Silicon Brains

Innovative computer devices are being inspired by the results of research on the brains of nature's creatures

T. J. SEJNOWSKI AND P. S. CHURCHLAND

Imagine a computing device that would revolutionize the stoop-labor sector of agriculture or perform many of the necessary but tedious tasks in other industries. Such agribots would need computational insides that are small, inexpensive, and enormously powerful.

Existing digital computers lack the efficiency, autonomy, flexibility, and adaptability required by the fictional agribots. However, the brains of birds, fish, mammals, and even insects prove that powerful, fast, flexible, and self-reliant computers can solve these problems (see the text box "Bee Smart" on page 142). Breakthroughs in neuroscience, combined with new computational devices such as analog VLSI chips, have made it possible to begin to reverse-engineer nature.

In addition to the sheer intellectual value of understanding ourselves, understanding how brains work could produce important economic benefits. If you know evolution's computational tricks and architectural ingenuities regarding speed, power, and flexibility, you can apply them to a variety of areas: image processing, speech recognition, free-form handwriting recognition, and holographic applications.

Current-generation neural networks capture some of the brain's general features (e.g., the parallel architecture). But neural networks represent only the beginning of brain-style computer technology.

Computing in Parallel

While digital technology is still very much in its heyday, there is tremendous potential in analog VLSI for addressing real-world problems. For example, current algorithms running on a digital machine can correctly read written numerals on credit-card sales forms about 60 percent of the time. The problem of machine-reading postal ZIP codes on letters is compounded by the problem of locating the digit, which is unsolved.

The crux of the difficulty is that digital machines are typically programmed to solve the segmentation problem (e.g., what character does a squiggle belong to?) and after that, to solve the recognition problem (e.g., is it a 0 or a 6?). Should the machine missolve or fail to solve the segmentation problem, recognition is doomed.

Brains, it appears, do not serialize the segmentation and recognition problems in lockstep fashion. As often as not, recognitional cues are used to solve the segmentation problem. In general, people believe that the brain's approach more closely resembles cooperative computation or constraint satisfaction than theorem proving. Of course, it takes a lot more computing to be able to solve the segmentation and recognition problems in parallel. With analog VLSI technology,

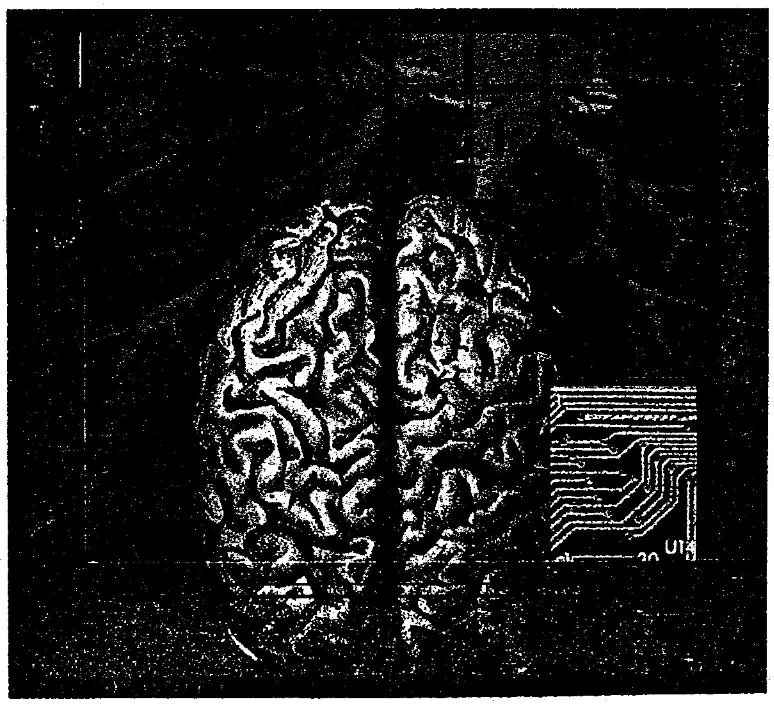
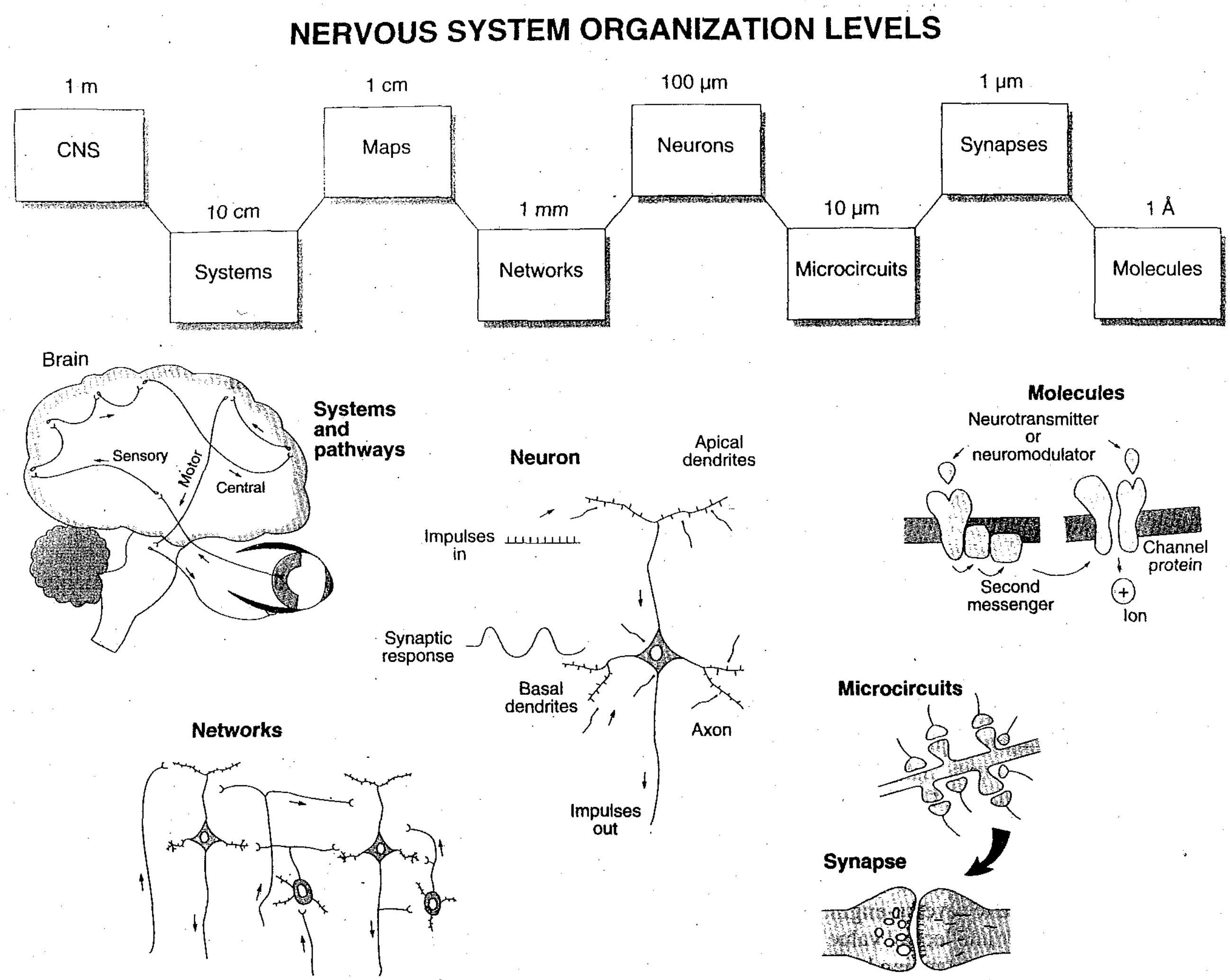


ILLUSTRATION: LYN BOYER PENNINGTON © 1992



The components of nervous systems include the brain and spinal cord, systems (e.g., the visual system), maps (e.g., the retina or the skin), networks (perhaps of many thousands of interconnected neurons), the individual neuron, microcircuits, synapses, and ion channels.

people are learning how to build machines that really compute in parallel.

Reverse Engineering the Brain

Computational neuroscience is the study of how the brain represents the world and how it computes. Being able to model the brain's neural circuits by computer is essential in finding out how neurons (i.e., the cellular components of nervous systems) interact with each other to produce complex effects (see the figure). Such effects include segregating a figure from its background, recognizing a banana from different angles, and following items moving in 3-D space.

Neuroscience contributes three main ingredients to this effort: anatomical parameters (e.g., the precise tree structure of various neuron types and the exact mode of connectivity between neurons in a particular real network), physiological parameters (e.g., the response characteristics of neurons, time constants, and synaptic strengths), and clues to the function of the human biological neural network and its computational mode of operation in executing that function.

Many techniques that neuroscientists use to study the brain involve intervention—lesioning or electrical stimulation. Analyzing a working model can provide neurobiologists with information about unsuspected mechanisms and interaction; they can then test the results under actual conditions.

This type of collaboration between computer modeling and neuroscience is already producing ideas for new and innovative computing procedures. It has resulted in architectural designs for interacting in real time, storing associative memory more efficiently, coordinating

Applications Benefiting from Brain Research

- Image processing
- Optical character recognition
- Speech recognition
- Handwriting recognition
- Holography

Glossary

axon Part of a neuron that conducts impulses away from the cell body.

Brownian motion The random movement of particles caused by the collision of the molecules in the fluid around those particles.

cortical structure Structures found in the cortex, a region of the brain.

cytoplasm The fluid outside the nucleus but within the membrane of a cell.

dendrite Part of a neuron that conducts electrical signals toward the cell body.

ion channel Proteins in the cell membrane that may reconfigure to let specific ions (e.g., Ca^{2+}) enter the cell in response to chemical or electrical signals.

lesion An abnormal change in an organ's structure due to injury, disease, or an experimental procedure.

maps Regions of the brain where the topography of neurons corresponds to the topography of the sensory surface (e.g., the retina or the skin).

mitochondria Structures found in the cytoplasm that produce energy through cellular respiration.

neurons The functional units of the brain (i.e., the cellular components of the nervous system). An individual neuron can be either excited or inhibited by inputs from other neurons.

photon A packet, or *quanta*, of electromagnetic energy (e.g., light).

photoreceptor A receptor for visible light stimuli.

pyramidal neuron A type of neuron found in cortical structures.

synapse The point of contact between adjacent neurons where nerve impulses are transmitted from one neuron to the other.

mixed modality, multiplexing, and understanding attention selectivity.

Simulate or Synthesize?

Digital machines are not yet powerful enough to faithfully simulate the nervous system's processes and do it in real time. One or the other is sacrificed. The problem is that the simulation strategy consists of compartmentalizing the phenomena and solving vast numbers of differential equations; thus, compared to the real thing, it is pitifully slow.

In a neuron, ions pass back and forth across a membrane, signals are integrated, and output spikes are produced—all in a matter of a few milliseconds. However, to simulate just 1 ms in a neuron's life, computers must solve thousands of coupled nonlinear differential equations.

To compound the problem, these equations use a wide variety of time scales. In the simulation, the time steps can be only as long as the shortest significant interval. Consequently, even a powerful workstation will take minutes to simulate 1 ms of real time of the electrical and chemical events occurring in a single neuron. You can circumvent this problem by constructing dedicated hardware for synthetic neurons and nervous systems. One strategy is to construct neuron-like chips.

To construct chips that compute as well as neurons do, you must first understand how neurons perform. The production of a spike in a neuron's axon is an all-or-nothing affair. Even axonal spiking is analog in some respects (e.g., when spikes occur, how frequently spikes happen, and how long it takes to repolarize them). The main analog integration of synaptic inputs occurs in dendrites.

Real circuits have many imperfections. Invariably, they do not meet the ideal, the components are not homogeneous, membranes leak, components malfunction or drop dead, and cross-coupling occurs. But you can't shun chip construction entirely in favor of performing simulations. The best long-term direction people should take seems to be to find out how real circuits obtain precision, speed, and power from imperfect and imprecise components. Somehow, neurons operate in real time and cope magnificently, probably by exploiting imperfections to their advantage. The coping capacity of real-world neurons is itself computationally interesting.

Neurons are organic. They use fatty molecules to make resistive membranes. Complex proteins make ion channels that let current flow across the membrane, and cytoplasm acts as the medium for transmitting current. Mitochondria are the neuron's miniature powerpacks; circulating oxygen is their energy source. But what can you use to construct synthetic neurons?

You can't shun chip construction entirely in favor of performing simulations.

Synthetic Neurons

Analog VLSI technology turns out to be well suited to constructing synthetic neurons for two reasons. One is theoretical, and the other is practical.

The device physics of doped silicon operating in subthreshold regions is comparable to the biophysics of ion channels in the neuron membrane. Therefore, you can implement the differential equations directly with analog circuits in CMOS VLSI. And the same techniques used to create digital VLSI chips can be adapted to make analog VLSI chips. Carver Mead of Caltech and Synaptics and Federico Faggin of Synaptics, industry pioneers

Bee Smart

Start by contrasting what a small honeybee can do with tasks that today's most powerful computers can't do, and add the fact that a honeybee's brain has only about 1 million neurons versus the human brain's 100 billion neurons. Then consider the following information:

Energy efficiency. A honeybee's brain dissipates less than 10 microwatts (10^{-6}). It is superior by about seven orders of magnitude to the most efficient of today's manufactured computers.

Speed. A honeybee's brain, roughly and conservatively, performs at about 10 TFLOPS (10,000 GFLOPS). The most powerful of today's computers approach speeds of only 10 GFLOPS (i.e., 1 billion operations per second).

Behavioral abilities. Honeybees harvest nectar from flowers and bring it back to the hive. They maximize foraging benefits and minimize foraging costs—for example, by recognizing high nectar sites and remembering which flowers they have already visited.

Honeybees can see, smell, fly, walk, and maintain balance. They can navigate long distances and predict changes in nectar location. They communicate the location of nectar sources to worker bees in the hive; they recognize intruders and attack; they remove garbage and dead bees from the hive; and, when the hive becomes crowded, a subpopulation will swarm in search of a new home.

Autonomy and self-reliance. Honeybees manage these activities entirely on their own without any help from superior beings. By contrast, a supercomputer needs the constant tender care of a cadre of maintainers and programmers.

Size. A honeybee's brain takes up only about a few cubic millimeters of space. It is a marvel of miniaturization. You cannot reach all the way around a supercomputer.

From this comparison, it seems we have a ways to go in allowing computers to perform some of the simpler things in life. Nature and its creatures are models for ways in which to improve our computing devices. (See photo A for an inside view of the insect visual system.)

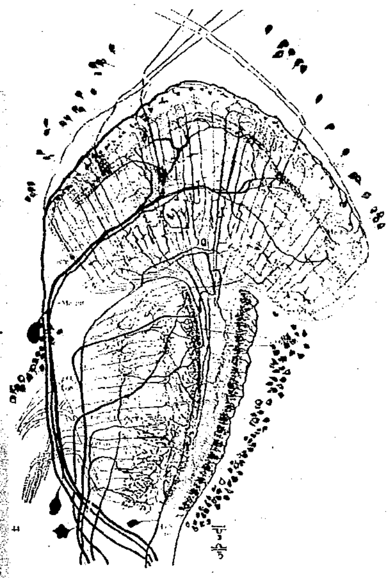


Photo A: *Anatomy of a fly's visual system. Like that of other insects, a fly's visual system is highly organized. The structure of the fly's system shows how the neurons (black) are organized into layers of visual processing. (Drawing by N. J. Strausfeld,* Atlas of an Insect Brain, *New York: Springer-Verlag, 1976)*

who played leading roles in digital chip technology, are now spearheading the development of analog chip technology for neural systems.

With analog VLSI, a chip can follow the brain's lead—for example, concurrently solving segmentation and recognition problems. As reported this year in *Nature*, Misha Mahwold and Rodney Douglas, both of Oxford and Caltech, achieved the first step in building silicon neurons (see the text box "Silicon Neurons" on page 144).

Using analog VLSI, Mahwold and Douglas created a chip that mimics selected properties of pyramidal neurons, a type of neuron found in cortical structures. Their silicon neuron consists of only one compartment (the cell body) and four types of ion channels in the membrane. By contrast, a real pyramidal neuron might have thousands of dendritic segments, as well as an axon, tens of thousands of synapses, and scores of various ion channels.

As a pilot project, however, the Mahwold/Douglas silicon neuron was successful on several counts. First, it ran in real time. This meant that Mahwold and Douglas could conduct experiments by tweaking parameters in real time, such as the density of a given type of channel. Second, the neuron's output behavior for varying amounts of current (displayed on an oscilloscope) closely resembled that of a real pyramidal cell under various physiological conditions. Third, the neuron consumes little power.

The successful debut of a single synthetic neuron has made possible several other potential developments: By adding more compartments (corresponding to dendrites) and a wider range of ion channels, you could improve the synthetic neuron's computational capabilities. Another possible development is that of building many neurons on a single chip. You could then explore synthetic neural circuits to learn more about the computational possibilities inherent in various parameters.

Ideally, you should be able to tweak thousands of parameters in real time; thus, interfaces need to be flexible and user friendly. Using synthetic circuits would mean that you could explore neurons in virtual reality rather than having to watch points appear on a graph on your screen.

A further refinement would be to make the chip able to learn from experience. Then, instead of having to hand-set neuronal connections, you could use a training regime. Mead and his group are currently developing trainable chips that can modify connectivity based on learning certain rules similar to those

Silicon Neurons

RODNEY DOUGLAS AND MISHA MAHWOLD

Neurons in the living body have electrical and chemical mechanisms that let them act together to represent and respond to behaviorally significant physical events. Over time, neurons have learned to manipulate how their membrane conducts various ions to produce electrical events that form a basis for computation.

Neuronal systems compute in fundamentally different ways than electronic computers do. Neurons are massively interconnected. The neurons shown in the reconstructed neocortical pyramidal cell (see photo A) receive input to their dendrites (green) from thousands of input cells and transmit to thousands of output cells via the synapses (white) made by their axons (red).

Neurons operate in the millisecond range rather than in the nanosecond range. The human brain generates 10^{16} operations per second (compared to the supercomputer's 10^9 operations per second). But the power consumption of the brain is only 10^{-15} joules per operation (compared to an electronic processor's roughly 10^{-7} joules per operation).

The perception of an object is an unsolved computational problem. The vast majority of neural computations as complex as perception are less than 100 operations deep. This fact reflects the essentially distributed nature of neural computation, in which algorithms express themselves as connectivity and processors are indistinguishable from memory. Certainly, it seems that people can learn about computing from the field of biology.

Neuroscientists are learning about neural computation through reverse engineering. They combine experimental neuroscience with neuromorphic systems made from analog CMOS VLSI technology. Fortunately, the physical properties of analog CMOS are similar to those governing the electrical behavior of neurons and neural systems; therefore, analog CMOS is a convenient medium for building neuromorphic systems, just as the properties of Lego make it appropriate for constructing structures and machines.

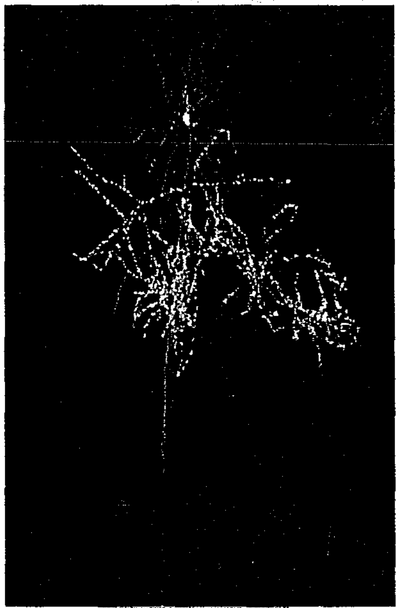


Photo A: *A reconstructed neocortical pyramidal cell from a cat. The cell's neurons receive input to their dendrites (green) from thousands of input cells and transmit to thousands of output cells via the synapses (white) made by its axons (red). (Photo courtesy of Rodney Douglas and Kevan Martin at MRC Cambridge)*

For example, we fabricated a generic silicon neuron that emulates the fluxes of the ionic currents that occur in real neurons. Consequently, the silicon neuron has the same computational properties (at the neuronal level) that real cells do. The neuron can emulate the behavior (i.e., personality) of any particular neuron in the nervous system simply by setting several parameters.

One exciting feature of the silicon neuron is that it behaves in real time regardless of its complexity or the number of neurons in the network. We are currently working to build many neurons, initially about 100 to 200 neurons, on a single chip.

In the not-too-distant future, we anticipate building networks of thousands of silicon neurons on multiple chips, with personalities and connectivity that can be modified in real time. Using these silicon neural networks, we will be able to emulate intelligent circuits in the brain (e.g., those of the visual system) and provide a test bed to investigate realistic learning mechanisms.

Rodney Douglas and Misha Mahwold are members of the Medical Research Council Anatomical Neuropharmacology Unit at Oxford and the Computation and Neural Systems Program at Caltech. You can reach them on BIX c/o "editors" or on the Internet at rjd@vax.oxford.ac.uk and mam@vax.oxford.ac.uk, respectively.

believed to underlie plasticity in nervous systems. Here, *plasticity* refers to a property of a neuron's body that undergoes a permanent change in shape, size, or composition under certain conditions.

Ultimately, you will want to create chips with subpopulations of neurons specialized for different tasks, in the manner that distinct brain regions—including visual cortex, auditory cortex, motor cortex, and so forth—are specialized. Learning from the ways that nature engineers specialization and integration functions should provide valuable information.

Following nature's lead may require that people model patterns of neuronal connectivity, both long-range (on the order of centimeters) and short-range (millimeters). Nervous systems are remarkably fault-tolerant: A circuit and its ability to function can survive the death of individual neurons within the circuit. Artificial systems might be able to achieve comparable fault tolerance if they are made to imitate the brain's connectivity, modifiability, and processing style.

continued

ANALOG VLSI VS. DIGITAL VLSI

Analog VLSI is strikingly superior to digital technology in terms of cost, power, and computation density. (Estimates by Federico Faggin.)

	Cost (MCS*/$)		Power (MCS/watt)		Computation density (MCS/ft.^{3})	
	1991	2000	1991	2000	1991	2000
Conventional digital	0.002	0.1	0.1	10	0.2	10
Special-purpose digital	0.1	4	10	10,000	10	1000
Dedicated digital	5	200	500	50,000	40	3000
Dedicated analog	500	20,000	50,000	5,000,000	4000	4,000,000
Human brain	10^9**		10^{10}		10^{11}	

* MCS = A million connection updates per second.
** This calculation assumes that the cost of a human brain is $10,000,000.

Neural Circuits in Silicon

Peripheral sensory organs (e.g., the eye) are highly specialized parts of the body that translate external physical signals into electrical activity. The retina is a powerful preprocessor that transforms information about photons into a form suitable for neural representation and computation.

The retina is a powerful preprocessor that transforms information about photons into a form suitable for neural representation and computation.

In a number of animals, sensory transducers and preprocessors are about as sensitive as they can get. For example, in primates, photoreceptors in the retina will respond to just a few photons; the human ear can pick up sound close to that of Brownian motion. Powerful analog preprocessors shape the information into a neural-friendly form—but can they be reverse-engineered?

Mead has built a family of silicon retinas. Each silicon retina is a VLSI chip that is a square centimeter in area, weighs about a gram, and consumes about a milliwatt of power. Between arrays of phototransistors etched in silicon, dedicated circuits execute smoothing, contrast enhancement, and motion processing. The chip operates in its subthreshold, analog mode.

Compared with a typical CCD (charge-coupled device) camera and standard digital image processor, the Mead chip is a paragon of efficiency in performance, power consumption, and compactness. A special-purpose digital equivalent would be about the size of a standard washing machine. Unlike cameras that must time sample, typically at 60 frames per second, the analog retina works continuously without needing to sample until the information leaves the chip already preprocessed.

Operations performed with Mead's chip capture some of the functions that real retinas perform; however, real retinas contain many more circuits than Mead's synthetic one. While it makes sense to build chips to maximize efficiency in the three critical elements (i.e., power, cost, and density), you must still push analog VLSI techniques a long way to approximate neural efficiency. The incentive to go forward with this technology will depend on whether the payoff looks promising in the long term (see the table).

Neuro-Revolutions

We are on the brink of two neuro-revolutions: one in the science of the brain and the other in the technology of brain-style computing. Knowledge grows exponentially: The more you have, the more you get—and the faster you get it. So it is in neuroscience. Almost every day, surprising discoveries about the organization and mechanisms of nervous systems are being reported.

The VLSI revolution has provided computer science with unprecedented tools to transform what we know about the brain into silicon. Silicon retinas are in production, silicon cochlea are nearing production, and oculobots (i.e., robotic eyes) are on the drawing board. Although it is nearly impossible to predict future technological breakthroughs, ever-more sophisticated neuro-engineering is in the offing. ■

BIBLIOGRAPHY

Churchland, P. S., and T. J. Sejnowski. *The Computational Brain.* Cambridge, MA: MIT Press, 1992.

Dowling, John. *Neurons and Networks: An Introduction to Neuroscience.* Cambridge, MA: Harvard University Press, 1992.

Levitan, Irwin B., and Leonard K. Kaczmarek. *The Neuron: Cell and Molecular Biology.* New York: Oxford University Press, 1991.

Mead, Carver. *Analog VLSI and Neural Systems.* Reading, MA: Addison-Wesley, 1989.

T. J. Sejnowski is an investigator with the Howard Hughes Medical Institute, director of the Computational Neurobiology Laboratory at the Salk Institute (La Jolla, CA), and a member of the biology department at the University of California–San Diego. P. S. Churchland is a member of the philosophy department at the University of California–San Diego and is an adjunct member of the Salk Institute. You can reach them on BIX c/o "editors" or on the Internet at tsejnowski@ucsd.edu and pschurchland@ucsd.edu, respectively.