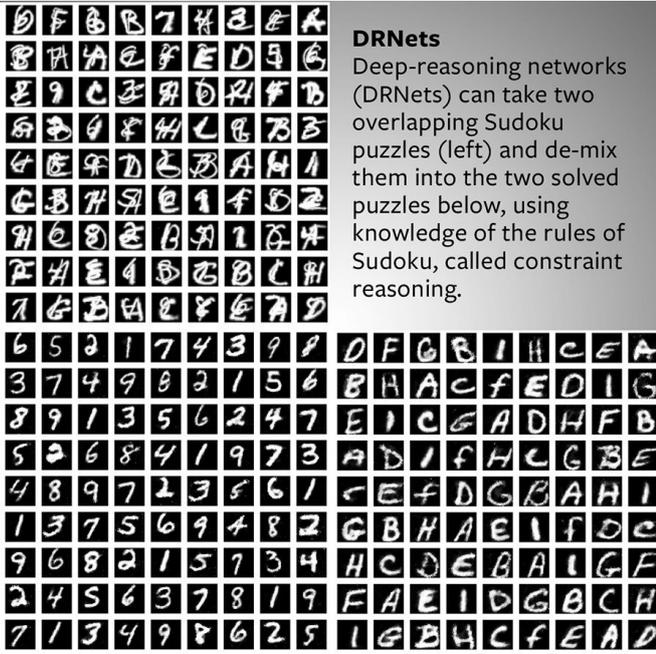


DRNets can solve Sudoku, speed scientific discovery

23 September 2021, by Tom Fleischman



IMAGES: DI CHEN, GOMES LAB/PROVIDED GRAPHIC: MATT FONDEUR/CORNELL UNIVERSITY

Credit: Cornell University

Say you're driving with a friend in a familiar neighborhood, and the friend asks you to turn at the next intersection. The friend doesn't say which way to turn, but since you both know it's a one-way street, it's understood.

That type of reasoning is at the heart of a new artificial-intelligence framework—tested successfully on overlapping Sudoku puzzles—that could speed discovery in [materials science](#), renewable energy technology and other areas.

An interdisciplinary research team led by Carla Gomes, the Ronald C. and Antonia V. Nielsen Professor of Computing and Information Science in the Cornell Ann S. Bowers College of Computing and Information Science, has developed Deep Reasoning Networks (DRNets), which combine

[deep learning](#)—even with a relatively small amount of data—with an understanding of the subject's boundaries and rules, known as "constraint reasoning."

Di Chen, a computer science doctoral student in Gomes' group, is first author of "Automating Crystal-Structure Phase Mapping by Combining Deep Learning with Constraint Reasoning," published Sept. 16 in *Nature Machine Intelligence*.

Gomes and John Gregoire, Ph.D. '09, a research professor at the California Institute of Technology, are the senior authors. Gregoire is a former postdoctoral researcher in the lab of co-author R. Bruce van Dover, the Walter S. Carpenter, Jr., Professor of Engineering.

DRNets, introduced at the 37th International Conference on Machine Learning, held virtually in July 2020, takes machine learning a step further by adding constraint reasoning—the ability to factor in rules and prior scientific knowledge, in order to solve problems with very little data as input.

You can teach a machine to recognize a dog by showing it 1,000 pictures of dogs, Gomes said, but scientific discovery is not like that.

"You are not going to have lots and lots of labeled data," she said. "And in general, the examples you have are not exactly what you are looking for, but then you reason about what you know scientifically about the domain, and you can infer new knowledge."

Gomes' group, which has been working on using AI and [machine learning](#) techniques to accelerate materials discovery for more than a decade, tested the DRNets framework by de-mixing overlapping handwritten Sudoku puzzles—grids with two numbers or letters in each box. The computer had to separate the puzzles into two solved Sudokus, without any training data, which it was able to

achieve with close to 100% accuracy.

noting that in the

The researchers then put DRNets to work on a real-world problem: automating crystal-structure phase mapping of solar-fuels materials, using X-ray diffraction (XRD) patterns. Crystal-structure phase mapping involves separating the source XRD signals of the desired crystal structures from "noisy" mixtures of XRD patterns, a task for which labeled training data are typically not available.

Sudoku experiments, "the machine has never seen what a '6' and 'D' overlap looks like, but can solve the problem by reasoning, using prior knowledge about Sudoku rules.

"In the same way," she said, "DRNets reason about thermodynamic rules and known crystal phases to demix the XRD patterns, without data to train on."

Using the understood thermodynamic rules, a few bits of unlabeled data, a total of 307 XRD patterns and minimal information regarding the elements of the chemical system—in this case, bismuth, copper and vanadium (Bi-Cu-V) oxide—DRNets was able to identify and separate a total of 13 crystal phases (single-phase materials) in 19 unique mixtures of the single-phase materials.

DRNets builds off the group's previous work involving citizen science related to species distribution, done in conjunction with the Cornell Lab of Ornithology's eBird program. The need to capture and interpret interactions between species and their local environments was the initial motivation and inspiration for the interpretable latent-space in the DRNets framework, said Gomes, a pioneer in the emerging field of computational sustainability.

DRNets' findings, verified using manual analysis, enable the discovery of complex mixtures of crystalline materials that convert solar energy into storable solar chemical fuels.

More information: Di Chen et al, Automating crystal-structure phase mapping by combining deep learning with constraint reasoning, *Nature Machine Intelligence* (2021). [DOI: 10.1038/s42256-021-00384-1](https://doi.org/10.1038/s42256-021-00384-1)

"The 13 phases and their mixtures comprise the scientific knowledge derived from the thousands of features in the measured XRD patterns," Gregoire said, emphasizing that human experts and prior algorithms "were unable to extract this knowledge from the XRD patterns due to the high level of complexity. Humans can reason about the physical rules and computers can process complex data, but scientific discovery requires integration of these approaches."

Provided by Cornell University

Said Gomes: "Verifying that a chemical system solution satisfies the physics rules is easier than producing it, the same way checking that a completed Sudoku is correct is easier than completing it."

Key to DRNets is the idea of an "interpretable latent space." Basically, it gives DRNets the ability to reason about the constraints of the domain—in this case materials science—from input data.

"This is really the big advancement of our methodology: We are doing this without having data for the computer to train on," Gomes said,

APA citation: DRNets can solve Sudoku, speed scientific discovery (2021, September 23) retrieved 8 December 2021 from <https://techxplore.com/news/2021-09-drnets-sudoku-scientific-discovery.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.