

Machines that learn better

May 18 2010, by Larry Hardesty



Cameron Freer, left, an instructor in pure mathematics; and Daniel Roy, right, a PhD student in the Department of Electrical Engineering and Computer Science. Photo: Jason Dorfman/CSAIL

(PhysOrg.com) -- In the last 20 years or so, many of the key advances in artificial-intelligence research have come courtesy of machine learning, in which computers learn how to make predictions by looking for patterns in large collections of training data. A new approach called probabilistic programming makes it much easier to build machine-learning systems, but it's useful for a relatively narrow set of problems. Now, MIT researchers have discovered how to extend the approach to a much larger class of problems, with implications for subjects as diverse as cognitive science, financial analysis and epidemiology.

Historically, building a machine-learning system capable of learning a new task would take a graduate student somewhere between a few weeks and several months, says Daniel Roy, a PhD student in the Department of Electrical Engineering and [Computer Science](#) who along with Cameron Freer, an instructor in pure mathematics, led the new research. A handful of new, experimental, probabilistic programming languages — one of which, Church, was developed at MIT — promise to cut that time down to a matter of hours.

At the heart of each of these new languages is a so-called inference algorithm, which instructs a machine-learning system how to draw conclusions from the data it's presented. The generality of the inference algorithm is what confers the languages' power: The same algorithm has to be able to guide a system that's learning how to recognize objects in digital images, or filter spam, or recommend DVDs based on past rentals, or whatever else an artificial-intelligence program may be called upon to do.

The inference algorithms currently used in probabilistic programming are great at handling discrete data but struggle with continuous data. For an idea of what that distinction means, consider three people of different heights. Their rank ordering, from tallest to shortest, is discrete: Each of them must be first, second, or third on the list. But their absolute heights are continuous. If the tallest person is 5 feet 10 inches tall, and the shortest is 5 feet 8 inches, you can't conclude that the third person is 5 feet 9 inches: He or she could be 5 feet 8.5 inches, or 5 feet 9.6302 inches or an infinite number of other possibilities.

Designers of probabilistic programming languages are thus avidly interested in whether it's possible to design a general-purpose inference algorithm that can handle continuous data. Unfortunately, the answer appears to be no: In a [yet-unpublished paper](#), Freer, Roy, and Nate Ackerman of the University of California, Berkeley, mathematically

demonstrate that there are certain types of statistical problems involving continuous data that no general-purpose algorithm could solve.

But there's good news as well: Last week, at the International Conference on Artificial Intelligence and Statistics, Roy presented a paper in which he and Freer not only demonstrate that there are large classes of problems involving continuous data that are susceptible to a general solution but also describe an inference algorithm that can handle them. A probabilistic [programming language](#) that implemented the algorithm would enable the rapid development of a much larger variety of machine-learning systems. It would, for instance, enable systems to better employ an analytic tool called the Pólya tree, which has been used to model stock prices, disease outbreaks, medical diagnoses, census data, and weather systems, among other things.

“The field of probabilistic programming is fairly new, and people have started coming up with probabilistic programs, but Dan and Cameron are really filling the theoretical gaps,” says Zoubin Ghahramani, professor of information engineering at the University of Cambridge. The hope, Ghahramani says, “is that their theoretical underpinnings will make the effort to come up with probabilistic programming languages much more solidly grounded.”

Chung-chieh Shan, a computer scientist at Rutgers who specializes in models of linguistic behavior, says that the MIT researchers' work could be especially useful for artificial-intelligence systems whose future behavior is dependent on their past behavior. For instance, a system designed to understand spoken language might have to determine words' parts of speech. If, in some context, it notices that a word tends to be used in an uncommon way — for instance, “man” is frequently used as a verb instead of a noun — then, going forward, it should have greater confidence in assigning that word its unusual interpretation.

Often, Shan explains, treating problems as having such “serial dependency” makes them easier to describe. But it also makes their solutions harder to calculate, because it requires keeping track of an ever-growing catalogue of past behaviors and revising future behaviors accordingly. Freer and Roy’s algorithm, he says, provides a way to convert problems that have serial dependency into problems that don’t, which makes them easier to solve. “A lot of models would call for this kind of picture,” Shan says. Roy and Freer’s work “is narrowing this gap between the intuitive description and the efficient implementation.”

While Freer and Roy’s algorithm is guaranteed to provide an answer to a range of previously intractable problems, Shan says, “there’s a difference between coming up with the right algorithm and implementing it so that it runs fast enough on an actual computer.” Roy and Freer agree, however, which is why they haven’t yet incorporated their algorithm into Church. “It’s fairly clear that within the set of models that our algorithm can handle, there are some that could be arbitrarily slow,” Roy says. “So now we have to study additional structure. We know that it’s possible. But when is it efficient?”

Provided by Massachusetts Institute of Technology

Citation: Machines that learn better (2010, May 18) retrieved 17 April 2024 from <https://phys.org/news/2010-05-machines.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.