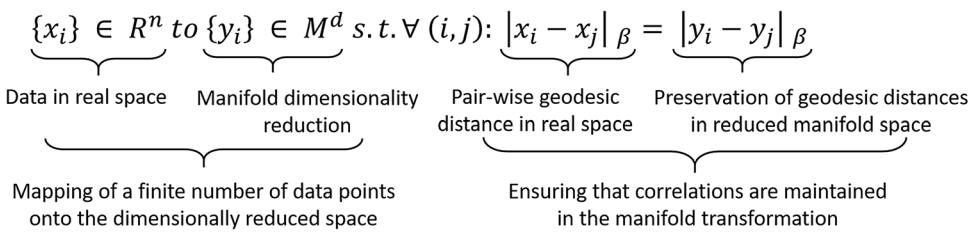

Fig. 3. Description of the nonlinear manifold process used. β is a norm, representative of the pairwise geodesic distances between elements *i* and *j*.

between each pair of points (x_i, x_j) in R^n from $[E_k]$.
4. Computation of the dissimilarity matrix $[A]$.
5. Solving the eigenvalue problem for $[A]$ to obtain the eigenpairs (λ, v). The product of the eigenpairs provide the low-dimensional embedding of the nonlinear manifold.

Given the complexity of information analyzed here, it is necessary to provide a brief definition of some general terms we use throughout this paper. First, we use "descriptor" as a term describing the controllable condition of the chemistry. For example, we can modify the size of an atom on a site by replacing that element with a different element. Thus, that controllable feature is defined by a descriptor related to size. Additionally, regarding the concept of descriptor versus attribute, the descriptor refers to a specific value (for example, ionic radius versus atomic radius), while we use the term "attribute" to refer, for example, to the broad category of size effects. We define "property" as the response resulting from the collective influence of those descriptors (for example, mechanical properties, which we cannot control directly through the chemical modification but rather which results from those changes). Finally, we refer to the dimensionality at two different points in the process. The first is our *input* dimensionality, which refers in this case to the number of descriptors we have input into our analysis (i.e., the descriptors corresponding with Fig. 2). The other use of dimensionality refers to the dimensionality of our *output*, which is significantly lower than our input dimensionality. For all analyses in this paper, the number of output dimensions is four, as this captured greater than 90% of the variance in all cases. For visualization purposes, only the first two or three dimensions are shown, although the connections shown are based on a four-dimensional analysis.

Uncertainty Assessment

Since the manifold in high-dimensional space can vary depending on the number of nearest neighbors chosen, a measure of statistical uncertainty in the geodesic distances can be obtained by varying the number of nearest neighbors to check for short-circuit errors as well as by ensuring the optimum number of dimensions for low-dimensional representation. In this case, we define uncertainty in terms of the sensitivity in the identified connections to the input parameters of neighborhood size and number of dimensions. For this reason, we explored the relationships and pathways for various nearest-neighbor values and with changing dimensions. Specifically, we modify the input parameters of number of nearest neighbors and number of dimensions and information captured by each respective IsoMap. By defining uncertainty in this way, we ensure we do not short-circuit the pathway.

Uncertainty is tracked by exploring permutations and combinations of how attributes to alloy chemistry appear to influence properties. Unlike using a large database perspective of more data, we are exploring what "statistical" proximity exists among groups of elements in a given alloy system. Our measure of proximity is defined by what are the nearest neighbors for a given element. Just as the Periodic Table maps proximity of elements associated with their intrinsic properties as single atoms, we strive to develop a mapping of elements that reflects their similarity/dissimilarity in the context of the alloy chemistry in which they reside.

A key issue is to assess the level of confidence we have on the nearest-neighbor relationships between elements. The uncertainty of the connections identified can be assessed by changing the number of nearest-neighbor connections, as well as the number of dimensions included in the analysis. The change of connections and neighboring lengths are correlated to the uncertainty in the results. To assess the sensitivity of the network to the available data, we performed a series of analyses after removing parts of the data. The logic is that, if we see significant change in the network (represented by the variance captured by the first dimensions) at the final selection of data, then the network may be sensitive to additional data inputs. The choice of descriptor combinations was optimized by statistically determining the smallest value that could minimize the residual variance $|dM - dG|$, while providing the maximum number of alternative paths. For each data point, we also compute the ratio of the distance to its closest and farthest neighbor. These ratios are then averaged over all data points to calculate a scale-invariant, global parameter, Δ[21] to estimate the measure of uncertainty introduced by sparsity in high-dimensional

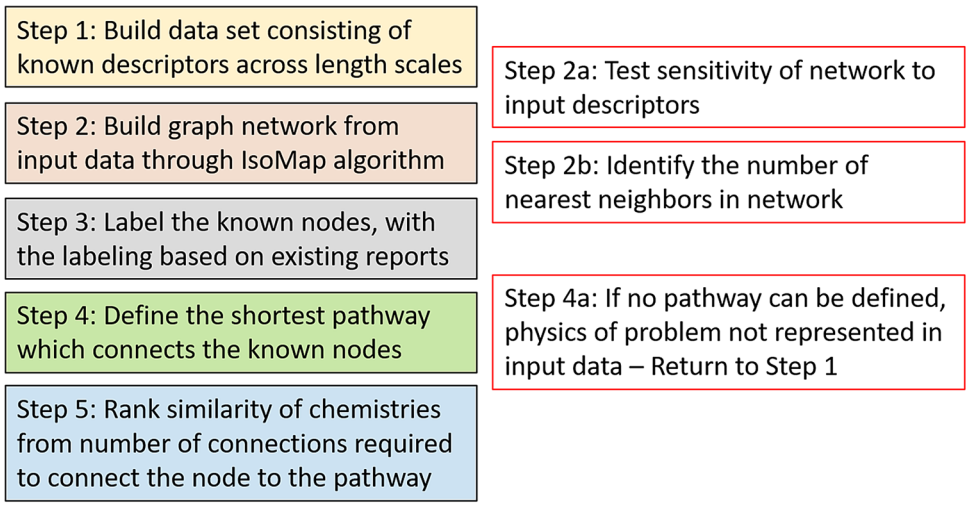

Fig. 4. Process for ranking the impact of elements on target alloy properties. This process was developed in a generalized manner and thus can be applied to multiple types of material properties. Beyond the standard application of graph networks, steps 3 and 4 link the unsupervised network with existing information from literature and handbooks, and therefore the network becomes labeled and target pathways are defined. The distance a node (representing a chemical addition) is from the pathway defines its similarity to those elements and corresponds to the probability of contributing to the property.

spaces, given that the data points must have sufficient density on the manifold.[22] Δ can range between 0 and 1, and a small value indicates a healthy variance in pairwise distances k. Once the graph is constructed, the geodesic distance dG between elements x_i and x_j for $|i - j| > k$ is estimated by calculating the pairwise Euclidean distances along the edges connecting all the intermediate vertices between x_i and x_j.

However, if we capture a near-steady-state behavior in a graph of amount of data included versus variance, we can consider our graph network to be reasonably robust. Therefore, in this case we choose $k = 4$ to ensure that the resulting graph embedding is neither overconnected, leading to loss of pairwise geodesic distances, nor are critical neighbors disconnected. Further, the comparison of connections under the different input parameters does not change significantly, demonstrating that the results presented here have low levels of uncertainty for every node. In this study, from our starting 40 descriptors, we tested nearly 400,000 combinations of descriptors and were able to select the graph that provides the greatest constancy of nearest-neighbor relationships between nodes (elements). That graph is then the most robust and minimizes the uncertainty.

The role of this step in our overall framework is laid out in Fig. 4. In the first step, we develop our descriptor set, as described above. The procedure laid out in this section was then used to assess these descriptors. The inclusion of descriptors beyond what is needed does not result in any issues, as we are tracking the robustness of the network and just need to ensure that additional descriptors do not result in changing results. The nonlinear manifold analysis is then performed. In this paper, the descriptor dimensionality is 40 and the number of nearest neighbor connections is 4. From this, we apply existing knowledge onto the network to provide new interpretations of the information. In each of our case studies, the known information maps onto the respective networks; however, we also discuss how this process can guide the selection of descriptors in cases where the network does not map the known relationships (as highlighted in step 4a). At the end of the analysis, we have an interpretable measure of the similarities and anticipated impact of the different applicable elements.

PERIODIC TABLES FOR HIGH-TEMPERATURE ALLOYS

We have previously introduced the use of graph theory for the design of new multicomponent alloys, where we identified new stable Co-based alloys with the subsequent literature showing the same results through experimental or computational approaches.[17] In this section, we provide additional examples of the utility of the graph network representations for a “Periodic Table for Alloys”. As noted above, the elemental selection for designing new superalloys has been an area of study for over 80 years. Using published literature data on selected commercial alloys over decades, we have found that the evolution of the addition of new alloying elements mapped as first and second nearest-neighbor correlations are mapped onto our network graphs (Figs. 5 and 6).

The Ni-based superalloy Periodic Table not only maps the apparent connectivity of minor alloying additions but also provides clues and suggestions regarding what trajectories in chemical space may serve as substitutions. For instance, Fig. 7 (using another projection of the high-dimensional network)

shows how the rare earth (RE) elements cluster in a very discrete manner (in dashed circle) compared with the rest of the alloy network, just as they do in the conventional Periodic Table. However, unlike the Periodic Table, the "Alloy Periodic Table" identifies the similarity/dissimilarity via nearest-neighbor connections of other elements that may hold promise as RE substitutions. So, using a dissimilarity metric relative to the RE cluster, Fig. 7 suggests via the connectivity of the network where to explore in a semisupervised manner. It would be reasonable for instance to explore refractory elements, leading to the subnetwork identified by the cluster in Fig. 7. Based on our network analysis that ranks the dissimilarity from rare earth elements via nearest-neighbor relationships, this refractory element cluster includes Re, which significantly improves creep life.[23] Different generations of superalloys have involved the use of different levels of Re and Ru combinations to enhance creep life, and they in fact coexist in our Alloy Periodic Table as first nearest neighbors. However, Re and Ru are expensive and rare, but this network suggests that exploring other nearby heavy elements such as W, Mo, and Ta may be a promising choice of elements to tune the composition without resorting to the use of rare earth elements. This in fact is consistent with recent suggestions in the literature, where these elements have been shown to impose a drag effect on dislocation movement, thus reducing the creep strain rate.[24] This example serves to highlight that our Alloy Periodic Table concept serves as a data-driven exploratory tool to identify trajectories of chemical pathways that rapidly identify new directions for alloy design. By coupling this identification with prior knowledge of the theoretical and mechanistic influence of nearby elements, one can be guided by physical principles. With additional or new data, this network can change and evolve, permitting a dynamic process of refinement in alloy design. These interpretations emphasize the amount of information contained within our input descriptor set described earlier. It should be noted, for instance, that although no information specific to creep behavior was included in our inputs, our analysis identifies correlations related to creep. This serves as an example of how the Alloy Periodic Table can serve as an exploratory data-mining tool for tracking potential chemistry–property correlations.

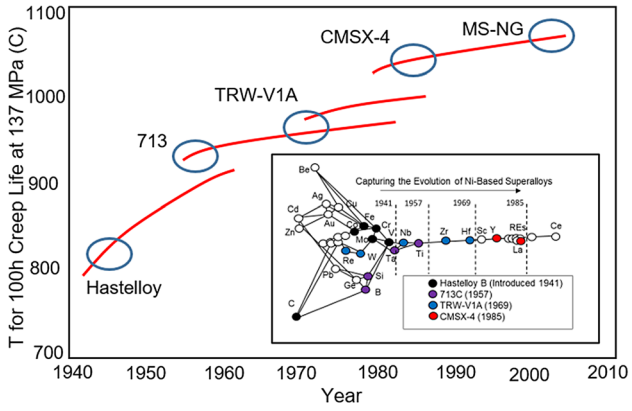

Fig. 5. Tracking the evolution of Ni-based superalloys. Our graph network (inset) captures the availability of commercial alloys corresponding with an increase in operating temperature. An expanded view of the elemental mapping for Nickel-based superalloys is shown in Fig. 6. Adapted from Ref. 4

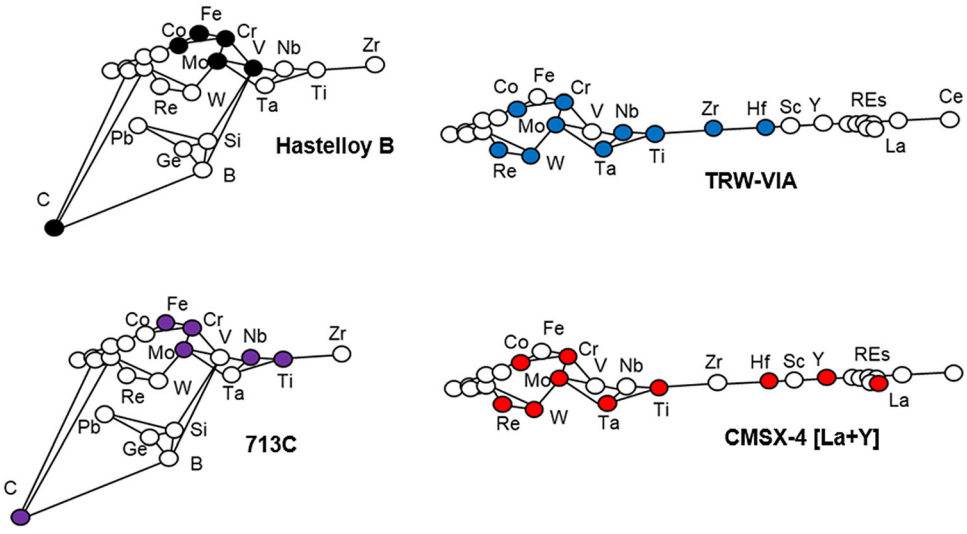

Fig. 6. Compositional mapping of some commercial alloys. Subsections of the graph shown in Fig. 5 are expanded for clarity for each alloy. The elemental nodes representing new alloying additions associated with each alloy are highlighted. In each case, we see how the networks capture, in terms of nearest-neighbor linkages, the chemical "trajectory" that was actually used in the design of these alloys.

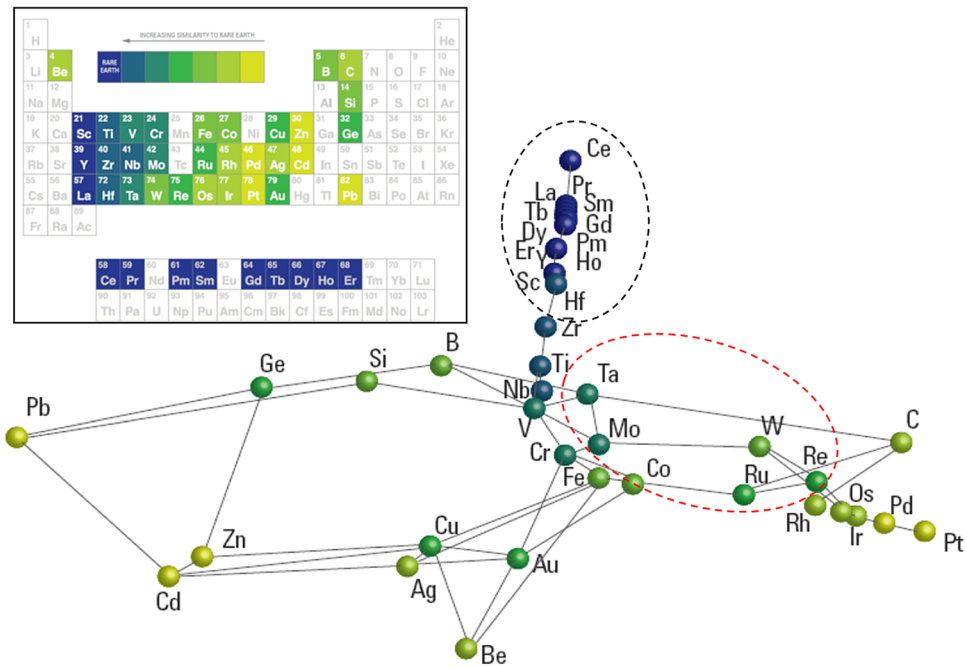

Fig. 7. Mapping of information of the design network onto the traditional Periodic Table. The three-dimensional network is shown here in a different projection from Figs. 5 and 6 to highlight the connections. Based on the number of nearest-neighbor connections between the RE elements and the other elements, a similarity/dissimilarity ranking was developed. The dashed line cluster captures the rare earth elements (keeping in mind that Hf is actually out of the plane of this projection and not in the rare earth cluster). The red outline cluster identifies refractory elements that serve as plausible chemical space for compositional tuning to enhance creep resistance.

We conclude with another example where we have produced a Periodic Table for alloying additions to TiAl. The γ-TiAl phase has the $L1_0$ structure, which causes low room-temperature ductility due to the limited number of slip systems. There have been many attempts at improving the ductility through alloying.[25–29] Typically, the improvements in the room-temperature ductility occur when there are multiple phases present, such as lamellar mixtures of the γ-TiAl tetragonal $L1_0$ phase and the hexagonal α_2 phase.[30–32] An approach to link alloy chemistry with phase formation and microstructure would have a significant impact on tailoring the design of alloys for targeted mechanical properties.

As an initial step in that direction, we developed a graph network for TiAl alloys, and consider the identification of elemental additions that may impact the behavior of single-phase systems (future work will report on the use of the graph network for chemical design of multiphase TiAl alloys). For our case study, we focus on the solid-solution strengthening and precipitation hardening of γ-TiAl through the additions of Nb, Cr, Mn, and V (typically 1–3% composition). The interpretation of the network can be applied to other impacts of chemical additions, such as increased diffusion activation energy and decreased dislocation climb rates for Hf, Mo, Ta, and W, or the precipitate formation and increased nucleation rate promoted by addition of interstitial elements such as B, C, N, Si, and Y. Finally, beta stabilizers such as Mo, Nb, and W have been shown to improve room-temperature ductility through beta solidification instead of peritectic reaction.[33] It should be noted though that there are competing mechanisms and the substitution rules cannot be simply defined, warranting the need for a machine learning analysis. For example, we include Nb in our list of beta stabilizers, but its addition at low Al content reduces ductility due to microsegregation while it improves ductility at high Al content. Thus, for the given set of descriptors, we provide one compression of this information that provides an interpretable framework for developing chemical design guidelines.

The resulting network is shown in Fig. 8a. After constructing the network, we track how the information regarding the role of elements maps out on this network. While we could apply this for any of the cases described, we focus here on the improvement of ductility through solid-solution strengthening contributed by Nb, Cr, Mn, and V, and therefore we are imposing known information[25,26] onto the network. These four elements are denoted by filled circles, and the general region of the network associated with these elements is circled.

Following step 4 from the flowchart described in Fig. 4, we seek to define a pathway connecting these points. The paths are defined as along the connecting edges between nodes. Nb and V are connected, as are Cr and Mn. However, one continuous pathway between all four nodes is not available.

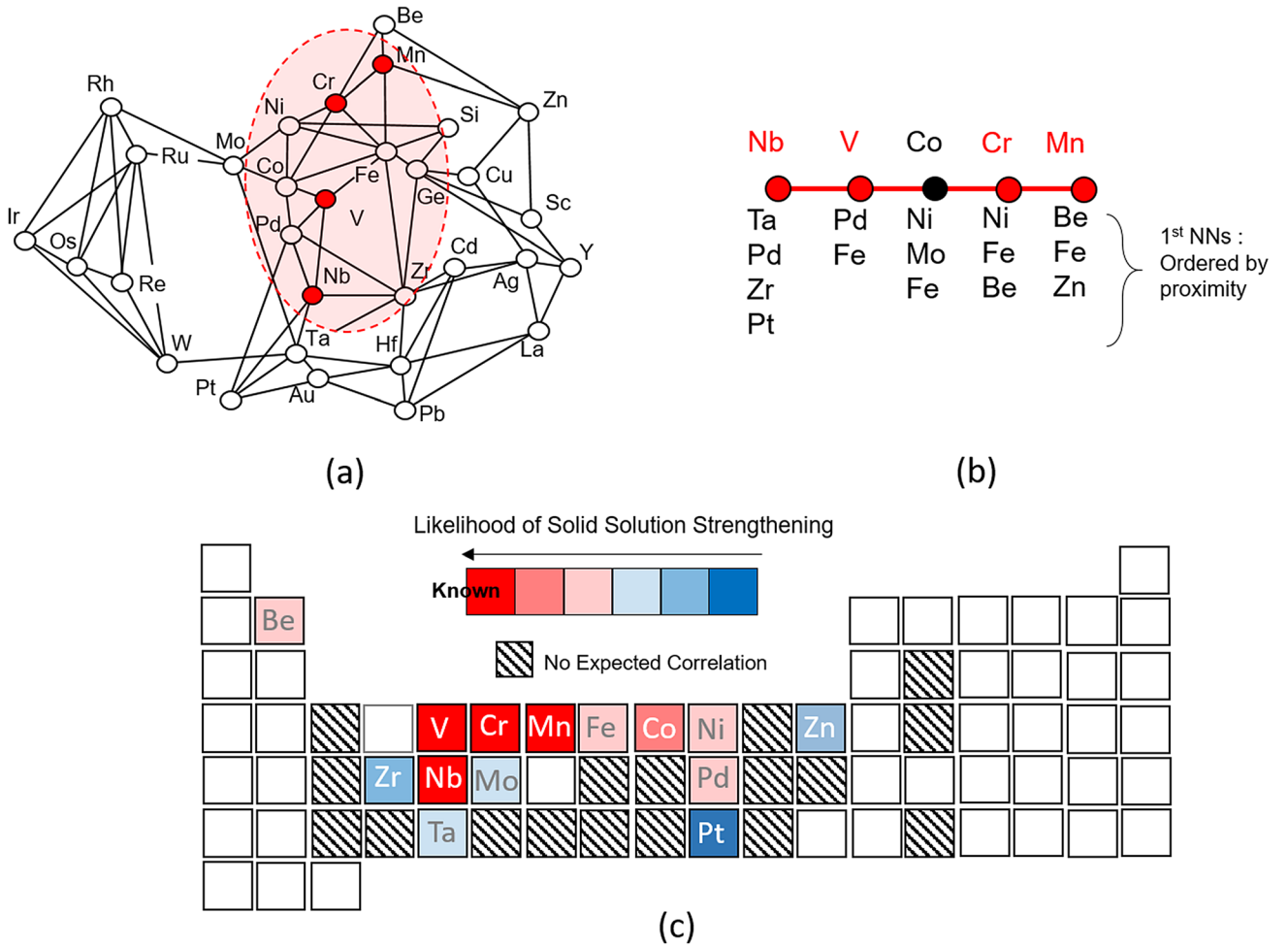

Fig. 8. Ranking of probability of alloying elements promoting solid-solution strengthening of Ti-Al alloys. (a) The network for Ti-Al-X, where each node represents the alloying element. Nb, V, Cr, and Mn have been shown to exhibit solid-solution strengthening. (b) From these previously reported findings, we define a pathway connecting these elements (with the addition of Co), as well as defining the NNs, which have the highest probability of having similar behaviors. (c) Based on these, a ranking of the likelihood of elements to promote solid-solution strengthening is made. Co, Fe, Ni, and Pd are identified as the most promising alloying elements.

However, we do not prevent additional nodes from being present, as we allow for information which has not yet been defined. Therefore, a pathway which includes an additional node does not require moving to step 4a. A single connecting pathway for Nb, V, Cr, and Mn is possible with just one additional node (either Co or Fe). We define the most likely additional node as that which minimizes the length of the pathway, and therefore we define Co as falling in the category with the other four elements.

The first NNs of these elements are those which are most likely to behave similarly to them and are therefore the most likely to contribute to solid-solution strengthening. These elements are listed in Fig. 8b, with the elements further ordered based on the distance between the nodes (e.g., Pd is closer to V than Fe, although both are connected to V). First, we identify Co as having the highest probability for similar behavior given that it falls on the defined pathway. The next highest probability is for those elements which are connected to the pathway multiple times (i.e., Be, Ni, Fe, and Pd). The ensuing probabilities follow the ordering of distances of the remaining NNs. There are few experimental reports on these alloys, thus representing the identification of new potentially promising compositions. In the case of Co, which we rank with the highest probability, studies are largely limited to addition to higher-order systems, although the Co addition still provides a beneficial response.[34] In the case of Ni, Fe, and Pd, there are reports on the beneficial impacts of these elements, although the exact mechanism has not been clearly defined.[35–39]

Considering that those elements which fall under the other suggested mechanism of decreasing dislocation climb (Hf, Mo, Ta, W, and Zr) are all low ranked or unranked, this therefore agrees with the interpretation of our network. In fact, a network with these elements could be constructed for these elements with the addition of a Ru node. Further, these two subnetworks do not intersect, suggesting

that these separate mechanisms are both being captured. In terms of beta stabilizers, Mo, Nb, and W are all connected through Ta, providing an alternate design pathway that has been shown to correspond with improved room-temperature ductility. An important point to highlight is that the network is based on the parameters input and may adapt with addition of further descriptors. The descriptors used here were selected with the objective of providing information across a range of length scales, while also selecting descriptors which are widely available. The interpretation of the network is then based on available domain knowledge. This serves the added benefit of identifying when additional descriptors are needed. In the cases discussed here of solid-solution strengtheners and beta stabilizers, the network captures the known relationships. However, if there is a property of interest for which a pathway cannot be defined, then additional data or other projections of the network may be needed to capture conflicting mechanisms. Our mapping not only provides a convenient representation of the complex alloying process but also captures known mechanisms, and so captures other elements which may promote the same mechanism, while also providing a representation of the data landscape.

CONCLUSION

It is appropriate to conclude by referencing R.W. Cahn's paper in *JOM* nearly half a century ago, entitled "Modern Practice in the Design of Strong Alloys".[40] He described the design of alloys as being driven by a process of discovery that was "*entirely sensual and the mathematics was only necessary to be able to communicate with other people*." He espoused on the role that intuition plays in our integration of theory with phenomenological and heuristic understanding of materials behavior. This intuition is needed to capture the uncertainty in establishing the multiscale linkages between chemistry and alloy performance. In this paper we have provided an example of how statistical/machine learning-based methods provide a mathematical framework to harness that intuition. Hence the "Periodic Table of Alloys" platform derived from statistical learning principles serves as a framework for exploring the choices and selection of minor alloying additions to accelerate chemical design of alloys for enhanced performance.

ACKNOWLEDGEMENTS

We gratefully acknowledge support from the National Science Foundation (NSF) under Grant No. 1640867; and from the Collaboratory for a Regenerative Economy (CoRE) at the University at Buffalo. K.R. acknowledges the Erich Bloch Endowed Chair at the University at Buffalo.

CONFLICT OF INTEREST

On behalf of all authors, the corresponding author states that there are no conflicts of interest.

ELECTRONIC SUPPLEMENTARY MATERIAL

The online version of this article (https://doi.org/10.1007/s11837-020-04388-x) contains supplementary material, which is available to authorized users.

REFERENCES

1. C.T. Sims, *JOM* 18, 1119 (1966).
2. C.T. Sims, *JOM* 21, 27 (1969).
3. M.J. Donachie, S.J. Donachie, A Guide to Engineering Selection of Superalloys for Design, *Mechanical Engineers Handbook* (Hoboken, NJ: Wiley, 2015), p. 299.
4. H. Long, S. Mao, Y. Liu, Z. Zhang, and X. Han, *J. Alloys Compd.* 743, 203 (2018).
5. C.E. Mohn and W. Kob, *Comput. Mater. Sci.* 45, 111 (2009).
6. G.H. Jóhannesson, T. Bligaard, A.V. Ruban, H.L. Skriver, K.W. Jacobsen, and J.K. Nørskov, *Phys. Rev. Lett.* 88, 255506 (2002).
7. C.C. Fischer, K.J. Tibbetts, D. Morgan, and G. Ceder, *Nat. Mater.* 5, 641 (2006).
8. S.V. Dudiy and A. Zunger, *Phys. Rev. Lett.* 97, 046401 (2006).
9. E. Alabort, D. Barba, M.R. Shagiev, M.A. Murzinova, R.M. Galeyev, O.R. Valiakhmetov, A.F. Aletdinov, and R.C. Reed, *Acta Mater.* 178, 275 (2019).
10. R.C. Reed, T. Tao, and N. Warnken, *Acta Mater.* 57, 5898 (2009).
11. L. Brewer, *Prediction of High Temperature Metallic Phase Diagrams* (Ann Arbor: University of Michigan, 1963).
12. C.T. Sims, A History of Superalloy Metallurgy for Superalloy Metallurgists, *Superalloys (Fifth International Symposium)* (Pittsburgh, PA: TMS 1984), p. 399.
13. J. Tenenbaum, V. de Silva, and J. Langford, *Science* 290, 2319 (2000).
14. P. Villars, *J. Less Comm. Met.* 99, 33 (1984).
15. P. Villars, *J. Less Comm. Met.* 92, 215 (1983).
16. A.R. Miedema, *J. Less Comm. Met.* 32, 117 (1973).
17. S. Srinivasan, S.R. Broderick, R. Zhang, A. Mishra, S.B. Sinnott, S.K. Saxena, J.M. LeBeau, and K. Rajan, *Sci. Rep.* 5, 17960 (2015).
18. A. Mullis, S. Broderick, A. Binnebose, N. Peroutka-Bigus, B. Bellaire, K. Rajan, and B. Narasimhan, *Mol. Pharm.* 16, 1917 (2019).
19. X. Zhen, T. Zhang, S. Broderick, and K. Rajan, *Mol. Syst. Des. Eng.* 3, 826 (2018).
20. A. Dasgupta, Y. Gao, S. Broderick, E.B. Pitman, and K. Rajan, *J. Phys. Chem. C* 124, 14158 (2020). https://doi.org/10.1021/acs.jpcc.0c01492.
21. W.J. Cukierski, D.J. Foran, Using Betweenness Centrality to Identify Manifold Shortcuts, *IEEE International Conference on Data Mining Workshops* (New York, NY: IEEE, 2008), p. 949.
22. M. Balasubramanian and E.L. Schwartz, *Science* 295, 7 (2002).
23. S. Sulzer, M. Hasselqvist, H. Murakami, P. Bagot, M. Moody, and R. Reed, *Metall. Mater. Trans. A* 51, 4902 (2020). https://doi.org/10.1007/s11661-020-05845-7.

24. X. Wu, S.K. Makineni, C.H. Liebscher, G. Dehm, J.R. Mianroodi, P. Shanthraj, B. Svendsen, D. Burger, G. Eggeler, D. Raabe, and B. Gault, *Nat. Commun.* 11, 389 (2020).
25. P.V. Cobbinahand and W.R. Matizamhuk, *Adv. Mater. Sci. Technol.* 2019, 4251953 (2019).
26. F. Appel, J.D.H. Paul, and M. Oehring, *Gamma Titanium Aluminide Alloys: Science and Technology* (Hoboken: Wiley, 2011).
27. R.B. Schwarz, P.B. Desch, S. Srinivasan, and P. Nash, *Nanostruct. Mater.* 1, 37 (1992).
28. H. Clemens and S. Mayer, *Adv. Eng. Mater.* 15, 191 (2013).
29. C. Kenel and C. Leinenbach, *Intermetallics* 69, 82 (2016).
30. S.-W. Kim, J.K. Hong, Y.-S. Na, J.-T. Yeom, and S.E. Kim, *Mater. Des.* 54, 814 (2014).
31. M. Schloffer, F. Iqbal, H. Gabrisch, E. Schwaighofer, F.-P. Schimansky, S. Mayer, A. Stark, T. Lippmann, M. Goken, F. Pyczak, and H. Clemens, *Intermetallics* 22, 231 (2012).
32. S.-W. Kim, Y.-S. Na, J.-T. Yeom, S.E. Kim, and Y.S. Choi, *Mater. Sci. Eng. A* 589, 140 (2014).
33. R. Chen, Q. Wang, Y. Yang, J. Guo, Y. Su, H. Ding, and H. Fu, *Intermetallics* 93, 47 (2018).
34. Y. Xia, P. Yu, G.B. Schaffer, and M. Qian, *Mater. Sci. Eng. A* 574, 176 (2013).
35. S. Nishikiori, K. Matsuda, and Y.G. Nakagawa, *Mater. Sci. Eng. A* 239–240, 592 (1997).
36. C. Qiu, Y. Liu, W. Zhang, B. Liu, and X. Liang, *Intermetallics* 27, 46 (2012).
37. L.A. MwambaI, L.A. Cornish, E. van der Lingen, and J.S. Afr, *Inst. Min. Met.* 7, 517 (2012).
38. C. Herzig, T. Przeorski, M. Friesel, F. Hisker, and S. Divinski, *Intermetallics* 9, 461 (2001).
39. J. Zhang, X. Su, E. Strom, Z. Zhong, and C. Li, *Mater. Sci. Eng. A* 329–331, 499 (2002).
40. R.W. Cahn, *JOM* 25, 28 (1973).