

Chemistry and computer science join forces to apply artificial intelligence to chemical reactions

February 5 2021, by Wendy Plump



Credit: CC0 Public Domain

In the past few years, researchers have turned increasingly to data science techniques to aid problem-solving in organic synthesis.

Researchers in the lab of Abigail Doyle, Princeton's A. Barton Hepburn Professor of Chemistry, collaborated with Professor of Computer Science Ryan Adams to develop [open-source software](#) that provides them with a state-of-the-art [optimization](#) algorithm to use in everyday work, folding what's been learned in the machine learning field into synthetic chemistry.

The [software](#) adapts key principles of Bayesian optimization to allow faster and more efficient syntheses of chemicals.

Based on the Bayes Theorem, a mathematical formula for determining conditional probability, Bayesian optimization is a widely used strategy in the sciences. Broadly defined, it allows people and computers use prior knowledge to inform and optimize future decisions.

The chemists in Doyle's lab, in collaboration with Adams, a professor of computer science, and colleagues at Bristol-Myers Squibb, compared human decision-making capabilities with the [software package](#). They found that the [optimization tool](#) yields both greater efficiency over human participants and less bias on a test reaction. Their work appears in the current issue of the journal *Nature*.

"Reaction optimization is ubiquitous in [chemical synthesis](#), both in academia and across the [chemical industry](#)," said Doyle. "Since chemical space is so large, it is impossible for chemists to evaluate the entirety of a reaction space experimentally. We wanted to develop and assess Bayesian optimization as a tool for synthetic chemistry given its success for related optimization problems in the sciences."

Benjamin Shields, a former postdoctoral fellow in the Doyle lab and the paper's lead author, created the Python package.

"I come from a synthetic chemistry background, so I definitely

appreciate that synthetic chemists are pretty good at tackling these problems on their own," said Shields. "Where I think the real strength of Bayesian Optimization comes in is that it allows us to model these high-dimensional problems and capture trends that we may not see in the data ourselves, so it can process the data a lot better.

"And two, within a space, it will not be held back by the biases of a human chemist," he added.

How it works

The software started as an out-of-field project to fulfill Shields' doctoral requirements. Doyle and Shield then formed a team under the Center for Computer Assisted Synthesis (C-CAS), a National Science Foundation initiative launched at five universities to transform how the synthesis of complex organic molecules is planned and executed. Doyle has been a principal investigator with C-CAS since 2019.

"Reaction optimization can be an expensive and time-consuming process," said Adams, who is also the director of the Program in Statistics and Machine Learning. "This approach not only accelerates it using state-of-the-art techniques, but also finds better solutions than humans would typically identify. I think this is just the beginning of what's possible with Bayesian optimization in this space."

Users start by defining a search space—plausible experiments to consider—such as a list of catalysts, reagents, ligands, solvents, temperatures, and concentrations. Once that space is prepared and the user defines how many experiments to run, the software chooses initial experimental conditions to be evaluated. Then it suggests new experiments to run, iterating through a smaller and smaller cast of choices until the reaction is optimized.

"In designing the software, I tried to include ways for people to kind of inject what they know about a reaction," said Shields. "No matter how you use this or machine learning in general, there's always going to be a case where human expertise is valuable."

The software and examples for its use can be accessed at this repository. GitHub links are available for the following: software that represents the chemicals under evaluation in a machine-readable format via density-functional theory; software for reaction optimization; and the game that collects chemists' decision-making on optimization of the test reaction.

"Bayesian reaction optimization as a tool for chemical synthesis," by Benjamin J. Shields, Jason Stevens, Jun Li, Marvin Parasram, Farhan Damani, Jesus I. Martinez Alvarado, Jacob M. Janey, Ryan P. Adams and Abigail G. Doyle, appears in the Feb. 3 issue of the journal *Nature*.

More information: Benjamin J. Shields et al. Bayesian reaction optimization as a tool for chemical synthesis, *Nature* (2021). [DOI: 10.1038/s41586-021-03213-y](https://doi.org/10.1038/s41586-021-03213-y)

Provided by Princeton University

Citation: Chemistry and computer science join forces to apply artificial intelligence to chemical reactions (2021, February 5) retrieved 26 June 2024 from <https://phys.org/news/2021-02-chemistry-science-artificial-intelligence-chemical.html>

<p>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.</p>
--