
Robotics

What Happens When You Give an AI a Working Memory?

It solves puzzles in a surprisingly human-like way.

by Will Knight October 12, 2016

A new kind of computer, devised by researchers at Google DeepMind in the U.K., could broaden the abilities of today's best AI systems by giving them an important new feature—a kind of working memory.

The researchers show that the computer, which consists of a large neural network connected to a unique form of memory, can perform relatively complex tasks by figuring out for itself what information to hold in its memory. The tasks include figuring out the best way to get from one station to another on London's spaghetti-like Underground transit network, after exploring diagrams of other types of networks and learning about the most salient features.

The Google DeepMind researchers call their system a differentiable neural computer. It is differentiable in the sense that its behavior—including what to store in memory—can be learned using the mathematical process, called backpropagation, that underlies the working of neural networks. As the network is trained with data, it will automatically store some information to a memory matrix.

“Like a conventional computer, it can use its memory to represent and manipulate complex data structures but, like a neural network, it can learn to do so from data,” the authors, which include Alex Graves, Greg Wayne, and Demis Hassabis, write in a paper published today in the journal *Nature*.

The advance is a step toward artificial intelligence that is a little more human-like in its abilities. While the technique is limited for now, systems built this way might someday perform useful work, says **Ruslan Salakhutdinov**, an associate professor at CMU who specializes in machine learning and AI. For example, a more advanced version might crawl Wikipedia and figure out what significant concepts, like names, places, and dates, to store in memory. Or it might allow a robot to use information learned in one setting in a completely new one. “It’s a very exciting piece of work,” Salakhutdinov says.

The latest machine-learning systems are brilliant at certain tasks, like recognizing faces in images or spoken words. And with practice they can learn to perform complex tasks like playing computer games to an expert level. But they require huge quantities of specific data for training, and unlike a human, cannot store much of what they have learned in memory for use later. This presents a problem in many areas, including language (see “**AI’s Language Problem**”).

Salakhutdinov notes, however, that making such a differentiable neural computer more complex could be difficult. That’s because in order to access its memory, it has to perform a complex calculation querying every stored piece. “It’s super difficult to get these things to work,” he says. “Scaling up can be a bit problematic.”

Interestingly, the work brings two fields of AI that have long been at loggerheads a little closer together. Early work in artificial intelligence involved programming machines to represent information symbolically, while the current vogue is to use very large neural networks that train themselves to perform tasks. For a long time, some AI traditionalists and cognitive scientists have questioned whether neural networks can do what humans do without gaining some deeper ability to represent information symbolically.

“I am most impressed by the network’s ability to learn ‘algorithms’ from examples,” says **Brenden Lake**, a cognitive scientist at New York

University who studies ways of making computers mimic human intelligence. This could expand the usefulness of deep learning. “Algorithms, such as sorting or finding shortest paths, are the bread and butter of classic computer science. They traditionally require a programmer to design and implement.”

But Lake points out that the system is still not quite human-like in the way it works. “People can pick up a new task from a limited amount of experience, especially if they are familiar with the domain,” he says. “In contrast, the differential neural computer is trained on tens or hundreds of thousands of examples of each task. I think the human ability to quickly learn new tasks will be one of the next major AI challenges.”

Learn more about AI at EmTech Digital 2017.

Register now