

A way to use artificial intelligence to predict chemical reactions

December 8 2017, by Bob Yirka

"Found in Translation"
Predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models

Take the Tour!

What do I get here?

Our model has predicted product for those reactants. You can get access to the attention weight by clicking on the link here-below.

Basically those weights indicates which inputs token atoms are influencing the prediction for the next token atoms.

BACK DONE

Predict Reaction Get Smiles Example 1 Example 2

[RXN PREDICTION]
BrCCOC1CCCCO1.Cc1[nH]c2ccc(C(=O)O)cc2c1C>>Cc1[nH]c2ccc(C(=O)OCCOC3CCCCO3)cc2c1C

Confidence: 0.87
> Attention Weight

[SMILES]
BrCCOC1CCCCO1.CC1=C(C)C2C(=CC=C(C(=2)C(O)=O)N1

(Phys.org)—A team of researchers with [IBM](#) has applied artificial intelligence to predict organic chemical reactions. In their paper uploaded to the preprint server *arXiv*, the group outlines their approach, which they describe as an improvement over other models.

Predicting what will happen when chemicals are mixed or treated in certain ways is difficult because of all the variables involved. But scientists would like to have a tool that does it anyway, because it would dramatically speed up development of useful new materials, especially drugs. In this new effort, the team at IBM has taken an entirely new approach to creating such a tool.

The new approach involves treating chemical reactions as a translation problem by rephrasing elements in such predictions as letters and words rather than atoms and molecules. That changes the problem from one of predicting how chemicals will react to translating words from one form to another—a problem that has been mostly solved by AI systems.

In using such an [approach](#), the group was able to feed [chemical](#) components into a [neural network](#) trained on a dataset of 395,496 reactions. The neural network then used what it had learned about prior reactions to make predictions about what would occur under new conditions. In practice, the system responded to such requests by offering a top-five list of possible outcomes. Testing showed that the top prediction turned out to be correct 80 percent of the time, though the team has thus far only trained it on molecules with 150 atoms or less. They plan to keep working on the system and have a current goal of improving its accuracy to 90 percent. They also have plans for modifying it so that parameters such as heat, pH levels and solvents can be included. They even envision one-day hosting contests between their system and human chemists to demonstrate how well it works.

The group notes that the development of such a system is not meant to serve as a replacement for chemists, but instead to serve as a tool for them, to develop products faster or more cheaply. They plan to put the system on a cloud server so that anyone who wishes to use it may do so.

The team presented their work at this week's [Neural Information](#)

[Processing Systems conference.](#)

More information: "Found in Translation": Predicting Outcomes of Complex Organic Chemistry Reactions using Neural Sequence-to-Sequence Models, arXiv:1711.04810 [cs.LG] arxiv.org/abs/1711.04810

Abstract

There is an intuitive analogy of an organic chemist's understanding of a compound and a language speaker's understanding of a word. Consequently, it is possible to introduce the basic concepts and analyze potential impacts of linguistic analysis to the world of organic chemistry. In this work, we cast the reaction prediction task as a translation problem by introducing a template-free sequence-to-sequence model, trained end-to-end and fully data-driven. We propose a novel way of tokenization, which is arbitrarily extensible with reaction information. With this approach, we demonstrate results superior to the state-of-the-art solution by a significant margin on the top-1 accuracy. Specifically, our approach achieves an accuracy of 80.1% without relying on auxiliary knowledge such as reaction templates. Also, 66.4% accuracy is reached on a larger and noisier dataset.

© 2017 Phys.org

Citation: A way to use artificial intelligence to predict chemical reactions (2017, December 8) retrieved 19 April 2024 from <https://phys.org/news/2017-12-artificial-intelligence-chemical-reactions.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.