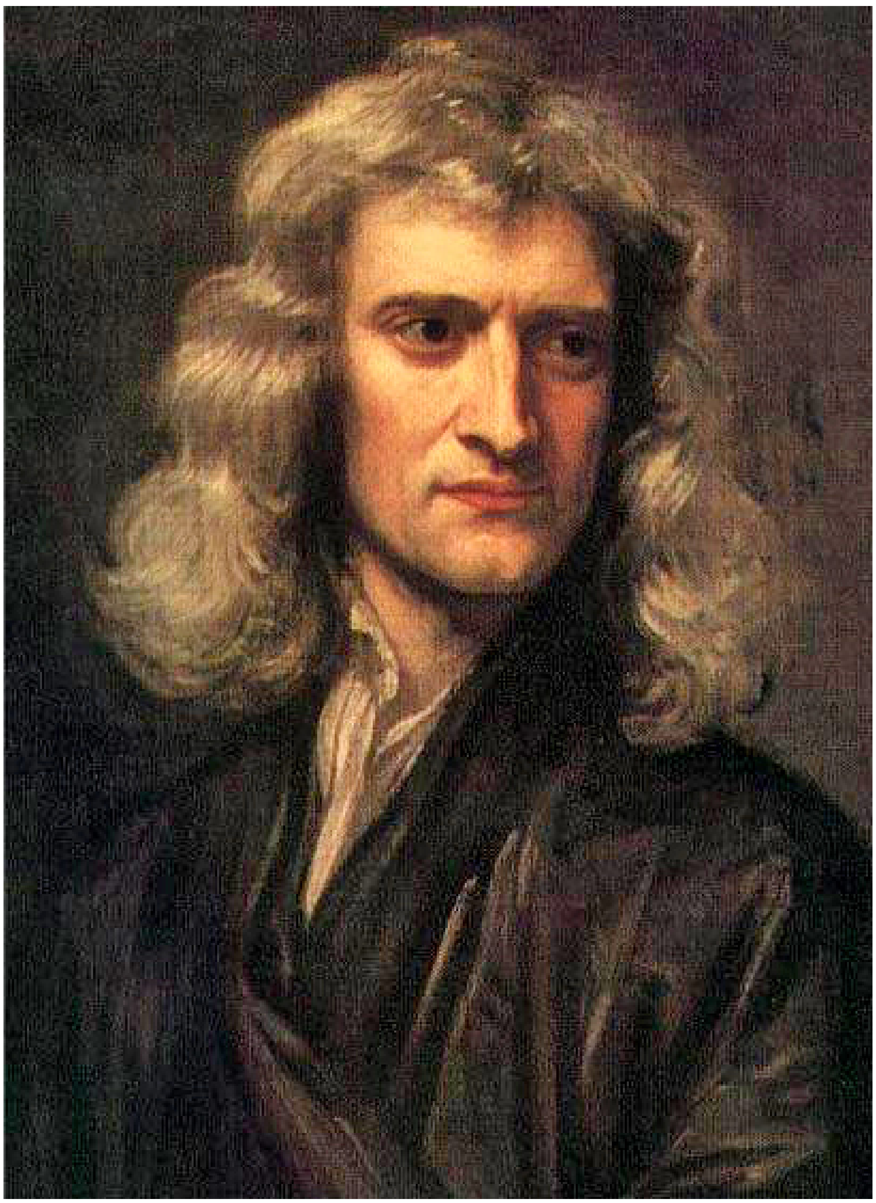
Biophysicists take small step in quest for 'robot scientist'

August 25 2015, by Carol Clark



The researchers dubbed their algorithm "Sir Isaac," in a nod to one of the greatest scientists of all time, Sir Isaac Newton. Credit: Public domain.

Biophysicists have taken another small step forward in the quest for an automated method to infer models describing a system's dynamics - a so-called robot scientist. *Nature Communications* published the finding - a practical algorithm for inferring laws of nature from time-series data of dynamical systems.

"Our algorithm is a small step," says Ilya Nemenman, lead author of the study and a professor of physics and biology at Emory University. "It could be described as a toy version of a robot scientist, but even so it may have practical applications. For the first time, we've taught a computer how to efficiently search for the laws that underlie arbitrary, natural dynamical systems, including complex, non-linear biological systems."

Nemenman's co-author on the paper is Bryan Daniels, a biophysicist at the University of Wisconsin.

Everything that is changing around and within us - from the relatively simple motion of celestial bodies, to weather and complex biological processes - is a dynamical system. A large part of science is guessing the laws of nature that underlie such systems, summarizing them in mathematical equations that can be used to make predictions, and then testing those equations and predictions through experiments.

"The long-term dream is to harness large-scale computing to make the guesses for us and speed up the process of discovery," Nemenman says.

While the quest for a true robot scientist, or computerized general intelligence, remains elusive, this latest algorithm represents a new approach to the problem.

"We think we have beaten any automated-inference algorithm that currently exists because we focus on getting an approximate solution to a problem, which we can get with much less data," Nemenman says.

In previous research, John Wikswo, a biophysicist at Vanderbilt University, along with colleagues at Cornell University, applied a software system to automate the scientific process for biological systems.

"We came up with a way to derive a model of cell behavior, but the approach is complicated and slow, and it is limited in the number of variables that it can track - it can't be scaled to more complicated systems," Wikswo says. "This new algorithm increases the speed of the necessary calculation by a factor of 100 or more. It provides an elegant method to generate compact and effective models that should allow prediction and control of complex systems."

Nemenman and Daniels dubbed their new algorithm "Sir Issac."

The real Sir Isaac Newton serves as a classic example of how the scientific method involves forming hypotheses, then testing them by looking at data and experiments. Newton guessed that the same rules of gravity applied to a falling apple and to the moon in orbit. He used data to test and refine his guess and generated the law of universal gravitation.

To test their algorithm, Nemenman and Daniels created an artificial, model solar system by generating numerical trajectories of planets and comets that move around a sun. In this simplified solar system, only the sun attracted the planets and comets.

"We trained our algorithm how to search through a group of laws which were limited enough to be practical, but also flexible enough to explain many different dynamics," Nemenman explains. "We then gave the algorithm some simulated planetary trajectories, and asked it what makes these planets move. It gave us the universal gravitational force. Not perfectly, but with very good accuracy. The error was just a few percent."

The algorithm also figured out that force changes velocity, not the position directly. "It gets Newton's First Law," Nemenman says, "the fact that in order to predict the possible trajectory of a planet, whether it stays near the sun or flies off into infinity, just knowing its initial position is not enough. The algorithm understands that you also need to know the velocity."

While most modern-day high school student know Newton's First Law, it took humanity 2,000 years beyond the time of Aristotle to discover it.

One limitation of the algorithm is inexactness. Getting an approximate model, however, is beneficial as long as the approximation is close enough to make good predictions, Nemenman says.

"Newton's laws are also approximate, but they have been remarkably beneficial for 350 years," he says. "We're still using them to control everything from electron microscopes to rockets."

Getting an exact description of any complex dynamical system requires large amounts of data, he adds. "In contrast, with our algorithm, we can get an approximate description by using just a few measurements of a system. That makes our method practical."

The researchers demonstrated, for example, that the algorithm can infer the dynamics of a caricature of an immune receptor in a leukocyte. This type of model could lead to a better understanding of the time-course for the response to an infection or a drug.

In another experiment, the researchers fed the algorithm data on concentrations of just three different species of chemicals involved in glycolysis in yeast. The algorithm generated a model that makes accurate predictions for the full system of this basic metabolic process to consume glucose, which involves seven chemical species.

"If you applied other methods of automatic inference to this system it would typically take tens of thousands of examples to reliably generate the laws that drive these chemical transformations," Nemenman says. "With our algorithm, we were able to do it with fewer than 100 examples."

With their experimental collaborators, the researchers are now exploring whether the algorithm can model more complex biological processes, such as the dynamics of insulin secretion in the pancreas and its relationship to the onset of a disease like diabetes. "The biology of insulin secreting cells is extremely complex. Understanding their dynamics on multiple scales is going to be difficult, and may not be possible for years with traditional methods," Nemenman says. "But we want to see if we can get a good enough approximation with our method to deliver a practical result."

The intuition of a genius mind like that of Isaac Newton is one quality that distinguishes human intelligence from even the highest-powered computer and algorithmic program.

"You can't give a machine intuition - at least for now," Nemenman says. "What we're hoping we can do is get our computer algorithm to spit out models of phenomena so that we, as scientists, can use them and our intuition to make useful generalizations. It's easier to generalize from models of specific systems then it is to generalize from various data sets directly."

More information: Automated adaptive inference of phenomenological dynamical models, *Nature Communications* 6, Article number: 8133 DOI: 10.1038/ncomms9133

Provided by Emory University