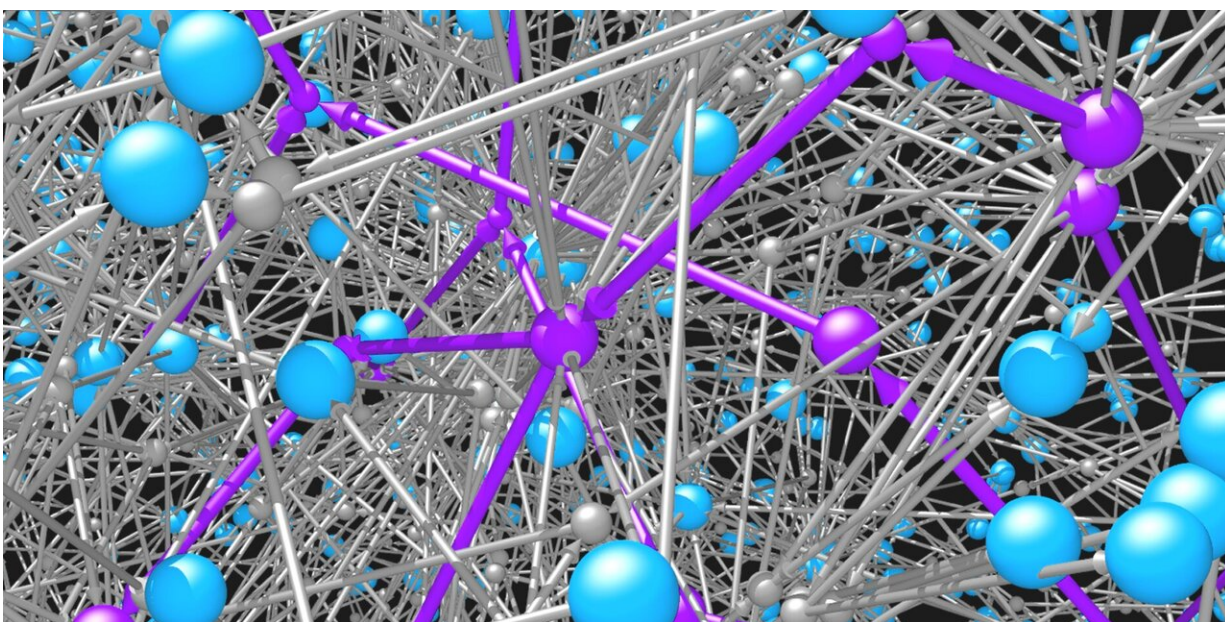


Deep learning techniques teach neural model to 'play' retrosynthesis

June 5 2019



Molecules (blue spheres) are connected to one another by the reactions (grey spheres and arrows) in which they participate. The network of possible organic molecules and reactions is impossibly vast. Intelligent search algorithms are needed to identify feasible pathways (purple) for synthesizing desired molecules. Credit: Mikolaj Kowalik & Kyle Bishop/Columbia Engineering

Researchers, from biochemists to material scientists, have long relied on the rich variety of organic molecules to solve pressing challenges. Some molecules may be useful in treating diseases, others for lighting our digital displays, still others for pigments, paints, and plastics. The unique

properties of each molecule are determined by its structure—that is, by the connectivity of its constituent atoms. Once a promising structure is identified, there remains the difficult task of making the targeted molecule through a sequence of chemical reactions. But which ones?

Organic chemists generally work backwards from the target molecule to the starting materials using a process called retrosynthetic analysis. During this process, the chemist faces a series of complex and inter-related decisions. For instance, of the tens of thousands of different chemical reactions, which one should you choose to create the target molecule? Once that decision is made, you may find yourself with multiple reactant molecules needed for the reaction. If these molecules are not available to purchase, then how do you select the appropriate reactions to produce them? Intelligently choosing what to do at each step of this process is critical in navigating the huge number of possible paths.

Researchers at Columbia Engineering have developed a [new technique](#) based on reinforcement learning that trains a [neural network model](#) to correctly select the "best" reaction at each step of the retrosynthetic process. This form of AI provides a framework for researchers to design chemical syntheses that optimize user specified objectives such synthesis cost, safety, and sustainability. The new approach, published May 31 by *ACS Central Science*, is more successful (by ~60%) than existing strategies for solving this challenging search problem.

"Reinforcement learning has created computer players that are much better than humans at playing complex video games. Perhaps retrosynthesis is no different! This study gives us hope that reinforcement-learning algorithms will be perhaps one day better than human players at the 'game' of retrosynthesis," says Alán Aspuru-Guzik, professor of chemistry and computer science at the University of Toronto, who was not involved with the study.

The team framed the challenge of retrosynthetic planning as a game like chess and Go, where the combinatorial number of possible choices is astronomical and the value of each choice uncertain until the synthesis plan is completed and its cost evaluated. Unlike earlier studies that used heuristic scoring functions—simple rules of thumb—to guide retrosynthetic planning, this new study used reinforcement learning techniques to make judgments based on the neural model's own experience.

"We're the first to apply reinforcement learning to the problem of retrosynthetic analysis," says Kyle Bishop, associate professor of chemical engineering. "Starting from a state of complete ignorance, where the model knows absolutely nothing about strategy and applies reactions randomly, the model can practice and practice until it finds a strategy that outperforms a human-defined heuristic."

In their study, Bishop's team focused on using the number of [reaction](#) steps as the measurement of what makes a "good" synthetic pathway. They had their [reinforcement learning](#) model tailor its strategy with this goal in mind. Using simulated experience, the team trained the model's neural network to estimate the expected synthesis cost or value of any given molecule based on a representation of its molecular structure.

The team plans to explore different goals in the future, for instance, training the model to minimize costs rather than the number of reactions, or to avoid [molecules](#) that could be toxic. The researchers are also trying to reduce the number of simulations required for the model to learn its strategy, as the training process was quite computationally expensive.

"We expect that our retrosynthesis game will soon follow the way of chess and Go, in which self-taught algorithms consistently outperform human experts," Bishop notes. "And we welcome competition. As with chess-playing computer programs, competition is the engine for

improvements in the state-of-the-art, and we hope that others can build on our work to demonstrate even [better performance](#)."

The study is titled "Learning retrosynthetic planning through simulated experience."

More information: John S. Schreck et al, Learning Retrosynthetic Planning through Simulated Experience, *ACS Central Science* (2019). [DOI: 10.1021/acscentsci.9b00055](https://doi.org/10.1021/acscentsci.9b00055)

Provided by Columbia University School of Engineering and Applied Science

Citation: Deep learning techniques teach neural model to 'play' retrosynthesis (2019, June 5) retrieved 18 April 2024 from <https://phys.org/news/2019-06-deep-techniques-neural-retrosynthesis.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.