



Mad Men: Is That All There Is?
By Sophie Gilbert, David Sims, and Lenika Cruz



Rolling Stone's Sensational Failure
By Conor Friedersdorf



How Apple Pay Gets People to Spend
By Joe Pinsker



Wolf Hall: A PBS Masterpiece
By Sophie Gilbert

Teaching a Computer Not to Forget

One of the keys to unlocking artificial intelligence will be to figure out why biological brains are so good at remembering old skills—even when learning new things.

ADRIENNE LAFRANCE | APR 6 2015, 11:00 AM ET

Imagine if every time you learned something new, you completely forgot how to do a thing you'd already learned.

Finally figured out that taxi-hailing whistle? Now you can't tie your shoes anymore. Learn how to moonwalk; forget how to play the violin. Humans do forget skills, of course, but it usually happens gradually.

Computers forget what they know more dramatically. Learning cannibalizes knowledge. As soon as a new skill is learned, old skills are crowded out. It's a problem computer scientists call "catastrophic forgetting." And it happens because computer brains often rewire themselves—forging new and different connections across neural pathways—every time they learn. This makes it hard for a computer to retain old lessons, but also to learn tasks that require a sequence of steps.

"Researchers will need to solve this problem of catastrophic forgetting for us to get anywhere in terms of producing artificially intelligent computers and robots," said Jeff Clune, an assistant professor of computer science at the University of Wyoming. "Until we do, machines will be mostly one-trick ponies."



The elements of a changing world, from technology and business to politics and culture
[Read More](#)

Catastrophic forgetting also stands in the way of one of the long-standing goals for artificial intelligence: to create computers that can compartmentalize different skills in order to solve diverse problems.

So what would it take for a computer brain to retain what it knows, even as it learns new things? That was the question Clune had when he and his colleagues set out to make an artificial brain act more like a human one. Their central idea: See if you can get a computer to organize—and preserve—what it knows within distinct modules of the brain,

rather than overwriting what it knows every time it learns something new.

"Biological brains exhibit a high degree of modularity, meaning they contain clusters of neurons with high degrees of connectivity within clusters, but low degrees of connectivity between clusters," the team explained [in a video](#) about their research, which was [published last week in the journal PLoS Computational Biology](#).

In humans and animals, brain modularity evolved as the optimal way to organize neural connections. That's because natural selection arranges the brain to minimize the costs associated with building, maintaining, and housing broader connections. "It is an interesting question as to how evolution solved this problem," Clune told me. "How did it figure out how to allow animals, including

VIDEO

A Stop-Motion Tour of New York City
A filmmaker animated hundreds of still photographs to create this Big Apple flip book

MORE IN TECHNOLOGY



Even Computers Are Unsure About Their March Madness Brackets
ADRIENNE LAFRANCE



Further Proof That It's Hard to Be a Mets Fan
ROBINSON MEYER



America's Immortal Cereal
RACHEL SMITH

us, to learn a new skill without overwriting the knowledge of a previously learned skill?"

In order to encourage modularity in a computer's brain, researchers incorporated what they call "connection costs"—essentially showing the computer that modularity is preferable. Then they measured the extent to which a computer remembered an old skill once it learned a new one.

And, indeed, modularity appears to help computers, like humans, compartmentalize what they know. Which means that modularity may be one of the keys to ending catastrophic forgetting. "This paper both illuminates how natural evolution may have solved this problem, and how machine intelligence may learn to overcome it as well: by having modular brains," Clune said. "I believe modularity will be one key focus going forward that will improve AI's ability to learn a variety of different skills, just like humans can."

"At present, there are more differences between human and computer intelligence than similarities," Clune said, "but we are slowly closing that gap."

PRESENTED BY

Jump to Comments



ADRIENNE LAFRANCE is a senior associate editor at *The Atlantic*, where she oversees the Technology Channel. Previously she worked as an investigative reporter for [Honolulu Civil Beat](#), [Nieman Journalism Lab](#), and [WBUR](#). [MORE](#)

[ALL POSTS](#) |



The Atlantic PHOTO



JUST IN

The Golden Age of Online Music Is Over (and Another Is Beginning)

SPENCER KORNHABER

How to Get Around Censors in China

HEATHER TIMMONS

Why Can't the Left Unseat Rahm Emanuel?

DAVIDA. GRAHAM

MOST POPULAR

1 A Centenarian's Advice to Young Women: Figure Out What You Want