

COMPUTER SCIENCE

# just add memory

New types of electronic components, closer to neurons than to transistors, are leading to tremendously efficient and faster “memcomputing”

*By Massimiliano Di Ventra and Yuriy V. Pershin*

## IN BRIEF

**Modern computers** all use a unit that does calculations and a separate memory unit that holds programs and data. Shuttling information back and forth takes lots of energy and time.

**A new idea, memcomputing** works in a way that is similar to the neurons in the human brain, which form computing and memory storage units that are physically the same.

**This could mean** a giant leap in computer speed and efficiency, as well as new computing architectures, so scientists are trying to learn the best ways to use different memcomputing components.

**Massimiliano Di Ventra** is a professor in the department of physics at the University of California, San Diego.

**Yuriy V. Pershin** is an associate professor in the department of physics and astronomy at the University of South Carolina.



# W

HEN WE WROTE THE WORDS YOU ARE NOW READING, WE WERE TYPING on the best computers that technology now offers: machines that are terribly wasteful of energy and slow when tackling important scientific calculations. And they are typical of every computer that exists today, from the smartphone in your hand to the multimillion-dollar supercomputers humming along in the world's most advanced computing facilities.

We were writing in Word, a perfectly fine program that you probably use as well. To write “When we wrote the words you are now reading,” our computer had to move a collection of 0’s and 1’s—the machine representation of a Word document—from a temporary memory area and send it to another physical location, the central processing unit (CPU), via a bunch of wires. The processing unit transformed the data into the letters that we saw on the screen. To keep that particular sentence from vanishing once we turned our computer off, the data representing it had to travel back along that bunch of wires to a more stable memory area such as a hard drive.

This two-step shuffle happens because, at the moment, computer memory cannot do processing, and processors cannot store memory. It is a standard division of labor, and it happens even in fancy computers that do the fastest kind of calculating, called parallel processing, with multiple processors. The trouble is that each of these processors is still hobbled by this limitation.

Scientists have been developing a way to combine the previously uncombinable: to create circuits that juggle numbers and store memories at the same time. This means replacing standard computer circuit elements such as transistors, capacitors and inductors with new components called memristors, memcapacitors and meminductors. These components exist right now, in experimental forms, and could soon be combined into a new type of machine called a memcomputer.

Memcomputers could have unmatched speed because of their dual abilities. Each part of a memcomputer can help compute the answer to a problem, in a new, more efficient version of today’s parallel computing. And because difficult problems are solved by the computer’s memory and stored directly into that

memory, they will also save all the energy that is currently required to transfer data back and forth within the machine. This brand-new type of computing architecture would change the way computers of all types operate, from the tiny chips in your phone to vast supercomputers. It is, in fact, a design that is close to the way the human brain works, holding memories and processing information in the same neurons.

These new memcomputing machines should be much swifter—taking mere seconds to do calculations that would take current machines decades—and smaller and use much less electricity. Complete memcomputers have not yet been built, but our experiments with the components indicate that they could have a huge impact on computer design, global sustainability, power use and our ability to answer vital scientific questions.

### AN ELECTRONIC, ENERGY-EFFICIENT BRAIN

IT TAKES A TINY BIT OF ELECTRICITY and a fraction of a second to shuffle data like our Word sentence within a machine. But if you think about what happens when the energy for this back and forth is multiplied across worldwide computing use, it is an enormous operation.

Between 2011 and 2012 power requirements for computer data centers around the globe grew by a staggering 58 percent. It is not just supercomputers. Add in every gadget in every house, from ovens to laptops to televisions, that now has some computing ability. Combined, the information and communication sectors now account for approximately 15 percent of global electricity consumption. By 2030 the global electricity use by consumer electronics will equal the current total residential electricity consumption of the U.S. and Japan *combined* and will cost \$200 bil-

## Three Memcomputer Building Blocks

Modern electrical circuits use three components that respond to electrical inputs. Resistors impede electric current that flows through them, capacitors store electrical charges and inductors convert current into a magnetic field. (In a computer, transistors are typically used instead of resistors.) When power is cut off, the components return to their original states. Memcomputing versions of these components, however, retain their changed state, and this “memory” permits complex calculations to occur quickly. (The red lines in the symbols represent the different effects on electricity.)



### Memristor

This device changes resistance depending on the amount of current flowing through it, and it retains that change. Thus, it can both process information and hold on to it, like a memory.



### Memcapacitor

Not only does this device store electrical charges, but it changes its state, or capacitance, depending on the history of voltages applied to it. Again, that gives it memory and processing ability. Any unused energy can be recycled for other computing operations, saving power.



### Meminductor

Like a memristor, this device lets current flow through, but it also stores energy like a memcapacitor. The combination of flow and storage lets meminductors process information and store it as a computer memory component.

lion in electricity annually. This power hogging is not sustainable.

We cannot fix it by shrinking transistors—the fundamental element of digital electronics—to smaller and smaller sizes. The International Technology Roadmap for Semiconductors has forecasted that the transistor industry most likely will hit a technological wall by 2016 because available component materials cannot go down any further in size and maintain their capabilities.

Scientific research on some urgent problems will also hit a wall. Important questions that can only be tackled by heavy-duty computation, such as the prediction of global weather patterns or forecasting the occurrence of diseases in various populations by exploring large genome databases, will require larger and larger amounts of computational power. Memcomputers, by avoiding the expensive, power-hungry and time-consuming process of constantly transferring data between a CPU and memory, should save a significant amount of energy.

They are not, of course, the first information-processing devices to handle calculations and storage in one place. The human brain does this very thing, and memcomputing takes its inspira-

tion from this fast, efficient organ sitting on top of our shoulders.

The average human brain, according to many estimates, can perform about 10 million billion operations per second and uses only 10 to 25 watts to do so. A supercomputer would require more than 10 million times that power to do the same amount of work. And a computer does not even come close to performing such complicated tasks as pattern recognition—like separating the sound of a dog barking from a car passing in the street—that we do in the noisy and unpredictable environment we live in. Unlike our present supercomputers, calculations in the brain are not performed in two places but are done by the same neurons and synapses. Less shuffling means less energy spent and less time lost moving information around. A computer can do calculations, one at a time, faster than humans can, but it takes all that brute-force transistor power to carry them out.

Traditionally computers have relied on their separation of powers to keep programs and the data they use from interfering with one another during processing. Physical changes in a circuit caused by new data—say, the letters we typed in Word—would change and corrupt the program or the data. This could be avoided if circuit elements in a processor could “remember” the last thing they did, even after the electricity is turned off. Data would still be intact.

### THREE PARTS OF A NEW MACHINE

MEMCOMPUTING COMPONENTS can do exactly that: process information and store it after the electricity stops. One of these new devices is a memristor. To understand it, imagine a pipe that changes its diameter depending on the direction of water flow. When water is flowing right

to left, the pipe gets wider, enabling more water to flow through it. When water is flowing left to right, the pipe gets narrower, and less water goes through it. If the water is turned off, the pipe maintains its most recent diameter—it remembers the amount of water that flowed through it.

Now replace the water with electric current and replace the pipe with a memristor. It changes its state depending on the amount of current flowing, as the water pipe changes diameter—a wider pipe has less electrical resistance, and a narrower pipe has more. If you think of resistance as a number and the change in resistance as a process of calculation, a memristor is a circuit element that can process information and then hold it after the current is turned off. Memristors can combine the work of the processing unit and of the memory in one place.

The notion of memristors came from Leon O. Chua, an electrical engineer at the University of California, Berkeley, in the 1970s. At the time, his theory did not appear to be very practical. Materials used to make circuits did not retain memory of their last state like the imaginary water pipe, so the idea seemed far-fetched. But over the decades engineers and materials scientists

were able to exert more and more control over the circuit materials they fabricated, imbuing them with new properties. In 2008 Hewlett-Packard engineer Stanley Williams and his colleagues produced memory elements that could shift resistance and hold their shifted state. They shaped titanium dioxide into an electrical component just tens of nanometers (billionths of a meter) wide. In a paper in *Nature*, the scientists showed that the component retained a state that was determined by the history of current flowing through it. The imaginary pipe was real. (*Scientific American* is part of Nature Publishing Group.)

It turns out that these devices can be fabricated with a large variety of materials and can be made just a few nanometers across. Smaller dimensions mean that more of them can be packed into a given area, so they can be crammed into almost any kind of gadget. Many of these components can be made in the same semiconductor facilities we now employ to make computer components and therefore can be fabricated on an industrial scale.

Another key component that could be used in memcomputing is a memcapacitor. Regular capacitors are devices that store electrical charges, but they do not change their state, or capacitance, no matter how many charges are deposited in them. In today's computers they are mainly used in a particular kind of memory, called dynamic random-access memory (DRAM), which stores computer programs in a state of readiness so they can be uploaded quickly to the processor when it calls for them. A memcapacitor, however, not only stores charges but changes its capacitance depending on past voltages applied to it. That gives it both memory and processing ability. Further, because memcapacitors store charges—energy—that power could be recycled during computation, helping to minimize energy consumption by the overall machine. (Memristors, in contrast, use all the energy put into them.)

Some types of memcapacitors, made of relatively costly ferroelectric materials, are already available on the market and are used as devices for data storage. But research laboratories are developing versions made of inexpensive silicon, keeping the manufacturing price low enough to use them throughout a computer.

The meminductor is the third element of memcomputing. It has two terminals, and it stores energy like a memcapacitor while letting current flow through it like a memristor. Meminductors, too, exist right now. But they are quite large because they rely on big wire magnetic coils, so they would be hard to use in small computers. Advances in materials could change that in the near future, however, as it did for memristors just a few years ago.

In 2010 we started trying to show that memcomputing could handle calculations better than current computer architecture. One problem we focused on was finding a way out of a maze. Devising programs for maze running has long been a way to test the efficiency of computer hardware. Conventional algorithms for solving mazes explore the maze in small consecutive steps. For instance, one of the best-known algorithms is the so-called wall follower. The program traces the wall of the maze through all its twists and turns, avoiding empty spaces where the wall ends, and moves, calculation after painstaking calculation, from the entrance to the exit. This step-by-step approach is slow.

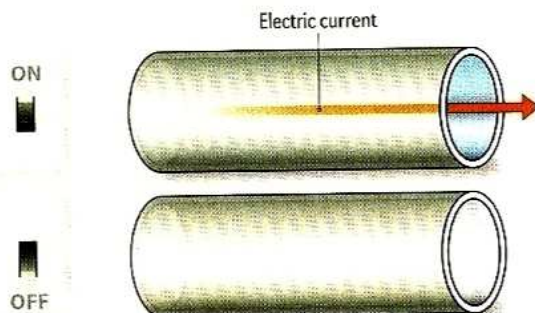
Memcomputing, we have shown in simulations, will solve the maze problem extremely fast. Consider a network of mem-

## The Power of Memory

One of the new memcomputing components, the memristor, responds in different ways to different levels of electric current. Three such changes are shown here. When the changes are retained, they form the basis for memory.

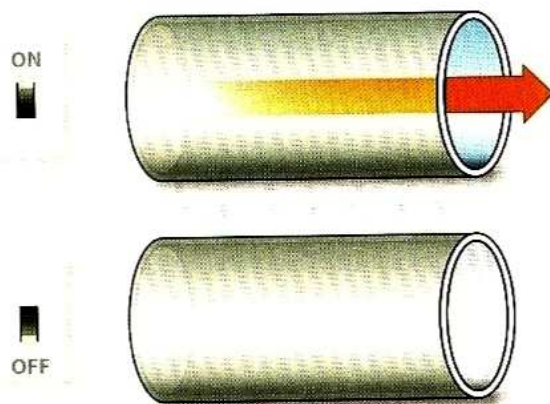
### Typical Resistance

When a weak electric current passes through a memristor, the device has a specific resistance level to the current. That level is shown here by the width of the tube. Resistance, in a computer, can be interpreted as a number, which leads to calculation. When the power is cut off, the tube stays the same width.



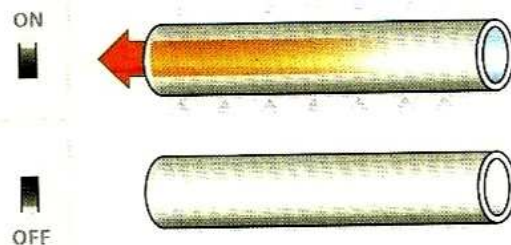
### Low Resistance

When a strong current passes through the memristor, the device lowers its resistance, shown by a wider tube. That change, interpreted as a different number than in the original configuration, permits information processing. When the power is off, the memristor retains this state, which allows it to function like memory. A traditional computer component would revert to its original state.



### High Resistance (Current Reversal)

Memristors can also increase resistance if the current through them is reversed, shown by a narrower tube. Again the tube does not revert to its original state when power is off, adding memory to its processing function.



ristors, one at each turn of the maze, all in a state of high resistance. If we apply a single voltage pulse across the entrance and exit points, the current will flow only along the solution path—it will be blocked by dead ends in other paths. As the current flows, it changes the resistances of the corresponding memristors. After the pulse disappears, the maze solution will be stored in the resistances of only those devices that have changed their state. We have computed *and* stored the solution in only one shot. All the memristors compute the solution in parallel, at the same time.

This kind of parallel processing is completely different from current versions of parallel computing. In a typical parallel machine today, a large number of processors compute different parts of a program and then communicate with one another to come up with the final answer. This still requires a lot of energy and time to transfer information between all these processors and their associated—but physically distinct—memory units. In our memcomputing scheme, it simply is not necessary.

Memcomputing really shows advantages when applied to one of the most difficult types of problems we know of in computer science: calculating all the properties of a large series of integers. This is the kind of challenge a computer faces when trying to decipher complex codes. For instance, give the computer 100 integers and then ask it to find at least one subset that adds up to zero. The computer would have to check all possible subsets and then sum all numbers in each subset. It would plow through each possible combination, one by one, which is an exponentially huge increase in processing time. If checking 10 integers took one second, 100 integers would take  $10^{27}$  seconds—millions of trillions of years.

As with the maze problem, a memcomputer can calculate all subsets and sums in just one step, in true parallel fashion, because it does not have to shuttle them back and forth to a processor (or several processors) in a series of sequential steps. The single-step approach would take just a single second.

Despite these advantages and despite the fact that components have already been made in labs, memcomputing chips are not yet commercially available. At the moment, early versions are being tested in academic facilities and by a few manufacturers to see if these untried designs are robust enough, over repeated use, to replace current memory chips made of standard transistors and capacitors. These chips are the kind you find in USB drives and solid-state memory drives. The tests can take a long time because the components need to last years without failure.

We think some memcomputing designs could be ready for use in the very near future. For instance, in 2013, together with two researchers at the Polytechnic University of Turin in Italy, Fabio Lorenzo Traversa and Fabrizio Bonani, we suggested a concept called dynamic computing random-access memory (DCRAM). The goal is to replace the standard type of memory that, as we have discussed, is used to hold programs and data just before a processor calls for them. In this conventional memory, each bit of information that makes up the program is represented by a charge stored on a single capacitor. That calls for a large number of capacitors to represent one program.

If we replace them with memcapacitors, however, all the different logic operations required by the program can be represented by a much smaller number of memcapacitors in this memory area. Memcapacitors can shift from one logic opera-

tion to another almost instantly when we apply different voltages to them. Computing instructions such as “do  $x$  AND  $y$ ,” “do  $x$  OR  $y$ ” and “ELSE do  $z$ ” can be handled by two memcapacitors instead of a large number of fixed regular capacitors and transistors. We do not have to change the basic physical architecture to carry out different functions. In computer terminology, this is called polymorphism, the ability of one element to perform different operations depending on the type of input signal. Our brain possesses this type of polymorphism—we do not need to change its architecture to carry out different tasks—but our current machines do not have it, because the circuits in their processors are fixed. And with memcomputing, of course, because this computation is occurring within a memory area, the time- and power-consuming shuffle back and forth to a separate processor is eliminated, and the result of the program’s calculations can be stored in the same place.

These systems can be built with present fabrication facilities. They do not require a major leap in technology. What may hold them up is the need to design new software to control them. We do not yet know the most efficient kinds of operating systems to command these new machines. The machines have to be built, and then various controlling systems have to be tested and optimized. This is the same design process that computer scientists went through with our present crop of machines.

Scientists also would like to find the best way to integrate these new memelements into our current computers. It might be a good idea to keep present processors to handle simple tasks—like computing that sentence in Word that began this article—while using memcomputing elements in the same machine for more intricate and hitherto time-consuming operations. We will need to build, test, rebuild and retest.

It is enticing, though, to consider where this technology could lead us. After building and testing, computer users might have a small device, maybe small enough to hold in your hand, that could tackle very complex problems involving, say, pattern recognition or modeling the earth’s climate at a very fine scale. Something it could do in one or a few computational steps, at very low energy and cost.

Wouldn’t you stand in line to get one? ■

---

#### MORE TO EXPLORE

**Missing Link of Electronics Discovered: Memristor.** JR Minkel; *ScientificAmerican.com*, May 1, 2008.

**Putting Memory into Circuit Elements: Memristors, Memcapacitors, and Meminductors.** M. Di Ventra, Y. V. Pershin and L. O. Chua in *Proceedings of the IEEE*, Vol. 97, No. 8, pages 1371–1372; August 2009.

**Experimental Demonstration of Associative Memory with Memristive Neural Networks.** Yuriy V. Pershin and Massimiliano Di Ventra in *Neural Networks*, Vol. 23, No. 7, pages 881–886; September 2010.

**Solving Mazes with Memristors: A Massively Parallel Approach.** Yuriy V. Pershin and Massimiliano Di Ventra in *Physical Reviews E*, Vol. 84, No. 4, Article No. 046703; October 14, 2011.

**The Parallel Approach.** Massimiliano Di Ventra and Yuriy V. Pershin in *Nature Physics*, Vol. 9, pages 200–202; April 2013.

**Dynamic Computing Random Access Memory.** F. L. Traversa et al. in *Nanotechnology*, Vol. 25, No. 28, Article No. 285201; July 2014.

#### FROM OUR ARCHIVES

**The Next 20 Years of Microchips.** The Editors; January 2010.

[scientificamerican.com/magazine/sa](http://scientificamerican.com/magazine/sa)