

THE MYTH OF MIND CONTROL

Before we can build cyborg soldiers and robohumans, researchers face a tough challenge: cracking the neural code. By **John Horgan**

All it took was a few jolts of electricity to turn ordinary rats into roborats and for pundits to leap to the conclusion that ordinary humans would soon be transformed into robohumans. Scientists at the State University of New York Downstate Medical Center in Brooklyn sparked a media frenzy six years ago when they demonstrated that rats with electrodes implanted in their brains could be steered like remote-controlled toy cars through an obstacle course. Using a laptop computer

equipped with a wireless transmitter, a researcher stimulated cortical cells governing whisker sensations and reinforced those signals by zapping the rats' pleasure centers. Presto! With this simple setup, the team had created living robots.

Publications around the world proclaimed the imminence of those familiar science fiction staples: surgically implanted devices that electronically monitor and manipulate our minds. *The Economist* warned that neurotechnology could be on the verge of "overturning the essential nature of humanity," and *New York Times* columnist William Safire brooded that neural implants might allow a "controlling organization" to hack into our brains. In a more positive vein, Rodney Brooks, an artificial-intelligence maven at MIT, predicted in *Technology Review* that by 2020 implants would let us carry out "thought-activated Google searches."

Hollywood's remake of *The Manchurian Candidate* raised the specter of a remote-controlled soldier turned politician. In fact, officials at the Defense Advanced Research Projects Agency (DARPA), which funded the roborat team, have suggested that cyborg soldiers could control weapons systems—or be controlled—via brain chips. "Implanting electrodes into healthy people is not something we're going to do anytime soon," says Alan Rudolph, the former head of the DARPA brain-machine research program. "But 25 years ago, no one thought we'd put a laser in the eye. This agency leaves the door open to what's possible."

Of course, that fails to answer the question: Just how realistic are these futuristic scenarios? To achieve truly precise mind reading and control, neuroscientists must master the syntax, or set of rules, that transforms electrochemical pulses coursing through the brain into perceptions, memories, emotions, and decisions. Deciphering this so-called neural code—think of it as the brain's software—is the ultimate goal of many scientists tinkering with brain-machine interfaces. "If you're a real neuroscientist, that's the game you want to play," says John Chapin, a physiology and pharmacology professor at Downstate Medical Center and a member of the roborat research team.

**IN ADDITION
TO BEING THE
MOST SIGNIFICANT
MYSTERY IN
SCIENCE, THE
NEURAL CODE MAY
ALSO BE THE
HARDEST TO
SOLVE.**

Chapin ranks the neural code right up there with two other great scientific mysteries: the origin of the universe and of life on earth. The neural code is arguably the most consequential of the three. The solution could, in principle, vastly expand our power to treat ailing brains and to augment healthy ones. It could allow us to program computers with human capabilities, helping them become more clever than HAL in *2001: A Space Odyssey* and C-3PO in *Star Wars*. The neural code could also represent the key to the deepest of all philosophical conundrums—the mind-body problem. We would finally understand how this wrinkled lump of jelly in our skulls generates a unique, conscious self with a sense of personal identity and autonomy.

In addition to being the most significant mystery in science, the neural code may also be the hardest to solve. Despite all they have learned in the past century, neuroscientists have made little headway in figuring out exactly how brain cells process information. "It's a bit like saying after 100 years of researching the body, 'Do you know if testes produce urine or sperm?'" says neuroscientist V. S. Ramachandran of the University of California at San Diego. "Our notions are still very primitive."

The neural code is often likened to the machine code that underpins the operating system of a digital computer. Like transistors, neurons serve as switches, or logic gates, absorbing and emitting electrochemical pulses. These pulses, called action potentials, resemble the basic units of information in digital computers. But the brain's complexity dwarfs that of any existing computer. A typical brain contains 100 billion cells—almost as numerous as the stars in the Milky Way galaxy. And each cell is linked via synapses to as many as 100,000 others. The synapses are awash in hormones and neurotransmitters that modulate the transmission of signals, and they constantly form and dissolve, weaken and strengthen in response to new experiences.

Assuming that the brain contains one quadrillion synapses processing on average 10 action potentials per second and that these transactions represent the brain's computational output, then the brain performs at 10 quadrillion operations per second. In comparison, IBM's Blue Gene/P supercomputer, introduced a year ago, can be configured to process up to 3 quadrillion operations per second. Some have argued that within a decade or so, computers may surpass the computational power of the brain. But many more computations may occur at a scale below or above that of individual synapses and neurons via genetic, hormonal, and other processes, says Steven Rose, a neurobiologist at the Open University in England. If so, the brain's total computational power could be many orders of magnitude greater than some believe. "The brain may use every possible means of carrying information," Rose says.

Optimists recall that in the middle of the 20th century, some biologists feared that the genetic code was too complex to crack. Then, in 1953, Francis Crick and James Watson unraveled the structure of DNA, and researchers quickly established that the double helix mediates an astonishingly simple genetic code governing the heredity of all organisms. The neural code is not likely to reveal such an elegant, universal solution. The brain is "so adaptive, so dynamic, changing so frequently from instant to instant," says Miguel Nicolelis, a neural-prosthesis researcher at Duke University, that "it may not be proper to use the term *code*."

Nicolelis has faith that science will one day ferret out all the brain's information-processing tricks—or at least enough of them to yield huge improvements in neural prostheses for people who are paralyzed, blind, or otherwise disabled. Yet he believes that certain aspects of our

minds may remain inviolable because our most meaningful thoughts and memories are written in a code, or language, that is unique to each of us. "There will always be some mystery," Nicoletis says.

If so, the bad news is that brain chips will never be sophisticated enough for us to learn new languages instantly or have a "mental telephone" conversation with a friend "simply by thinking about talking," as *Popular Science* magazine has prophesied. The good news is that we are not on the verge of what the *Boston Globe* has called a "Matrix-like cyberpunk dystopia" in which we all become robohumans, controlled by implants that "impose false memories" and "scan for wayward thoughts."

All the loose speculation provoked by roborats is ironic considering that the experiment was just a small-scale replay of a major media event that occurred almost 45 years ago. In 1964, José Delgado, a neuroscientist from Yale University, stood in a Spanish bullring as a bull with a radio-equipped array of electrodes implanted in its brain charged toward him. When Delgado pushed a button on a radio transmitter he was holding, the bull stopped in its tracks. Delgado pushed another button and the bull obediently turned to the right and trotted away. *The New York Times* hailed the event as "probably the most spectacular demonstration ever performed of the deliberate modification of animal behavior through external control of the brain."

Delgado implanted similar electrode arrays, or "stimoceivers," in the brains of cats, monkeys, chimpanzees, and even human psychiatric patients. He showed that he could jerk the limbs of patients like marionettes and induce sensations such as euphoria, sexual arousal, sleepiness, terror, and rage. In his 1969 book, *Physical Control of the Mind: Toward a Psychocivilized Society*, Delgado extolled the promise of brain stimulation techniques for curbing violent aggression and other maladaptive traits.

Delgado's work, partly funded by the Pentagon, provoked fear of government plots to transform citizens into robots. He dismissed this "Orwellian possibility," pointing out that the technology was still much too unreliable and crude for precise mind control. The major impediment to progress, he wrote, is that "our present knowledge regarding the coding of information...is so elemental."

In Delgado's heyday neuroscientists believed that the brain employed just a single, simple coding scheme discovered in the 1930s by Lord Edgar Adrian, a British neurobiologist. After isolating sensory neurons from frogs and eels, Adrian showed that as the intensity of a sensory stimulus increases, so does a neuron's firing rate, which can peak at 200 spikes per second. In the next few decades, experiments confirmed that the nervous systems of all animals employ this method of conveying information, called a rate code. Researchers also demonstrated that specific neurons are dedicated to extremely specific tasks, such as seeing vertical lines, hearing sounds of a specific pitch, or flexing a finger. Together, these findings suggested that controlling the brain might be a simple matter of delivering the right jolt of electricity to the right clusters of brain cells.

It turns out that things are not so simple. Research has undermined two basic assumptions about how the brain processes information. One is the view of neurons as drones single-mindedly carrying out specific tasks. In fact, cells can be retrained for different jobs, switching from facial expressions to finger flexing or from seeing red to hearing squeaks. Our neural circuits keep shifting "massively and continuously," not only during childhood but throughout our lives, says Michael Merzenich of the University of California at

San Francisco, whose research has helped expose just how plastic neurons really are.

Neuroscientists are rethinking another long-held dogma: that the firing rate serves as a brain cell's sole means of expression. Rate codes are extremely inefficient. They are analogous to a language that conveys information only through modulations in a voice's volume, and they imply that the brain is inherently noisy and wasteful. In the traditional view, what counts as a genuine signal is a surge in the firing rate of a cell from, say, 2 to 50 times a second; variations in the intervals between successive spikes are considered irrelevant. But just as geneticists suspect that the "junk DNA" riddling our genomes actually serves hidden functions, some neuroscientists believe that information may lurk within the fluctuating gaps between spikes. Schemes of this sort, which are known as temporal codes, imply that significant information may be conveyed within the span of just a couple of spikes.

Another time-sensitive code involves groups of neurons firing in precise lockstep, or synchrony. Some evidence suggests that synchrony helps us focus our attention. If you are at a noisy cocktail party and suddenly hear someone nearby talking about you, your ability to eavesdrop on that conversation and ignore all the others around you could result from the synchronous firing of cells. "Synchrony is an effective way to boost the power of a signal and the impact it has downstream on other neurons," says Terry Sejnowski, a computational neurobiologist at the Salk Institute. He speculates that the abundant feedback loops linking neurons allow them to synchronize their firing before passing messages on for further processing.

Wolf Singer of the Max Planck Institute for Brain Research, in Frankfurt, supports a code involving synchronous oscillations, or many neurons firing at both the same rate and the same time. He thinks this may play an important role in cognition.

Then there is the chaotic code championed by Walter J. Freeman of the University of California at Berkeley. For decades he has contended that far too much emphasis has been placed on individual neurons and action potentials, for reasons that are less

**THE UNIQUENESS
OF EACH INDIVIDUAL
REPRESENTS A
BARRIER TO
SCIENCE'S ATTEMPTS
TO UNDERSTAND
AND CONTROL
THE MIND.**

empirical than pedagogical. The action potential “organizes data, it is easy to teach, and the data are so compelling in terms of the immediacy of spikes on a screen.” But spikes are ultimately just “errand boys,” Freeman says; they serve to convey raw sensory information, but then much more subtle, larger-scale processes immediately take over.

The most vital components of cognition, Freeman believes, are the electrical and magnetic fields, generated by synaptic currents, that constantly ripple through the brain. These fields are chaotic, in the sense that they conceal a complex order and are extremely sensitive to minute influences—the so-called butterfly effect. A sound enters the ear and triggers a stream of action potentials, which nudge the waves of electrical activity coursing through the cortex into a particular chaotic pattern, or attractor. The result is fantastically precise, almost instant comprehension. “You pick up the telephone and hear a voice,” Freeman says, “and before you even know the meaning of the words, you know who you’re talking to and what her emotional state is.”

None of these alternatives to rate codes has been proven yet. In fact, so little is known about how the brain processes information that “it’s difficult to rule out any coding scheme at this time,” says neuroscientist Christof Koch of Caltech. Koch and Itzhak Fried, who is both a neuroscientist and a practicing neurosurgeon at UCLA Medical School, have uncovered evidence for a coding scheme long ago discarded as implausible. This scheme has been disparaged as the “grandmother cell” hypothesis because in its reductio ad absurdum version, it implies that our memory banks dedicate a single neuron to each person, place, or thing that inhabits our thoughts, such as Grandma. Most theorists assume that such a complex concept would have to be supported by large populations of cells, each of which corresponds to one component of the object (the bun, the bifocals, the leather miniskirt).

Yet in 2004 and 2005, Fried and Koch found neurons that act very much like grandmother cells. Their subjects were epileptics who had electrodes temporarily inserted into their brains to provide information that could guide surgical treatment. The researchers monitored the output of the electrodes while showing the patients images of animals, people, and other things. A neuron in the amygdala of one patient spiked only in response to three quite different images of Bill Clinton—a line drawing, a presidential portrait, and a group photograph. A cortical cell in another patient responded in a similar way to images of characters from *The Simpsons*.

Most recently, Fried studied the activity of single neurons in the hippocampus and surrounding areas while epileptics first watched television episodes and later recalled these episodes. The same hippocampal neurons fired whether the patients were actually watching the show or merely recalling it.

It makes intuitive sense, Koch says, that our brains should dedicate some cells to people and things frequently in our thoughts. He adds that his findings might seem less surprising if one realizes that neurons are much more than simple threshold switches that fire whenever incoming pulses from other neurons exceed a certain level. A typical neuron receives input from thousands of other cells, some of which inhibit rather than encourage the neuron’s firing. The neuron may in turn encourage or suppress firing by some of those same cells in complex positive or negative feedback loops.

In other words, a single neuron may resemble less a simple switch than a customized minicomputer, sophisticated enough to distinguish your grandmother from Grandma Moses. If this view

WITH THOUGHTS ALONE, MONKEYS WERE ABLE TO MANEUVER A MECHANICAL ARM TO GRAB FOR FOOD IN FRONT OF THEM.

is correct, meaningful messages might be conveyed not just by hordes of neurons screaming in unison but by a small group of cells whispering, perhaps in a terse temporal code. Discerning such faint signals within the cacophony of the brain will be “incredibly difficult,” Koch says, no matter how far neurotechnology advances.

Efforts to detect the whispers amid the cacophony are further complicated by the improvisational dexterity of the brain. Studies of the motor cortex, which underlies body movement, have shown that the brain invents entirely new coding schemes for novel situations. In the 1980s researchers discovered neurons in a monkey’s motor cortex that peaked in their firing rate when the monkey moved its hand in a specific direction. Rather than falling silent when the hand diverged even slightly from its so-called preferred direction, the cells’ firing rate diminished in proportion to the angle of divergence.

Several teams, including one led by Andrew Schwartz of the University of Pittsburgh, have sought to exploit these findings to create neural prostheses for paralyzed patients. These teams have demonstrated that electrodes implanted in a monkey’s motor cortex can detect signals accompanying a specific arm movement; these same signals, after being processed by a computer, can be used to manipulate a robot arm that might be in another room—or even, in one experiment, a robot’s legs on another continent. If the monkey’s arm is tied down, the monkey learns to control the robot arm through pure thought—but with an entirely different set of neural signals. In a June 2008 study published in *Nature*, Schwartz and his team used monkeys’ cortical signals to control a multijointed prosthetic device as it interacted with the physical environment. With thoughts alone, the monkeys were able to maneuver the mechanical arm to reach for and grab food located in front of them. The food reached their mouths about two-thirds of the time.

These findings dovetail with others showing that neurons’ coding behavior shifts in different contexts. “What you’re aiming at is sort of a moving target,” Schwartz explains. “If you make an estimate of something at one point in time, that doesn’t mean it’s going to stay that way.”

The mutability of the neural code is not necessarily bad news for neural-prosthesis designers. In fact, the brain's capacity for inventing new information-processing schemes is thought to explain the success of artificial cochleas, which have been implanted in the ears of approximately 100,000 hearing-impaired people around the world in the past few decades. Commercial versions typically employ an array of electrodes, each of which channels electrical signals corresponding to a different pitch toward the auditory nerve. Like an old telephone party line, the electrodes can stimulate not just a single neuron but many simultaneously.

When cochlear implants were introduced in the mid-1980s, many neuroscientists expected them to work poorly, given their crude design. But the devices work well enough for some deaf people to converse over the telephone, particularly after an adjustment period during which channel settings are tweaked to provide the best reception. Patients' brains somehow figure out how to make the most out of the strange signals.

There are surely limits to the brain's ability to make up for scientists' ignorance, as the poor performance of other neural prostheses suggests. Artificial retinas, light-sensitive chips that mimic the eye's signal-processing ability and stimulate the optic nerve or visual cortex, have been tested in a handful of blind subjects who usually "see" nothing more than phosphenes, or flashes of light. And like Schwartz's monkeys, a few paralyzed humans have learned to transmit commands to computers via chips embedded in their brains, but the associated prostheses are still slow and unreliable.

Nevertheless, the surprising effectiveness of artificial cochleas—together with other evidence of the brain's adaptability and opportunism—has fueled optimism about the prospects for brain/machine interfaces. "This is very relevant to why we think we are going to be successful," says Ted Berger of the University of Southern California in Los Angeles, who is leading a project to create implantable brain chips that can restore or enhance memory. "We don't need a perfectly accurate model of a memory cell," he says. "We probably just have to be close, and the rest of the brain will adapt around it."

Berger's experiments use slices of rat brain in petri dishes. For more than a decade, he has embedded electrodes in slices of the hippocampus—which plays a role in learning and memory—and recorded neurons' responses to a wide range of electrical stimuli. His observations have made him a firm believer in temporal codes; hippocampal cells seem to be exquisitely sensitive not only to the rate but also to the timing of incoming pulses. "The evidence for temporal coding is indisputable," Berger says.

He and his team have created a prototype of the world's first memory implant chip. In order to create the chip, Berger bombarded live rat hippocampal neurons with electric impulses and recorded the electrical responses from the cells, collecting the "vocabulary" of the tissue. This information was programmed into the chip, enabling it to "listen" to brain signals, decode them, and respond appropriately with its own chain of electrical pulses, just like a network of neurons would. Experimentally, the chip responds to neural signals in exactly the same way as does the living brain tissue, suggesting that the chip may have the ability to communicate with living brain cells.

Berger boldly predicts that someday chips like his might restore memory capacity to stroke victims. But in some respects Berger is quite modest. He acknowledges that his memory chip could not be

used to identify and manipulate specific memories. Rather, it can simulate "how neurons in a particular part of the brain change inputs into outputs. That's very different from saying that I can identify a memory of your grandmother in a particular series of impulses." To achieve this sort of mind reading, scientists would have to compile a "dictionary" for translating specific neural patterns into specific memories, perceptions, and thoughts. "I don't know that it's not possible," Berger says. "It's certainly not possible with what we know at the moment."

"Don't count on it in the 21st century, or even in the 22nd," says neuroscientist Bruce McNaughton of the University of Arizona. With arrays of as many as 50 electrodes, McNaughton has monitored neurons in the hippocampus of rats as they move through a maze. Once a rat learns to navigate a maze, its neurons discharge in the same patterns whenever it goes through it. Remarkably, when the rat sleeps after a hard day of maze running, the same firing pattern often unfolds; the rat is presumably dreaming of the maze. This pattern could be said to represent—at least partially—the rat's memory of the maze.

McNaughton emphasizes that the same maze generates different firing patterns in different rats; even in the same rat, the pattern changes if the maze is moved to a different room. He thus doubts whether science can compile a dictionary for decoding the neural signals corresponding to human memories, which are surely more complex, variable, and context sensitive than those of rats. At best, McNaughton suggests, one might construct a dictionary for a single person by monitoring the output of all her neurons for years while recording all her behavior and her self-described thoughts. Even then, the dictionary would be imperfect, and it would have to be constantly revised to account for the individual's ongoing experiences. This dictionary would not work for anyone else.

Delgado hinted at the problem more than 30 years ago in *Physical Control of the Mind* when he raised the knotty question of meaning. With improved stimoceivers and a better understanding of the neural code, he said, scientists might determine what we are perceiving—a piece of music, say—based on our neural output. But no conceivable technology will be subtle enough to discern all the memories, emotions, and meanings aroused in us by our perceptions, because these emerge from "the experiential history of each individual." You hear a stale pop tune, I hear my wedding song.

This is one point on which many neuroscientists agree: The uniqueness of each individual represents a fundamental barrier to science's attempts to understand and control the mind. Although all humans share a "universal mode of operation," Freeman says, even identical twins have divergent life histories and hence unique memories, perceptions, and predilections. The patterns of neural activity underpinning our selves keep changing throughout our lives as we learn to play checkers, read *Thus Spake Zarathustra*, fall in love, lose a job, win the lottery, get divorced, take Prozac.

Freeman thinks the prospects are good for developing relatively simple neural prostheses, such as devices that improve vision in the blind. But he suspects that our brains' complexity and diversity rule out more ambitious projects, such as mind reading. If artificial-intelligence engineers ever succeed in building a truly intelligent machine based on a neural coding scheme similar to ours, "we won't be able to read its mind either," Freeman says. We and even our cyborg descendants will always be "beyond Big Brother, and I'm very grateful for that."

Additional reporting by Shannon Sweeney