Teaching machines to teach themselves

December 4 2017, by Arend Hintze

Credit: AI-generated image ([disclaimer](#))

Are you tired of telling machines what to do and what not to do? It's a large part of regular people's days – operating dishwashers, smartphones and cars. It's an even bigger part of life for researchers like me, working on artificial intelligence and machine learning.

Much of this is even more boring than driving or talking to a virtual assistant. The most common way of teaching computers new skills –

such as telling apart photos of dogs from ones of cats – involves a lot of human interaction or preparation. For instance, if a computer looks at a picture of a cat and labels it "dog," we have to tell it that's wrong.

But when that gets too cumbersome and tiring, it's time to build computers that can teach themselves, and retain what they learn. My research team and I have taken a first step toward the sort of learning that people imagine the robots of the future will be capable of – learning by observation and experience, rather than needing to be directly told every little step of what to do. We expect future machines to be as smart as we are, so they'll need to be able to learn like we do.

Setting robots free to learn on their own

In the most basic methods of training computers, the machine can use only the information it has been specifically taught by engineers and programmers. For instance, when researchers want a machine to be able to classify images into different categories, such as telling apart cats and dogs, we first need some reference pictures of other cats and dogs to start with. We show these pictures to the machine, and when it guesses right we give positive feedback, and when it guesses wrong we apply negative feedback.

This method, called reinforcement learning, uses external feedback to teach the system to change its internal workings in order to guess better next time. This self-change involves identifying the factors that made the biggest differences in the algorithm's decision, reinforcing accuracy and discouraging wrong decisions.

Another layer of advancement sets up another computer system to be the supervisor, rather than a human. This lets researchers create several dog-cat classifier machines, each with different attributes – perhaps some look more closely at color, while others look more closely at ear or nose

shape – and evaluate how well they work. Each time each machine runs, it looks at a picture, makes a decision about what it sees and checks with the automated supervisor to get feedback.

Alternatively or in addition, we researchers turn off the classifier machines that don't do as well, and introduce new changes to the ones that have done well so far. We repeat this many times, introducing small mutations into successive generations of classifier machines, slowly improving their abilities. This is a digital form of Darwinian evolution – and it's why this type of training is called a "genetic algorithm." But even that requires a lot of human effort – and telling cats and dogs apart is an extremely simple task for a person.

Learning like people

Our research is working toward a shift from a present in which machines learn simple tasks with human supervision, to a future in which they learn complicated processes on their own. This mirrors the development of human intelligence: As babies we were equipped with pain receptors that warned us about physical damage, and we had an instinct to cry when hungry or otherwise in need.

Human babies learn a lot on their own, and also learn a lot from direct instruction by parents specifically teaching vocabulary and specific behaviors. In the process, they learn not only how to interpret positive and negative feedback, but how to tell the difference – all on their own. We're not born knowing that the phrase "good job" means something positive, and that the threat of a "timeout" implies negative consequences. But we figure it out – and quite quickly. As adults, we can set our own goals and learn to accomplish them fully autonomously; we are our own teachers.

Our brains add each new experience and insight to our abilities and

memories, using a [capability called neuroplasticity](#) to [make and store new connections between neurons](#). There are [several ways](#) to use [neuroplasticity in computational systems](#), but these computational methods all still rely on feedback from an outside supervisor – something externally tells them what is right and wrong. (The method called "[unsupervised learning](#)" is not quite accurately named: It doesn't involve algorithms that can change themselves, and used a process quite different from what humans would understand as "learning.")

Figuring out a maze puzzle

The recent research my group and I have conducted takes a first step toward AI systems with neuroplasticity that do not require supervision. A key problem in doing this involves how to get a computer to give itself feedback that is somehow meaningful or effective.

We didn't actually know how to do that – in fact, it's one of the things we're learning about while analyzing our results. We use [Markov Brains](#), a type of artificial neural network, as the basis of our research. But instead of designing them directly, we used another [machine learning](#) technique, a genetic algorithm, to train these Markov Brains.

The challenge we set was to solve a [maze](#) using four buttons, which moved forward, backward, left and right. But the controls' functions changed for each new maze – so the button that meant "forward" last game might mean "left" or "backward" in the next. For a person solving this challenge, the reward would be not only in navigating through the maze but also in figuring out how the buttons had changed – in learning.

Evolving a good solution-finder

In our setup, the Markov Brains that solved mazes fastest – the ones that

learned the controls and moved through the maze most quickly – survived the genetic selection process. At the beginning of the process, each algorithm's actions were pretty much random. Just as with human players, randomly hitting buttons will only rarely get through the maze – but that strategy will succeed more often than doing nothing at all, or even just pressing the same button over and over.

If our research had involved keeping the buttons and maze structure constant, the Markov Brains would eventually learn what the buttons meant and how to get through the maze most quickly. They would immediately hit the correct sequence of buttons, without paying attention to the environment. That's not the sort of learning we're aiming for.

By randomizing both the button configurations and the maze structure, we force the Markov Brains to pay more attention, pressing a button and noticing the change to the situation – what direction that button moved through the maze, and whether that is toward a dead end or a wall or an open pathway. This is more advanced learning, to be sure. But a Markov Brain that evolved to navigate using only one or two button configurations could still do well: It would solve at least some mazes very quickly – even if it didn't solve others at all. That doesn't provide the adaptability to the environment that we're looking for.

The genetic algorithm, which decides which Markov Brains to select for further evolution and which to discontinue, is the key to optimizing response to the environment. We told it to select the Markov Brains that were the best overall solvers of mazes (rather than those that were blindingly fast on some mazes but utterly unable to solve others), choosing generalists over specialists.

Over many generations, this process produces Markov Brains that are particularly observant of the changes that result from pressing a particular button and very good at interpreting what those mean:

"Pressing the button that moves left took me into a dead end; I should press the button that moves right to get out of there."

It is this ability to interpret observations that liberates the [genetic algorithm](link)-Markov Brain system from the outside feedback of supervised learning. The Markov Brains have been selected specifically for their ability to create internal feedback that changes their structure in ways that lead to pressing the correct button at the correct time more often. Technically, we evolved Markov Brains to be able to learn by themselves.

This is indeed very similar to how humans learn: We try something, look at what happened and use the results to do better the next time. All of that happens within our brains, without the need for an external guide.

Our work [adds a new method to the field of machine learning](link), and in our view takes a major step toward developing what is called "general artificial intelligence," systems that can learn new information and new skills on their own. It also opens the door for using computer systems to [test how learning actually happens](link).

This article was originally published on [The Conversation](link). Read the [original article](link).

Provided by The Conversation

Citation: Teaching machines to teach themselves (2017, December 4) retrieved 9 April 2024 from https://techxplore.com/news/2017-12-machines.html