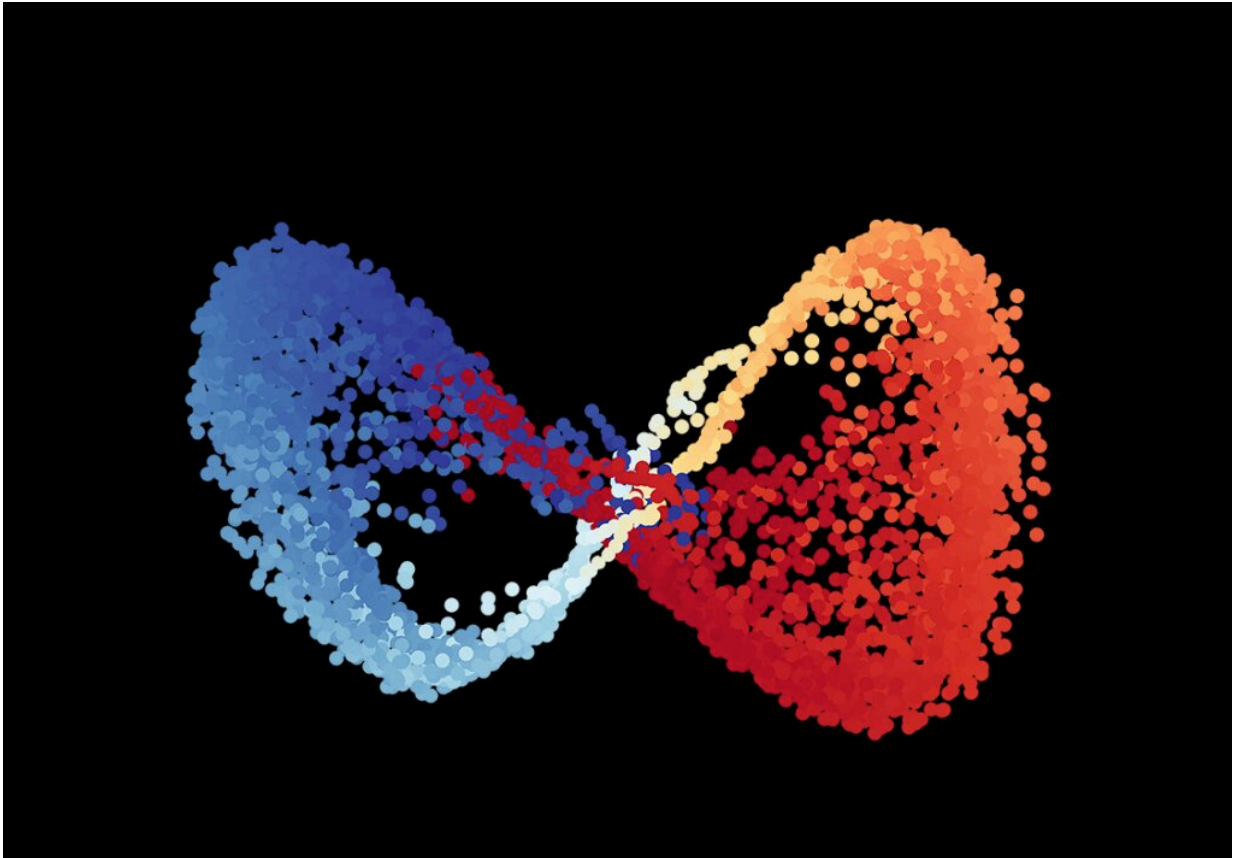# Roboticists discover alternative physics

July 26 2022, by Holly Evart

Latent embeddings from a framework colored by physical state variables. Credit: Boyuan Chen/Columbia Engineering

Energy, mass, velocity. These three variables make up Einstein's iconic equation $E=MC^2$. But how did Einstein know about these concepts in the first place? A precursor step to understanding physics is identifying

relevant variables. Without the concept of energy, mass, and velocity, not even Einstein could discover relativity. But can such variables be discovered automatically? Doing so could greatly accelerate scientific discovery.

This is the question that researchers at Columbia Engineering posed to a new AI program. The program was designed to observe physical phenomena through a video camera, then try to search for the minimal set of fundamental variables that fully describe the observed dynamics. The study was published on July 25 in *Nature Computational Science*.

The researchers began by feeding the system raw video footage of phenomena for which they already knew the answer. For example, they fed a video of a swinging double pendulum known to have exactly four "state variables"—the angle and angular velocity of each of the two arms. After a few hours of analysis, the AI produced the answer: 4.7.

"We thought this answer was close enough," said Hod Lipson, director of the Creative Machines Lab in the Department of Mechanical Engineering, where the work was primarily done. "Especially since all the AI had access to was raw video footage, without any knowledge of physics or geometry. But we wanted to know what the variables actually were, not just their number."

The researchers then proceeded to visualize the actual variables that the program identified. Extracting the variables themselves was not easy, since the program cannot describe them in any intuitive way that would be understandable to humans. After some probing, it appeared that two of the variables the program chose loosely corresponded to the angles of the arms, but the other two remain a mystery.

"We tried correlating the other variables with anything and everything we could think of: angular and linear velocities, kinetic and potential

energy, and various combinations of known quantities," explained Boyuan Chen Ph.D., now an assistant professor at Duke University, who led the work. "But nothing seemed to match perfectly." The team was confident that the AI had found a valid set of four variables, since it was making good predictions, "but we don't yet understand the mathematical language it is speaking," he explained.

After validating a number of other physical systems with known solutions, the researchers fed videos of systems for which they did not know the explicit answer. The first videos featured an "air dancer" undulating in front of a local used car lot. After a few hours of analysis, the program returned eight variables. A video of a lava lamp also produced eight variables. They then fed a video clip of flames from a holiday fireplace loop, and the program returned 24 variables.

A particularly interesting question was whether the set of variable was unique for every system, or whether a different set was produced each time the program was restarted.

"I always wondered, if we ever met an intelligent alien race, would they have discovered the same physics laws as we have, or might they describe the universe in a different way?" said Lipson. "Perhaps some phenomena seem enigmatically complex because we are trying to understand them using the wrong set of variables. In the experiments, the number of variables was the same each time the AI restarted, but the specific variables were different each time. So yes, there are alternative ways to describe the universe and it is quite possible that our choices aren't perfect."

The researchers believe that this sort of AI can help scientists uncover complex phenomena for which theoretical understanding is not keeping pace with the deluge of data—areas ranging from biology to cosmology. "While we used video data in this work, any kind of array data source

could be used—radar arrays, or DNA arrays, for example," explained Kuang Huang, Ph.D., who co-authored the paper.

The work is part of Lipson and Fu Foundation Professor of Mathematics Qiang Du's decades-long interest in creating algorithms that can distill data into scientific laws. Past software systems, such as Lipson and Michael Schmidt's Eureqa software, could distill freeform physical laws from experimental data, but only if the variables were identified in advance. But what if the variables are yet unknown?

Lipson, who is also the James and Sally Scapa Professor of Innovation, argues that scientists may be misinterpreting or failing to understand many phenomena simply because they don't have a good set of variables to describe the phenomena.

"For millennia, people knew about objects moving quickly or slowly, but it was only when the notion of velocity and acceleration was formally quantified that Newton could discover his famous law of motion F=MA," Lipson noted. Variables describing temperature and pressure needed to be identified before laws of thermodynamics could be formalized, and so on for every corner of the scientific world. The variables are a precursor to any theory.

"What other laws are we missing simply because we don't have the variables?" asked Du, who co-led the work.

The paper was also co-authored by Sunand Raghupathi and Ishaan Chandratreya, who helped collect the data for the experiments.