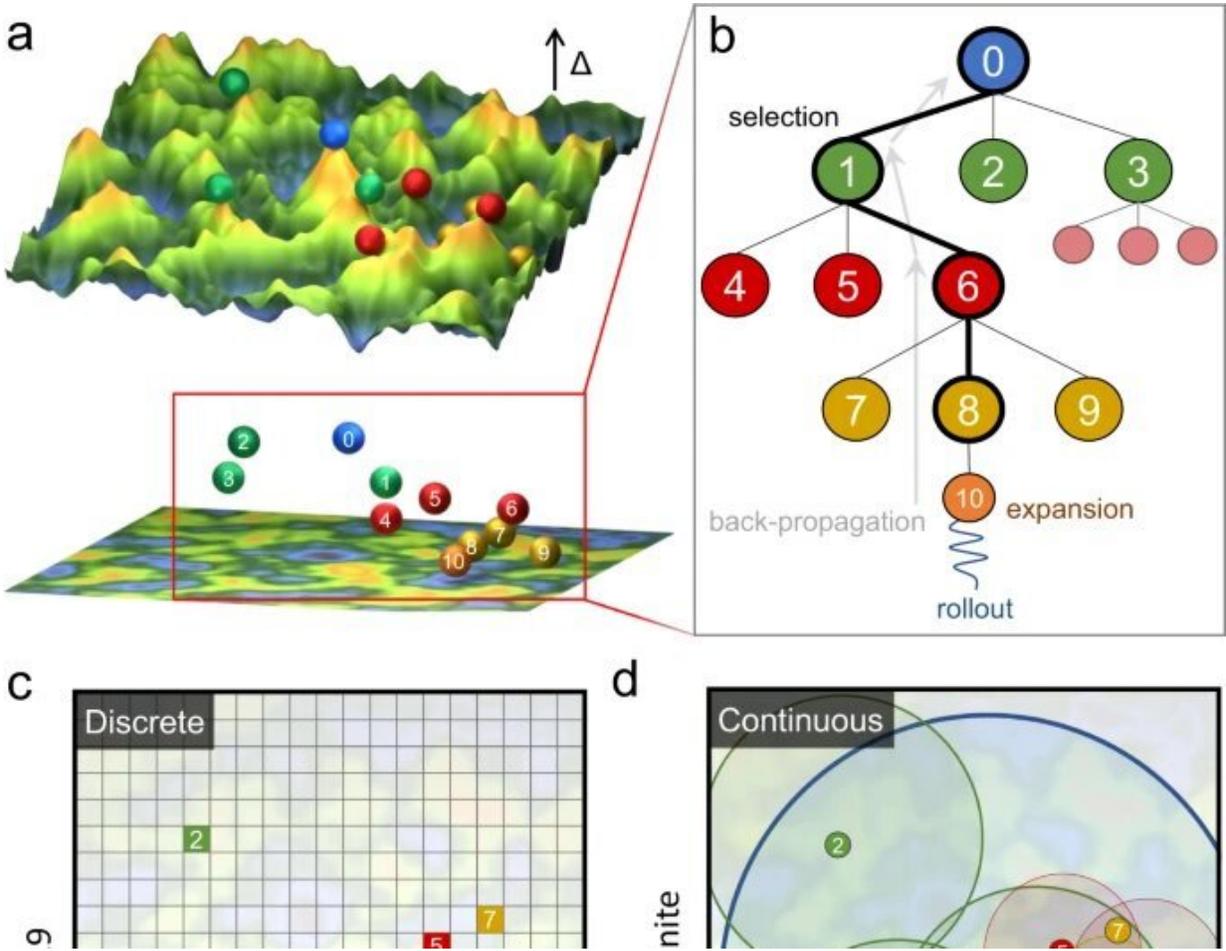


Machine learning explores materials science questions and solves difficult search problems

May 24 2022, by Elizabeth Ball



Schematic of the continuous action MCTS algorithm applied for exploration of high-dimensional potential parameter surfaces. a Top: Simplistic representation of an objective landscape for a two-parameter search problem. In-plane axes correspond to (two) independent model parameters. The out-of-plane axis

corresponds to objective values, which is defined as the weighted sum of the error in model predicted energies of clusters with respect to target energies. This objective is minimized by our c-MCTS algorithm. The spheres represent candidates of different model parameters within an MCTS run, where differences in their vertical positions indicate differences in their objective values. Bottom: Slightly tilted view of the above with the surface represented as a contour map below the spheres. The numbering on the spheres corresponds to their node positions in the MCTS tree shown in b. These numbers roughly correspond to the order that the candidates are explored. b Schematic showing the root, parent, child nodes, and their relationship within an MCTS tree structure. A typical MCTS search involves node selection, expansion, simulation (playout), and back-propagation. Different coloring of the nodes indicates different depths in the MCTS tree. The algorithm balances between exploration (lateral expansion of nodes) and exploitation (depth expansion of nodes). As shown in a, the objective value of an MCTS run is expected to decrease quickly along with the depth of the tree. c Search space of a traditional MCTS algorithm, e.g., game board, is discrete. In the context of parameter optimization, for two discrete parameters each of 19 possible values, the search space consists of a finite 361 search positions. d The problem of parameter search, such as the objective surface illustrated in a, generally involves parameters that are continuous, which corresponds to infinite possible search positions. We handle this challenge by applying a range-funneling technique to the MCTS algorithm where the search neighborhood at each tree-depth becomes smaller and smaller such that the algorithm can converge to the optimal solutions. Credit: *Nature Communications* (2022). DOI: 10.1038/s41467-021-27849-6

Using computing resources at the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory (Berkeley Lab), researchers at Argonne National Laboratory have succeeded in exploring important materials science questions and demonstrated progress using machine learning to solve difficult search problems.

By adapting a machine-learning algorithm from [board games](#) such as AlphaGo, the researchers developed force fields for nanoclusters of 54 elements across the periodic table, a dramatic leap toward understanding their [unique properties](#) and proof of concept for their search method. The team published its results in *Nature Communications* in January.

Depending on their scale—bulk systems of 100+ nanometers versus nanoclusters of less than 100 nanometers—materials can display dramatically different properties, including optical and magnetic properties, discrete energy levels, and enhanced photoluminescence. These properties may lend themselves to new scientific and industry applications, and scientists can learn about them by developing force fields—computational models that estimate the potential energies between atoms in a molecule and between molecules—for each element or compound. But materials scientists can spend years using traditional physics-based methods to explore the structures and forces between atoms in nanoclusters of a single element.

"We wanted to look at the nanoscale dynamics, and for that, usually we'd use some quantum calculus and density functional theory, but those are computationally very expensive calculations," said materials scientist Sukriti Manna, primary author on the paper, of the painstaking work of searching for and finding the parameters of potential models.

Applying machine learning is one potential way to cut that cost. However, the available algorithms come from discrete search spaces like games, where the number of search branches and possible outcomes is finite. In a continuous action space like force fields for chemical element nanoclusters, the number of possible search branches is infinite, and brute force—the ability to run every scenario to find the best outcome—simply doesn't work.

Working smarter, not harder

To make an existing algorithm work smarter, not harder, machine learning specialist Troy Loeffler used a type of [reinforcement learning](#) called Monte Carlo tree search (MCTS). Reinforcement learning is a form of machine learning that allows an algorithm to interact directly with its environment, learning through punishment and reward, with the goal of gaining the most cumulative reward over time. MCTS uses an "explore and exploit" method—initially searching randomly, then learning to ignore less productive search paths, or playouts, and focus on more productive ones. Loeffler also introduced a few new functions to make the algorithm more efficient: a uniqueness function to eliminate redundant searches, a window scaling scheme to correlate the tree depth to the action space to provide a useful bit of structure, and playout expansion, which teaches the algorithm to prioritize random searches that were closer to something that had already proven productive.

"A lot of the work we did was actually developing the algorithm for continuous action spaces, where you don't have nice, discrete board game spaces; you have parameters that can move anywhere on the particular landscape," said Loeffler. "The core idea is that you're using a combination of both complete randomness and a bit of a deterministic element, with the AI, to figure it out."



Two representations show the algorithm's effectiveness at predicting force fields for 54 elements across the periodic table. Credit: NERSC

The combination worked, yielding force fields for 54 elements in a fraction of the time it once would have taken to find parameters for just one element and proving that reinforcement learning can be a useful tool in continuous action spaces.

The team used the Cori supercomputer at NERSC to perform their calculations and generate both training and fitting datasets, primarily using Vienna Ab initio Simulation Package (VASP) software for atomic-scale materials modeling and the classical molecular-dynamics code LAMMPS. This project is just one of many at NERSC from the Theory and Modeling team at Argonne, who frequently take advantage of NERSC's computational power, minimal queues, and reliable maintenance.

"For elements such as carbon, boron, and phosphorous, we require a lot

of datasets and we require good quality, and for this particular work I use NERSC for generating tons of huge datasets because of their structural diversity. Cori is a very fast computer, and when I was using it, the queue time was very short, so we got that work done very quickly," said Manna. In addition, he said, "if we have 100% workload, for computational time, we depend on NERSC for 90% of that workload."

Machine learning specialist Rohit Batra, who developed a machine learning framework to analyze the error trends in potential functions across the periodic table, concurred. "I'm a big fan of Cori—I use it for several purposes," he said. "It's very well-maintained. Sometimes, in other clusters, there can be issues that cause them to be offline for quite a while, but I think NERSC is very well-maintained and very reliable in that way."

The future of MCTS goes deep and wide

Now that the use of MCTS in continuous search space has been demonstrated, what comes next? From a materials science perspective, there's more work to do exploring more complex materials.

"From an application perspective, a force field development perspective, we've explored elemental stuff and a few binary alloys, so in the near future we'll look into combinations, like oxides and sulfites, and develop those force fields," said Manna. "Because of the powerful algorithm, all we need is time and other training data sets."

But materials science isn't the only application of MCTS broken open by this work—and part of the next stage involves testing the breadth and boundaries of the algorithm's utility.

"We're taking MCTS and applying it to a lot of different situations," said Loeffler. "We have 10 or 11 different projects that we or our

collaborators are interested in using the algorithm for," including further games-oriented research and additional force-field fitting. Thus far, it's a process that has met with success, and its future looks bright, he added. "We're looking for a lot of things to try it on. But so far, everything we've tried it on, it's worked incredibly well."

More information: Sukriti Manna et al, Learning in continuous action space for developing high dimensional potential energy models, *Nature Communications* (2022). [DOI: 10.1038/s41467-021-27849-6](https://doi.org/10.1038/s41467-021-27849-6)

Provided by National Energy Research Scientific Computing Center

Citation: Machine learning explores materials science questions and solves difficult search problems (2022, May 24) retrieved 25 April 2024 from <https://techxplore.com/news/2022-05-machine-explores-materials-science-difficult.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.