The Memory

Researchers are closing in on the rules that the brain uses to lay down memories. Discovery of this memory code could lead to the design of smarter computers and robots and even to new ways to peer into the human mind

By Joe Z. Tsien

KEY CONCEPTS

- The brain relies on large populations of neurons acting in concert to represent and form a memory of an organism's experiences.
- In the mouse hippocampus (an area critical to memory formation), subsets of such populations—dubbed "neural cliques"—have been shown to respond to different aspects of an event. Some represent abstract, general information about a situation; others indicate more selective features.
- The same hierarchical organization used to lay down memories could be applied by the brain to convert collections of electrical impulses into perception, knowledge and behavior. If so, the memory work brings investigators closer to uncovering the universal neural code: the rules the brain uses to identify and make sense of the body's experiences.
- The author and his colleagues have converted recordings of clique activity into binary code. Such digitization of brain signals could create a foundation for assembling a codebook of the mind—a tool for cataloguing thoughts and experiences and comparing them across individuals and, perhaps, species. *—The Editors*

Anyone who has ever been in an earthquake has vivid memories of it: the ground shakes, trembles, buckles and heaves; the air fills with sounds of rumbling, cracking and shattering glass; cabinets fly open; books, dishes and knickknacks tumble from shelves. We remember such episodes—with striking clarity and for years afterward—because that is what our brains evolved to do: extract information from salient events and use that knowledge to guide our responses to similar situations in the future. This ability to learn from past experience allows all animals to adapt to a world that is complex and ever changing.

For decades, neuroscientists have attempted to unravel how the brain makes memories. Now, by combining a set of novel experiments with powerful mathematical analyses and an ability to record simultaneously the activity of more than 200 neurons in awake mice, my colleagues and I have discovered what we believe is the basic mechanism the brain uses to draw vital information from experiences and turn that information into memories. Our results add to a growing body of work indicating that a linear flow of signals from one neuron to another is not enough to explain how the brain represents perceptions and memories [see "Seeking the Neural Code," by Miguel A. L. Nicolelis and Sidarta Ribeiro; Scientific American, December 2006]. Rather the coordinated activity of large populations of neurons is needed.

Furthermore, our studies indicate that neuronal populations involved in encoding memories also extract the kind of generalized concepts that allow us to transform our daily experiences into knowledge and ideas. Our findings bring biologists closer to deciphering the universal neural code: the rules the brain follows to convert collections of electrical impulses into perception, memory, knowledge and, ultimately, behavior. Such understanding could allow investi-

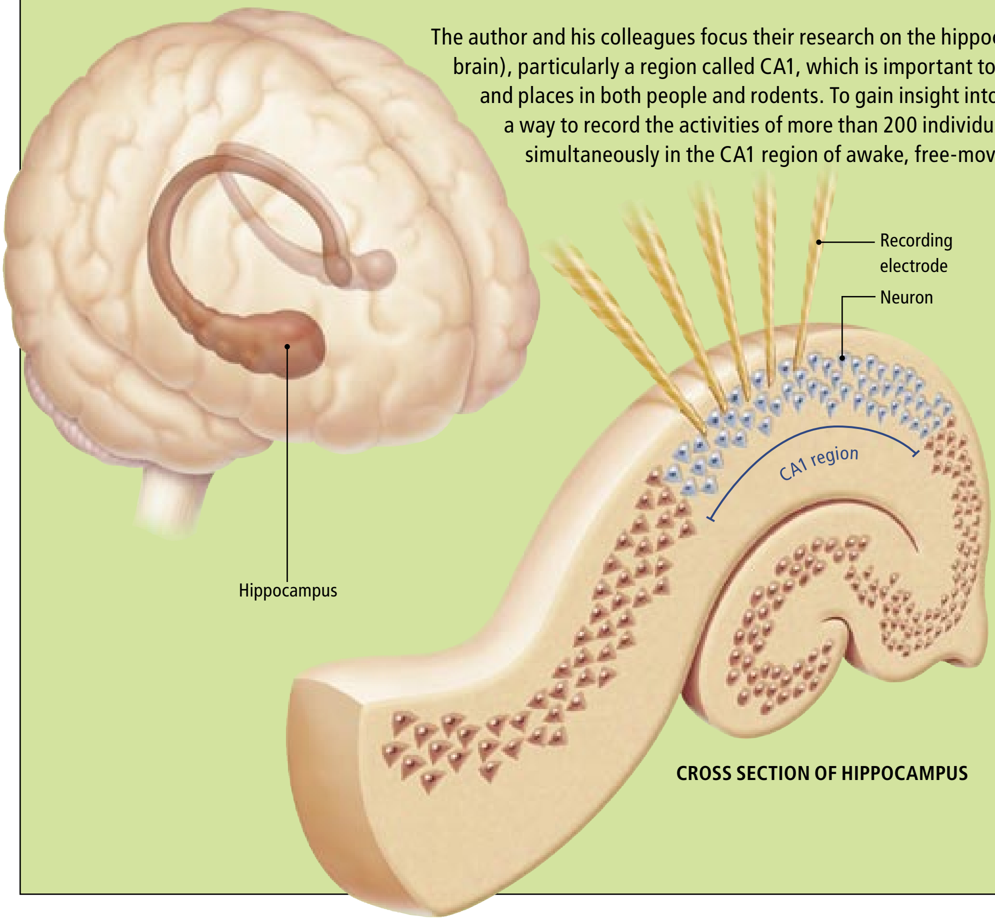
TOM DRAPER DESIGN; IMAGESHOP/CORBIS (*photograph*)

Code

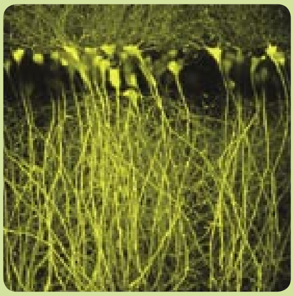
[BASICS]

A SEAT OF MEMORY

The author and his colleagues focus their research on the hippocampus (shown in human brain), particularly a region called CA1, which is important to forming memories of events and places in both people and rodents. To gain insight into the process, they developed a way to record the activities of more than 200 individual nerve cells, or neurons, simultaneously in the CA1 region of awake, free-moving mice.



▼ CA1 neurons branch dramatically, a feature well captured in this micrograph depicting mouse nerve cells coaxed to produce a yellow fluorescent protein.



ALICE Y. CHEN (illustrations); M. FUHRMANN AND JOCHEM HERMS Center for Neuropathology and Prion Research, Ludwig Maximilians University Munich (micrograph)

gators to develop more seamless brain-machine interfaces, design a whole new generation of smart computers and robots, and perhaps even assemble a codebook of the mind that would make it possible to decipher—by monitoring neural activity—what someone remembers and thinks.

We designed experiments that take advantage of what the brain seems to do best: laying down memories of dramatic events.

Doogie Raises Questions

My group's research into the brain code grew out of work focused on the molecular basis of learning and memory. In the fall of 1999 we generated a strain of mice engineered to have improved memory [see "Building a Brainier Mouse," by Joe Z. Tsien; SCIENTIFIC AMERICAN, April 2000]. This "smart" mouse—nicknamed Doogie after the brainy young doctor in the early-1990s TV dramedy *Doogie Howser, M.D.*—learns faster and remembers things longer than wild-type mice. The work generated great interest and debate and even made the cover of *Time* magazine. But our findings left me asking, What exactly is a memory?

Scientists knew that converting perceptual experiences into long-lasting memories requires a brain region called the hippocampus. And we even knew what molecules are critical to the process, such as the NMDA receptor, which we altered to produce Doogie. But no one knew how, exactly, the activation of nerve cells in the brain represents memory. A few years ago I began to wonder if we could find a way to describe mathematically or physiologically what memory is. Could we identify the relevant neural network dynamic and visualize the activity pattern that occurs when a memory is formed? And could we discern the organizing principles that enable neuronal populations to extract and record the most vital details of an experience?

To learn something about the neural code involved in memory, we first needed to design better brain-monitoring equipment. We wanted to continue working with mice, in part so that we could eventually conduct experiments in animals with genetically altered abilities to learn and remember, such as the smart mouse Doogie and mutant mice with impaired memory. Researchers had monitored the activities of hundreds of neurons in awake monkeys, but investigators working with mice had managed at best to record from only 20 or 30 cells at once—mostly because the mouse brain is not much bigger than a peanut. So Longnian Lin, then a postdoctoral fellow in my lab, and I developed a recording device that allowed us to monitor the activities of much larger numbers of individual neurons in the awake, freely behaving mouse.

We then designed experiments that take advantage of what the brain seems to do best: laying down memories of dramatic events that can have profound influences on one's life. Witnessing the 9/11 terrorist attacks, surviving an earthquake or even plummeting 13 stories in Disney's Tower of Terror are things that are hard to forget. So we developed tests that would mimic this type of emotionally charged, episodic event. Such experiences should produce memories that are long-lasting and strong. And encoding such robust memories, we reasoned, might involve a large number of cells in the hippocampus, thus making it more likely that we would be able to find cells activated by the experience and gather enough data to unravel any patterns and organizing principles involved in the process.

The episodic events we chose include a lab version of an earthquake (induced by shaking a small container holding a mouse), a sudden blast of air to the animal's back (meant to mimic an owl attack from the sky) and a brief vertical free fall inside a small "elevator" (which, when we first started doing these experiments, was provided by a cookie jar we had in the lab). Each animal was subjected to seven episodes of each event separated by periods of rest over several hours. During the events—and the intervening rest periods—we recorded activity from as many as 260 cells in the CA1 region of the hippocampus, an area that is key to memory formation in both animals and humans [*see box on next two pages*].

Startling Patterns

After collecting the data, we first attempted to tease out any patterns that might encode memories of these startling events. Remus Osan—another postdoctoral fellow—and I analyzed the recordings using powerful pattern-recognition methods, especially multiple discriminant analysis, or MDA. This mathematical method collapses what would otherwise be a problem with a large number of dimensions (for instance, the activities of 260 neurons before and after an event, which would make 520 dimensions) into a graphical space with only three dimensions. Sadly for classically trained biologists, the axes no longer correspond to any tangible measure of neuronal activity, but they do map out a mathematical subspace capable of discriminating distinct patterns generated by different events.

When we projected the collected responses of all recorded neurons from an individual animal into this three-dimensional space, four distinct "bubbles" of network activity popped out: one associated with the resting brain state, one with the earthquake, one with the air puff and one with the elevator drop. Thus, each of our startling episodes resulted in a distinct pattern of activity in the CA1 neural ensembles. The patterns, we believe, represent integrated information about perceptual, emotional and factual aspects of the events.

To see how these patterns evolved dynamically as the animals endured their various experiences, we then applied a "sliding window" technique to hours of recorded data for each animal—moving through the recordings moment by moment and repeating the MDA analysis for each half-second window. As a result, we were able to visualize how the response patterns changed as the animal laid down memories of each event while it happened. In an animal that went through an earthquake, for example, we could watch the ensemble activity begin in the rest bubble, shoot out into the earthquake bubble and then return to the resting state, forming a trajectory with a characteristic triangular shape.

This temporal analysis revealed something even more interesting: the activity patterns associated with those startling experiences recurred spontaneously at intervals ranging from seconds to minutes after the actual event. These "replays" showed similar trajectories, including the characteristic geometric shape, but had smaller amplitudes than their original responses. The recurrence of these activation patterns provides evidence that the information traveling through the hippocampal system was inscribed into the brain's memory circuits—and we imagine the replay corresponds to a recollection of the experience after the fact. This ability to qualitatively and quantitatively measure spontaneous reactivations of memory-encoding patterns opens a door to being able to monitor how newly formed memory traces are consolidated into long-lasting memories and to exam-

In an animal that went through an earthquake, we could watch the neural ensemble activity begin in the rest bubble, shoot out into the earthquake bubble and then return to the resting state.

[THE AUTHOR]

COURTESY OF JOE Z. TSIEN

Joe Z. Tsien, professor of pharmacology and biomedical engineering and director of the Center for Systems Neurobiology at Boston University, has made major contributions to the understanding of learning and memory and is a pioneer in the development of techniques for knocking out genes or proteins at a specific time in a specific tissue. He made headlines in 1999, when, at Princeton University, he generated a smart mouse strain called Doogie, which learned faster and remembered things longer than standard laboratory mice. Tsien, who moved to Boston University in 2004, recently founded the Shanghai Institute of Brain Functional Genomics at East China Normal University, his alma mater.

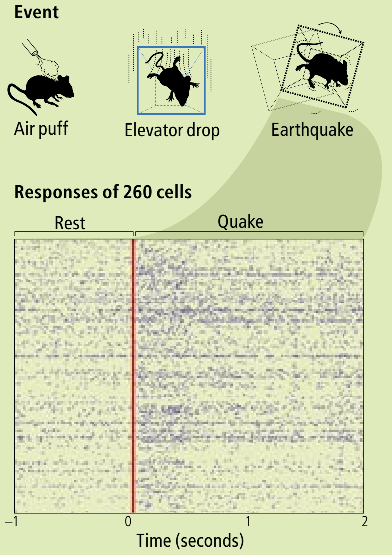
[FINDINGS]

FIRST STEPS TO UNCOVERING THE MEMORY CODE

To gain insight into the code that the brain uses to lay down memories, the author and his co-workers analyzed brain signals in a series of innovative ways.

❶ RECORDED EXPERIENCES

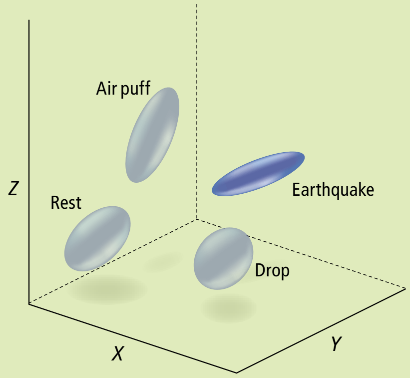
The team exposed mice to three startling experiences—a puff of air on the back, a fall in a container (the "elevator" drop), and shaking in a cage (the "earthquake")—while a recorder plotted firing from a large set of CA1 neurons. Each row in the plot below (from the quake) captures firing of a single cell over time.



❷ APPLIED PATTERN-RECOGNITION ALGORITHMS

Software translated the data from an individual mouse into a 3-D plot that represented the activity of the full ensemble of recorded neurons when the animal was at rest and undergoing startling events. Such plots enabled researchers to "read" what was happening to an animal simply by watching the recorded signal move within that 3-D space. (See a movie clip at www.SciAm.com/ontheweb)

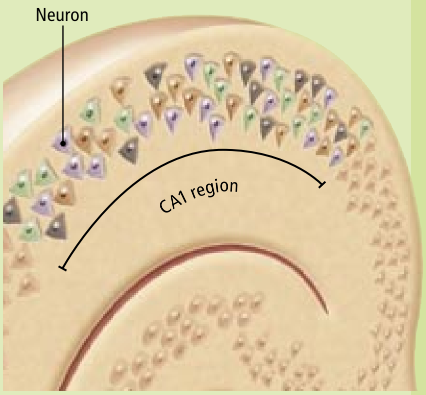
Plot of overall CA1 responses



❸ DISCOVERED CODING CLIQUES

Further analyses revealed that neuron ensembles active during an event contain subsets—termed neural cliques. The cells in a clique all show very similar firing patterns and are not part of the other cliques.

Schematic view of cliques encoding the earthquake experience
(each color represents one clique)



ine how such processes are affected in smart mice and learning-impaired ones.

Our findings suggest a number of things about the organizing principles that govern the encoding of memory.

The Power of Cliques

With the patterns indicative of specific memories in hand, we sought to understand how the neurons among those we were "tapping" actually work together to encode these different events. By coupling another mathematical tool called hierarchical clustering analysis with the sequential MDA methods, Osan and I discovered that these overall network-level patterns are generated by distinct subsets of neural populations that we have dubbed "neural cliques." A clique is a group of neurons that respond similarly to a select event and thus operate collectively as a robust coding unit.

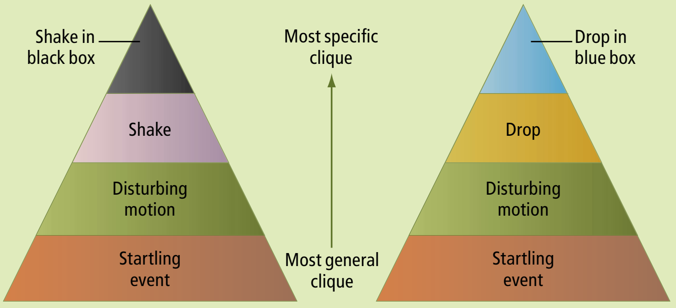
Furthermore, we found that each specific event is always represented by a set of neural cliques that encode different features ranging from the general to the specific. Notably, an earthquake episode activates a general startle clique (one that responds to all three startling stimuli), as well as a second clique that responds only to the events involving motion disturbance (both the earthquake and the elevator drop), a third clique that is activated exclusively by shaking and a fourth clique that indicates where the event took place (we put the animal in one of two different containers before each quake). Thus, information about these episodic events is represented by neural clique assemblies that are invariantly organized hierarchically (from general to specific). We think of the hierarchical arrangement as forming a feature-encoding pyramid whose base encodes a general characteristic (such as "startling event") and whose apex represents more specific information (such as "shaking" or "shaking in the black box") [*see panel 4 in box on opposite page*].

The CA1 region of the hippocampus receives inputs from many brain regions and sensory systems, and this feature most likely influences what type of information a given clique encodes. For example, the clique that responds to all three startling events could be integrating infor-

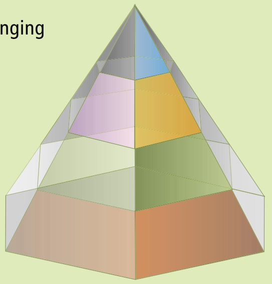
ALICE Y. CHEN (1, 3); LUCY READING-IKKANDA (2, 4, 5)

④ FOUND ORGANIZATION OF MEMORIES

Other analyses showed that each clique encodes a different aspect of an experience, ranging from the general to the specific. The author conceives of this hierarchical organization as a pyramid with the most general clique at bottom, as is shown below for two events. (The sizes of the pyramid "layers" do not signify the number of neurons in the cliques.)



Earthquake pyramid **Elevator drop pyramid**



Startling event polyhedron

Any given pyramid can be a component of a polyhedron representing all events of a given category, such as "all startling events."

⑤ TRANSLATED BRAIN ACTIVITY INTO BINARY CODE

The investigators then represented clique activity as a string of binary code that revealed details of the event an animal experienced. In the string fragments shown here, a 1 means a particular clique was active and a 0 signifies inactivity. Binary translations of neural activity could prove useful in many realms, such as helping investigators to peer into minds of those who cannot speak or to advance development of robots controlled by thoughts alone.

Clique	General startle	Motion	Air puff	Drop	Shake
Earthquake binary code	1	1	0	0	1
Elevator drop binary code	1	1	0	1	0

mation from the amygdala (which processes emotions such as fear or the experience of novelty), thereby encoding that "these events are scary and shocking"; the cliques that are activated by both the earthquake and the elevator drop, on the other hand, could be processing input from the vestibular system (which provides information about motion disturbance), thus encoding that "these events make me lose my balance." Likewise, the cliques that respond only to a particular event occurring at a particular place could be integrating additional input from place cells (neurons that fire when a creature passes through a particular familiar spot in its environment), thereby encoding that "this earthquake took place in the black container."

The Road to Knowledge

Our findings suggest a number of things about the organizing principles that govern the encoding of memory. First, we believe that neural cliques serve as the functional coding units that give rise to memories and that they are robust enough to represent information even if some individual neurons in the ensemble vary somewhat in their activity. Although the idea that memories and perception might be represented by neural populations is not new, we think we have the first experimental data that reveal how such information is actually organized within the neural population. The brain relies on memory-coding cliques to record and extract different features of the same event, and it essentially arranges the information relating to a given event into a pyramid whose levels are arranged hierarchically, from the most general, abstract features to the most specific aspects. We believe, as well, that each such pyramid can be thought of as a component of a polyhedron that represents all events falling into a shared category, such as "all startling events."

This combinatorial, hierarchical approach to memory formation provides a way for the brain to generate an almost unlimited number of unique network-level patterns for representing the infinite number of experiences that an or-

PRINCETON UNIVERSITY (*mouse*)

WHAT'S NEXT?

In future work, the author hopes to explore these questions, among others, in mice:

- **Do subpopulations of neurons within a given clique encode different aspects of an event? For instance, do cliques that record memories of fear include a subset that responds to the intensity of the fear, while another subset just notes the frightening nature of the event?**
- **How do memory traces—recurrences of the firing patterns that occurred when a memory was first laid down—differ soon after an event and much later on? How do false memories arise over time?**
- **How might the binary codes extracted from electrical signaling in the brain be used to download memories and thoughts directly into computers and to control robotic machines or assist real-time monitoring of learning processes?**

TO YOU, A DISH TO ME, A NEST

Recently published work supports the idea that some neural cliques in the hippocampus indeed encode abstract concepts. Some cells in mice turn out to react to items having varied shapes and textures only if the things have accessible depressions and can thus function as a nest. Cover the depressions, and the cells no longer respond.

▼ **MOUSE RELAXES** in a dish it views as a nest.

LONGNIAN LIN

ganism might encounter during life—similar to the way that the four "letters" or nucleotides that make up DNA molecules can be combined in a virtually unlimited number of patterns to produce the seemingly infinite variety of organisms on earth. And because the memory code is categorical and hierarchical, representing new experiences might simply involve substituting the specific cliques that form the tops of the memory pyramids to indicate, for example, that the dog barking behind the hedge this time is a poodle instead of a German shepherd or that the earthquake took place in California rather than in Indonesia.

The fact that each memory-encoding pyramid invariably includes cliques that process rather abstract information also reinforces the idea that the brain is not simply a device that records every detail of a particular event. Instead neural cliques in the memory system allow the brain to encode the key features of specific episodes and, at the same time, to extract from those experiences general information that can be applied to a future situation that may share some essential features but vary in physical detail. This ability to generate abstract concepts and knowledge from daily episodes is the essence of our intelligence and enables us to solve new problems in the ever changing world.

Consider, for instance, the concept of "bed." People can go into any hotel room in the world and immediately recognize the bed, even if they have never seen that particular bed before. It is the structure of our memory-encoding ensembles that enables us to retain not only an image of a specific bed but also a general knowledge of what a bed is. Indeed, my colleagues and I have seen evidence of this in mice. During the course of our experiments, we accidentally discovered a small number of hippocampal neurons that appear to respond to the abstract concept of "nest." These cells react vigorously to all types of nests, regardless of whether they are round or square or triangular or made of cotton or plastic or wood. Place a piece of glass over the nest so the animal can see it but can no longer climb in, and the nest cells cease to react. We conclude that these cells are responding not to the specific physical features of the nest—its appearance or shape or material—but to its functionality: a nest is someplace to curl up in to sleep.

The categorical and hierarchical organization of neural cliques most likely represents a general mechanism not only for encoding memory but also for processing and representing other types of information in brain areas outside the hippocampus, from sensory perceptions to conscious thoughts. Some evidence suggests this supposition is true. In the visual system, for example, researchers have discovered neurons that respond to "faces," including human faces, monkey faces or even leaves that have the shape of a face. Others have found cells that respond only to a subclass of faces. Back in the hippocampus, researchers studying patients with epilepsy have discovered a subset of cells that increase their firing rates in response to images of famous people. Itzhak Fried of the University of California, Los Angeles, further made the fascinating observation that one particular cell in a patient's hippocampus seems to respond only to the actress Halle Berry. (Perhaps it is part of a Halle Berry clique!) Together such observations support the notion that the general-to-specific hierarchical organization of information-processing units represents a general organizing principle throughout the brain.

Remember 11001?

Our work with mice also yielded a way for us to compare patterns from one brain to another—and even to pass information from a brain to a computer. Using a mathematical treatment called matrix inversion, we were able to translate the activities of neural clique assemblies into a string of binary code, where 1 represents an active state and 0 represents an inactive state for each coding unit within a given assembly we examined. For example, the memory of an earth-

ON HUMAN MIND READING

Our increasing ability to read the minds of mice raises an intriguing possibility: if enough neurons in a human brain could be recorded simultaneously, such recordings could well be able to reveal human thoughts.

Of course, to be practical, this technology would have to be noninvasive. Existing tools, such as EEG monitors and functional magnetic resonance imaging devices, are noninvasive but are not sensitive enough. They record averaged signals from or oxygen consumption by millions of nerve cells. Using such tools would be like listening in on a crowded football stadium from the outside; noise would simply overwhelm any individual conversations.

If a sensitive method existed, it could potentially be used to determine whether someone who seems to be in a vegetative state is actually able to think or whether someone with Alzheimer's disease who can no longer talk is able to understand conversation. Such "mind reading" might also be helpful for diagnosing mental disorders or assessing how well some medications are working. Much better lie detectors would also be possible.

With such benefits, though, would come major moral, philosophical and societal questions that would have to be addressed. Each of us might like to read other people's minds, but who among us would want our own mind read by others? *—J.Z.T.*


DOOGIE

MOLECULES OF MEMORY

In 1949 Canadian psychologist Donald O. Hebb postulated that a memory is produced when two nerve cells interact in a way that somehow strengthens future signaling through the synapse—the contact point between two neurons. But it was not until the 1980s that scientists saw Hebb's rule in action in brain slices. By stimulating neuron pairs in the hippocampus with electrodes, Holger Wigström of Göteborg University in Sweden and his colleagues found that activating a presynaptic neuron (a signaling cell) at the same time as a postsynaptic neuron (the signal's recipient) led to enhanced synaptic efficacy: the postsynaptic neuron came to respond more vigorously to the same amount of input from its presynaptic partner. The researchers suggested that the NMDA receptor—a protein complex found in the membranes of postsynaptic neurons—acted as the coincidence detector responsible for this synaptic strengthening.

To test this hypothesis, my laboratory decided to genetically manipulate one version of the NMDA receptor, which comes in different forms. We confirmed that adult mice lacking NMDA receptors in the hippocampus showed profound memory deficits. But we also showed the opposite is true: when we boosted the production of a specific NMDA receptor subunit (known as NR2B) in the hippocampus and cortex, the resulting mouse strain—which we named Doogie—learned faster and retained memories longer than unaltered mice did.

We believe that NMDA receptor activation—and reactivation—may serve to inscribe the ensemble activity patterns of the neural cliques that encode memories, thereby linking memory traces from the molecular level to the network level. *—J.Z.T.*

quake might be recorded as "11001," where the first 1 represents activation of the general startle clique, the second 1 represents activation of the clique that responds to a motion disturbance, the first 0 indicates lack of activity in the air puff clique, the second 0 indicates lack of activity in the elevator drop clique and the final 1 shows activation of the earthquake clique. We have applied a similar binary code to the neural ensemble activity from four different mice and were able to predict, with up to 99 percent accuracy, which event they had experienced and where it had happened. In other words, by scanning the binary code we could read and compare the animals' minds mathematically.

Such a binary code of the brain could also provide a potentially unifying framework for studying cognition, even across animal species, and could greatly facilitate the design of more seamless, real-time brain-to-machine communication. For example, we have arranged a system that converts the neural activity of a mouse experiencing an earthquake into a binary code that instructs an escape hatch to open, allowing the animal to exit the shaking container. We believe our approach provides an alternative, more intuitive decoding method for powering the kinds of devices that have already allowed patients with neural implants to control a cursor on a computer screen or a monkey to move a robotic arm using signals recorded from its motor cortex. Moreover, real-time processing of memory codes in the brain might, one day, lead to downloading of memories directly to a computer for permanent digital storage [see "A Digital Life," by Gordon Bell and Jim Gemmell; Scientific American, March].

In addition, we and other computer engineers are beginning to apply what we have learned about the organization of the brain's memory system to the design of an entirely new generation of intelligent computers and network-centric systems, because the current machines fail miserably in the type of cognitive decision making that humans find easy, such as recognizing a high school classmate even though he has grown a beard and aged 20 years. Someday intelligent computers and machines equipped with sophisticated sensors and with a logical architecture similar to the categorical, hierarchical organization of memory-coding units in the hippocampus might even do more than imitate, perhaps exceeding our human ability to handle complex cognitive tasks.

For me, our discoveries raise many interesting—and unnerving—philosophical possibilities. If all our memories, emotions, knowledge and imagination can be translated into 1s and 0s, who knows what that would mean for who we are and how we will operate in the future. Could it be that 5,000 years from now, we will be able to download our minds onto computers, travel to distant worlds and live forever in the network? ■

SCI AM

More on the Web

For an extended version of this article, visit **www.SciAm.com/ontheweb**

HUIMIN WANG

MORE TO EXPLORE

The Oxford Handbook of Memory. E. Tulving and F.I.M. Craik. Oxford University Press, 2000.

Identification of Network-Level Coding Units for Real-Time Representation of Episodic Experiences in the Hippocampus. L. Lin, R. Osan, S. Shoham, W. Jin, W. Zuo and J. Z. Tsien in *Proceedings of the National Academy of Sciences USA*, Vol. 102, No. 17, pages 6125–6130; April 26, 2005.

Organizing Principles of Real-Time Memory Encoding: Neural Clique Assemblies and Universal Neural Codes. L. Lin, R. Osan and J. Z. Tsien in *Trends in Neurosciences*, Vol. 29, No. 1, pages 48–57; January 2006.

Neural Encoding of the Concept of Nest in the Mouse Brain. L. Lin, G. Chen, H. Kuang, D. Wang and J. Z. Tsien in *Proceedings of the National Academy of Sciences USA*, Vol. 104, No. 14, pages 6066–6071; April 3, 2007.