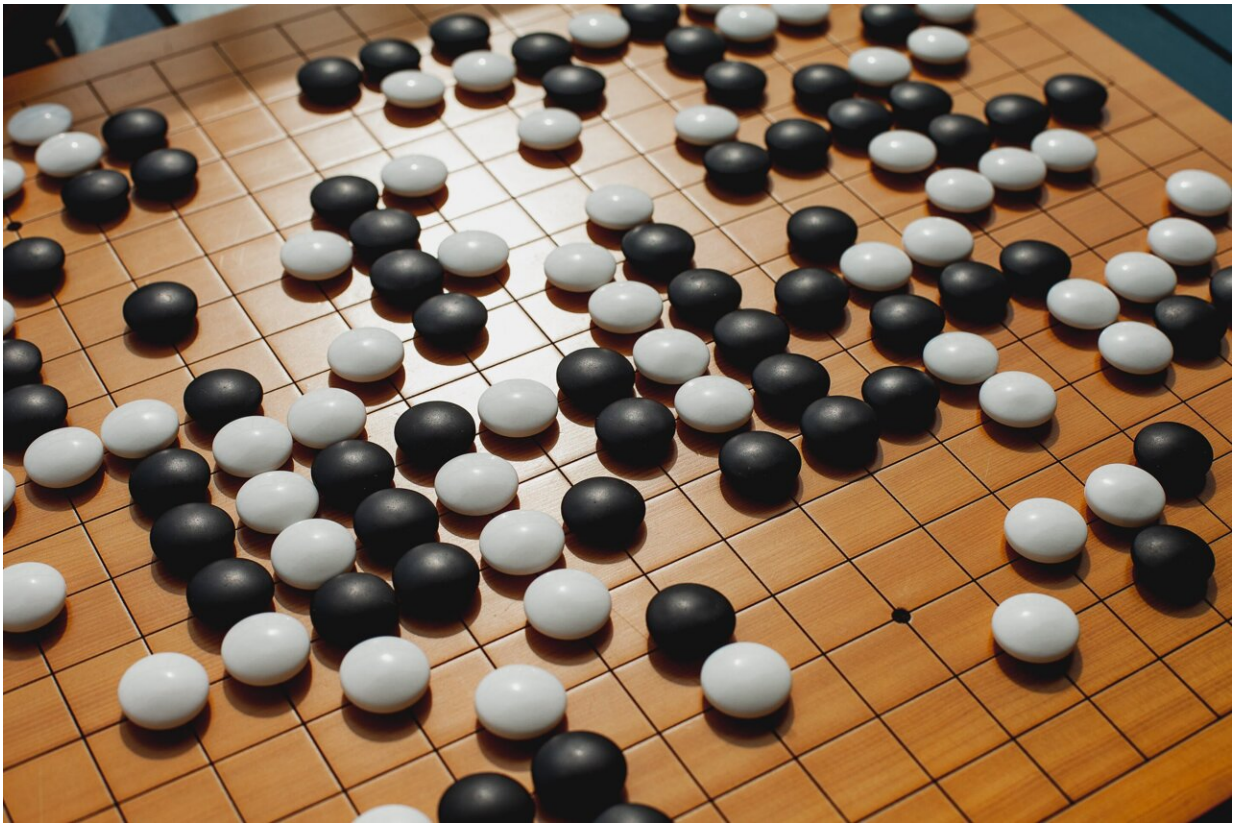
Reinforcement learning: From board games to protein design

April 20 2023, by Ian Haydon



Credit: Unsplash/CC0 Public Domain

Scientists have successfully applied reinforcement learning to a challenge in molecular biology. The team of researchers developed powerful new protein design software adapted from a strategy proven

adept at board games like Chess and Go. In one experiment, proteins made with the new approach were found to be more effective at generating useful antibodies in mice.

The findings, reported April 21 in *Science*, suggest that this breakthrough may soon lead to more potent vaccines. More broadly, the approach could lead to a new era in protein design.

"Our results show that reinforcement learning can do more than master board games. When trained to solve long-standing puzzles in protein science, the software excelled at creating useful molecules," said senior author David Baker, professor of biochemistry at the UW School of Medicine in Seattle and a recipient of the 2021 Breakthrough Prize in Life Sciences.

"If this method is applied to the right research problems," he said, "it could accelerate progress in a variety of scientific fields."

The research is a milestone in tapping artificial intelligence to conduct protein science research. The potential applications are vast, from developing more effective cancer treatments to creating new biodegradable textiles.

Reinforcement learning is a type of machine learning in which a computer program learns to make decisions by trying different actions and receiving feedback. Such an algorithm can learn to play chess, for example, by testing millions of different moves that lead to victory or defeat on the board. The program is designed to learn from these experiences and become better at making decisions over time.

To make a reinforcement learning program for protein design, the scientists gave the computer millions of simple starting molecules. The software then made ten thousand attempts at randomly improving each

toward a predefined goal. The computer lengthened the proteins or bent them in specific ways until it learned how to contort them into desired shapes.

Isaac D. Lutz, Shunzhi Wang, and Christoffer Norn, all members of the Baker Lab, led the research. Their team's *Science* manuscript is titled "Top-down design of protein architectures with reinforcement learning."

"Our approach is unique because we use reinforcement learning to solve the problem of creating protein shapes that fit together like pieces of a puzzle," explained co-lead author Lutz, a doctoral student at the UW Medicine Institute for Protein Design. "This simply was not possible using prior approaches and has the potential to transform the types of molecules we can build."

As part of this study, the scientists manufactured hundreds of AI-designed proteins in the lab. Using electron microscopes and other instruments, they confirmed that many of the protein shapes created by the computer were indeed realized in the lab.

"This approach proved not only accurate but also highly customizable. For example, we asked the software to make spherical structures with no holes, small holes, or large holes. Its potential to make all kinds of architectures has yet to be fully explored," said co-lead author Shunzhi Wang, a postdoctoral scholar at the UW Medicine Institute for Protein Design.

The team concentrated on designing new nano-scale structures composed of many protein molecules. This required designing both the protein components themselves and the chemical interfaces that allow the nano-structures to self-assemble.

Electron microscopy confirmed that numerous AI-designed nano-

structures were able to form in the lab. As a measure of how accurate the design software had become, the scientists observed many unique nano-structures in which every atom was found to be in the intended place. In other words, the deviation between the intended and realized nano-structure was on average less than the width of a single atom. This is called atomically accurate design.

The authors foresee a future in which this approach could enable them and others to create therapeutic proteins, vaccines, and other molecules that could not have been made using prior methods.

Researchers from the UW Medicine Institute for Stem Cell and Regenerative Medicine used primary cell models of blood vessel cells to show that the designed protein scaffolds outperformed previous versions of the technology. For example, because the receptors that help cells receive and interpret signals were clustered more densely on the more compact scaffolds, they were more effective at promoting blood vessel stability.

Hannele Ruohola-Baker, a UW School of Medicine professor of biochemistry and one of the study's authors, spoke to the implications of the investigation for regenerative medicine: "The more accurate the technology becomes, the more it opens up potential applications, including vascular treatments for diabetes, brain injuries, strokes, and other cases where blood vessels are at risk. We can also imagine more precise delivery of factors that we use to differentiate stem cells into various cell types, giving us new ways to regulate the processes of cell development and aging."

More information: Isaac D. Lutz et al, Top-down design of protein architectures with reinforcement learning, *Science* (2023). DOI: 10.1126/science.adf6591.
www.science.org/doi/10.1126/science.adf6591

Provided by University of Washington School of Medicine