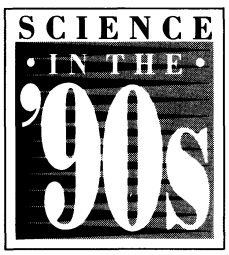
Neuroscience Models the Brain

After many halting inititial steps, mathematical simulations of brain functions have come close enough to reality that they are now a powerful guide to experimentation



Third in a series

HOW DRAMATICALLY THINGS CAN CHANGE in neuroscience! A decade ago a computer scientist who approached a neurobiologist with a new model of how the brain works was likely to get a shrug—or perhaps something less polite. To those studying the biological circuitry of living nervous systems, abstract conceptions of how those circuits might work seemed to bear little relation to reality.

Today a computer scientist offering a model of the brain is likely to be received warmly, particularly if the model can make testable predictions about the behavior of nervous systems. Furthermore, neurobiologists are sharpening up their own computer skills to get into the game themselves. Indeed, the marriage of computational models and experimentation promises to be a key trend in neuroscience in the 1990s.

It should not be assumed that modeling is on the verge of unlocking the brain's mysteries. In fact, so far computer models have not brought any major revelations about how the brain works. But they have begun to help researchers chisel away at some unsolved questions, such as how neural activity shapes brain development and how specific groups of neurons process information. The optimistic view is that such modeling will pick up steam, and as more is learned about the nervous system, increasingly sophisticated models will be created, bringing into reach larger questions—how we learn a language, for example, or how we recognize faces.

This interest in computational approaches is not surprising. Neuroscientists are quick to snatch up any new trick that may shed light on the workings of the brain. Techniques and tools such as monoclonal antibodies, recombinant DNA, and patch clamping (a method of recording ion flow through single channels in cell membranes) were eagerly and quickly adopted by the neuroscience community. But many experimental neuroscientists have been slow to accept the notion that computation and modeling have something tangible to offer.

Part of the reason for the skepticism is that neuroscientists remember previous models that didn't live up to their billing. For instance, computer simulations devised in the 1970s as models of how neural activity influences the development of the visual cortex came under criticism for simply describing what was known rather than making predictions that could be tested in experiments. Such experiences made those who study the brain leery of the idea that computer models could be anything more than high-tech simulations of the obvious or elaborate but possibly uninformed guesses about how the brain might work.

What has changed? One difference is the increasing sophistication of computer modeling techniques. That sophistication, coupled with the growing power of the computers that are available to every lab scientist, has brought greater modeling power close to hand. Another equally important factor is the rapid increase in knowledge of the biology of the nervous system—often without a suitable theoretical framework in which to understand it.

"The problem is that we're just awash in data," says Allen Selverston of the University of California at San Diego. "But how much of that is important for the operation of the system and how much isn't important? If you want to know how a car works, you don't have to analyze the paint pigments on the body. We have to start to know what are the important parameters. Modeling is one way to try to do that."

And once those parameters are identified, models can point the way to experimentation, according to Terrence Sejnowski of the Salk Institute. "Models won't solve problems by themselves," notes Sejnowski, who has been a leader in bringing computer scientists together with neurobiologists. "You still have to do experiments. But having a model greatly amplifies your intuition. . . . Doing an experiment is often very difficult; it's critically important that you pick ones that are going to give you the maximum payoff. A model lets you play with a lot of different experiments in a reduced, simpler form."

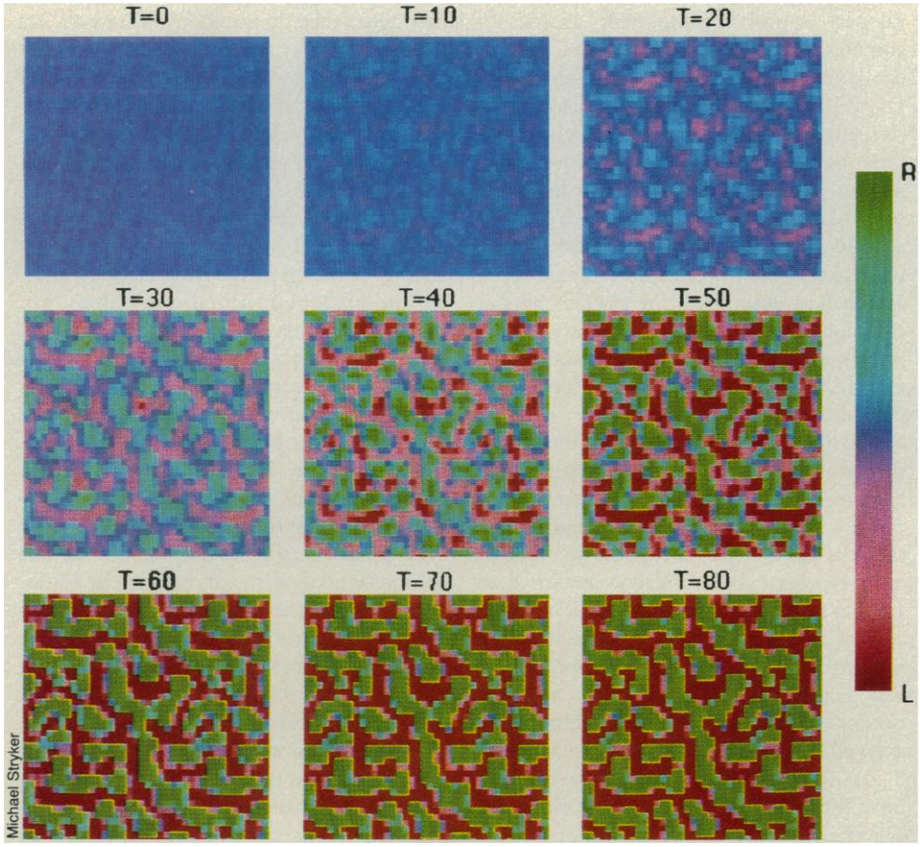
An excellent example of the power of models to identify key parameters and point the way to experiments comes from work by Michael Stryker of the University of California at San Francisco. Stryker used modeling to analyze the factors that influence the formation of ocular dominance columns, patches of nerve cells in the visual cortex that respond to signals from only one eye. In a newborn animal, the columns don't exist; all cells in the region receive impulses from both eyes. But in each small patch of cortex the connections from one eye gradually weaken while those from the other eye strengthen, yielding a pattern of patches of cells that serve one eye or the other.

Stryker knew that the development of ocular dominance columns was probably influenced by the extent of branching of the neurons coming into the cortex, the correlation of their firing patterns, and the inhibitory or excitatory interplay of neurons in the area of the cortex where the columns are forming. What he didn't know was which of those three variables was most important.

Then Kenneth Miller joined Stryker's lab as a graduate student and decided to tackle the problem by modeling. With the help of Stanford mathematician Joseph Keller, Miller, previously trained in theoretical physics, devised a model consisting of a series of equations that predicts the relative influence of the three factors. Solution of the equations indicated that interactions in the cortex would be the dominant influence, a hypothesis Stryker and Miller are now testing experimentally.

An essential element of this work, according to Stryker, was cooperation between a neurophysiologist, a mathematician, and a student trained in both areas. Such students are the leading edge of the next wave of neuroscientists, says Max Cynader, a neurophysiologist at the University of British Columbia, who has had a fruitful collaboration with Steven Zucker, who specializes in computational vision at McGill University, and Allan Dobbins, a graduate student with Zucker. The work was greatly enhanced by Dobbins, whom Cynader describes as "one of those people of the future: both a neurobiologist and an engineer—and a computational wizard as well."

The group decided to take a computation-



Separate but equal. *Computer simulation by Michael Stryker models formation of ocular dominance columns. Patches of visual cortex initially respond to signals from either eye (blue) but later respond only to signals from left eye (red) or right eye (green).*

al approach to the problem of how the brain uses cues from shading generated when light falls on a curved surface to perceive the curvature of objects. This capacity is likely to be a key to how we see in three dimensions and distinguish objects—and no one had identified the cells that are responsible.

As a first step, Cynader, Zucker, and Dobbins formulated a model to address the two-dimensional problem of detecting curved lines. The model was based on an analysis of the computations the brain would have to make to determine whether a line is curved. It predicted that the process requires several cell types, including cells that specifically respond to straight bars of different lengths.

Such cells were already known, Cynader says. They had been named "end-stopped" cells, because of their presumed function in end point detection. But no one had shown that end-stopped cells sense curvature. Directed by the model's prediction, Dobbins conducted experiments in cats and found that the cells indeed respond when the eye is trained on curved lines. They are now revising the model to address the problem in three dimensions.

Both Stryker's and Cynader's models hew closely to the physiology of their respective biological systems. And that is no accident: in each case parameters of the model were derived directly from the physiological characteristics of the cells. Many neuroscientists argue that if computational models are to be taken seriously as reflections of what's going on in the brain, such a direct relation is vital. But a different approach, using computer-simulated networks of neurons, which also shows promise for making experimental predictions, has been criticized for being insufficiently rooted in physiology.

Neural networks have a long history, although they have only recently achieved acceptance in the general neuroscience community. During the 1950s and 1960s there was a flurry of excitement over computer-simulated neural nets that could learn to recognize patterns by adjusting the strength of the connections between individual units in the network. The excitement died when the nets proved to be quite limited in what they could learn.

In 1983, Sejnowski, then at Johns Hopkins University, and Geoffrey Hinton, then at Carnegie-Mellon University, introduced the three-layer neural net, which was a vast improvement on its predecessors. In the three-layer net a middle layer of units connects the input and output layers. When the net is given an input, it sends signals through the hidden layer to produce an output. That output is checked against the "correct" output, and a learning algorithm is used to reduce error by strengthening or weakening connections in the net.

The great advance in this form of neural net was the middle layer, often referred to as "hidden," because its state is not apparent from an examination of the input or the output. But the collective activity of the hidden layer determines the output of the entire network. The most important property of the hidden layer is that it tends to distribute tasks: the computation is shared by many units rather than being delegated to only one. After the network has been "trained" to do a job, the hidden layer retains the record of how the network distributed the task in the process.

The three-layer neural net drew considerable attention in 1987 when Sejnowski used an improved learning algorithm, called back-propagation of errors (back-prop for short), to develop a program called NETtalk, which can learn overnight to convert written text to recognizable spoken speech. But while NETtalk pointed out the commercial potential of neural nets in "smart" machines, many neuroscientists remained doubtful that such networks could advance understanding of the nervous system.

One reason for their caution is the fact that the brain is unlikely to use a learning method like back-prop. "Alas, the back-prop nets are unrealistic in almost every respect," wrote Francis Crick in the 12 January 1989 issue of *Nature*. Crick has warned repeatedly against taking such models too literally. He points out that taking back-propagation at face value would require that individual neurons have the ability to both excite and inhibit their neighbors and be able to transmit impulses in the backward as well as forward direction—both distinctly unphysiological characteristics.

But physiologically faithful or not, neural nets have made some excellent predictions about how neurons behave. One success story comes from the lab of Richard Andersen at Massachusetts Institute of Technology. Andersen and other researchers had

spent years recording from cells in the brains of monkeys, trying to find the neurons responsible for computing the location of an object in space. The monkeys were trained to keep their eyes fixed on one point while an object appeared elsewhere in their visual space. By then directing the monkey's gaze to different spots, the experimenters could move the image of the object to different parts of the monkey's retina, although the object itself remained at the same point in space.

Even though the image falls on different parts of the retina, the monkey knows that the position of the object has not actually changed. The researchers were looking for neurons that keep track of the object's "true" position by firing only when a specific combination of image position on the retina and eye rotation corresponded to that particular point in space.

What they found was a population of cells that seemed to divide the labor in an unexpected way. No cells responded simply to a position in space. Rather, groups of cells would respond only to stimulation of specific parts of the retina and then only if the eye was in a certain position. The upshot of this was that no position in space was uniquely covered by any one group of neurons. "One place in space was covered by a number of groups of cells," says Andersen.

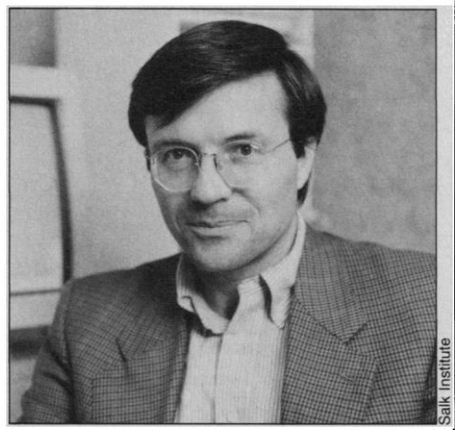
Given this overlap, it wasn't intuitively obvious that these cells could be doing the job of localization. Andersen and his collaborator David Zipser, of the University of California, San Diego, turned to a neural net to expand their intuition. They trained the net to locate an object on the basis of retina and eye position, then examined the hidden layer to see how the units there responded to various combinations of positions. What they found was the same kind of distribution they had seen when recording from nerve cells.

Andersen points out that their result does not prove that the cells they identified are actually performing the location task; it only shows that they could be. But the more tests the model passes, the stronger the case for it becomes. The next step is to add another variable the brain must deal with—head position. "The model makes some very powerful predictions about what you should find in the recordings," Andersen says. His laboratory is now conducting experiments with monkeys to see if the cells handle changes in head position as the model suggests they do.

Many other workers are also finding neural nets helpful in simplifying analysis of complex data. Thomas Anastasio, now at the University of Southern California, and David Robinson, of Johns Hopkins University, used a neural network to help determine how the brain combines visual information with head position cues from the inner ear to control three kinds of eye movement: the vestibulo-ocular reflex, which keeps your eyes focused on an object as you turn your head; the smooth pursuit response, by which your eye follows a moving object; and the saccadic response, in which the eye jumps suddenly to a different part of the visual space.

Recording from the vestibular nucleus of the monkey, the brain center where such movements are computed, Anastasio (then a postdoc with Robinson) found that most of the cells did not fire during just one type of eye movement, but instead were active to varying degrees during all three. As in the case of the locational task, the computing responsibility seemed to be distributed in a way that was not obvious, but a neural network trained to compute eye movements showed a remarkably similar distribution.

Anastasio and Andersen both say the value of the computer simulations was that, by making sense of cellular recordings, they provided a reasonable hypothesis about how neurons are actually doing certain complex processing tasks. But while the models provided the springboard for an important conceptual leap, they left another question open: Even if a neural net tells us how brain cells share the responsibility for a certain task, does it say anything about how the cells actually assume that form of distributed activity?



Master mingler. *Terrence Sejnowski has brought together computer scientists and biologists.*

This may be the most controversial element of neural net modeling. Neuroscientists agree that the brain is unlikely to use the precise survey and adjustment of connections required by the learning algorithm of back-propagation. But one can't ignore the fact that neural net models do a good job of mimicking some brain functions. How can this be explained?

Some researchers believe the brain may use a biological correlate of back-prop—an analogous means of reducing error by altering the strength of neural connections. For example, Robinson points out that the vestibulo-ocular reflex, which remains able to adapt and change into adulthood, contains neurons that are perfect candidates for a means of relaying information about error. The neurons he is referring to fire in response to retinal slip—the movement of an image across the retina that occurs when the eye does not accurately track an object. The neurons send their signals back to the vestibular nucleus where they could be acting to improve the accuracy of the eye movement, in effect implementing a learning algorithm that is more biological than back-prop, but has a similar outcome.

The next step in the refinement of neural net models, and other computational models of brain function, is to use information gathered from physiological systems to bring the models closer to biological reality. Already researchers are customizing the connections in their neural nets to match the neural connections in the systems they study. And others are replacing back-prop with learning algorithms that can work through separate neural pathways—creating feedback systems like the ones frequently found in the brain. Such systems are not only "more physiological," they also allow the neural networks to do at least one important brain-like thing: create a pattern of output that changes over time.

As neuroscience enters the 1990s, an important dialectic between model and experiment is becoming clear. Computational models have become powerful enough to make specific predictions—that are far from intuitively obvious—about how systems of neurons do their jobs. In turn, what is known about the physiology of brain cells is being used to modify the models and bring them into closer conformation with biological reality. As the models become more physiologically based, they will be able to make even stronger predictions. If this interaction is only in its infancy, and it has not yet produced revelations, it seems clear that in the next decade it will begin to do so.

■ MARCIA BARINAGA