The Brain in the Machine

By BRUCE BOWER

Biologically inspired computer models renew debates over the nature of thought

Neural networks — a group of computer models of how the brain might work — have generated much interest, not to mention hype, in the past few years. Yet while their ability to illuminate the dark recesses of the mind may have been exaggerated by ardent proponents, there remains a strong belief in some quarters that neural networks will link up with emerging studies of brain cells in action to produce new insights into how the human brain makes sense of the world and generates complex thoughts.

In fact, according to a report in the Sept. 9 SCIENCE, this field of "computational neuroscience" has already arrived.

Its ultimate aim is to explain how the brain uses electrical and chemical signals to represent and process information, say three researchers involved in neural network modeling: biophysicist Terrence J. Sejnowski of Johns Hopkins University in Baltimore, computer scientist Christof Koch of the California Institute of Technology in Pasadena and philosopher Patricia S. Churchland of the University of California, San Diego. Although this goal is not new, they contend science is now in a better position to serve as a matchmaker between the computer hardware of neural networks, or "connectionist" models, and the three pounds of "wetware" encased in the human skull.

At the philosophical heart of network modeling lies the notion that the mind emerges from the brain's behavior. Thus, it makes sense to imitate in computer setups the structure and biological wiring of the brain to reproduce mental abilities.

The appeal of this approach, says Yale University psychologist Denise Dellarosa, "has its roots in an idea that will not die" — associationism. Put simply, associationism posits that humans learn through repetition to recognize people, things and events as more or less related to each other and as familiar or novel. Generalizing from examples, recognizing familiar faces in a crowd and driving a car are a few of the many tasks that characterize the effortless nature of associative learning.

Eighteenth-century philosophers David Hume and George Berkeley and psychologists in later centuries — including William James and B.F. Skinner — have championed, in their own ways, the cause of cognition as a building of associations through experience, Dellarosa says.

Neural networks attempt to simulate associative learning involved in vision, language processing, problem solving and motor control. Mathematical calculations adjust the strength of connections linking up "neuron-like" processing units. A given stimulus fed into the network activates all the units at the same time, including feedback mechanisms that stimulate or suppress designated connections. If the statistical assumptions guiding the connections are on target, a correct response is produced gradually after hundreds or thousands of trials.

One example of plugging neurobiology into a connectionist model was recently reported by Sejnowski and Hopkins colleague Sidney Lehky (SN: 3/5/88, p.149). Their neural network calculates curvature from shading in an image and behaves much as two types of neurons in the cat's visual cortex do. It relies on a procedure called back propagation. The system contains a layer of input units, a layer of output units and a layer of intermediate or "hidden" units that gradually acquire the right electrical responses — after several thousand trials — to accomplish the computational task. Error signals are sent back through the network as training proceeds to adjust connections between units and guide the system toward a correct response.

Turning this approach around, other researchers test computational approaches with data from brain studies. At last summer's International Conference on Neural Networks, held in San Diego by the Institute of Electrical and Electronics Engineers (IEEE), Bill Betts of the University of Southern California in Los Angeles reported that cells in the toad's visual center appear to operate in a manner modeled by the neural networks of Boston University's Stephen Grossberg. The neurons fire electrical impulses that activate prey-catching behavior only if given enough visual input to overcome the impulse-suppressing effects of another type of cells. Grossberg's model uses inhibitory mechanisms to establish a cutoff point that network activity must meet or exceed before information enters the system's memory; weak signals are suppressed and strong signals are enhanced.

Studies of small, related groups of brain cells in invertebrates further support the validity of neural networks, asserts biologist Eve Marder of Brandeis University in Waltham, Mass. For example, bony teeth that grind and shred food in the lobster's stomach are activated in different ways by a small group of neurons on the stomach surface, Marder reports in the Sept. 22 NATURE. Chemical and electrical properties of these neurons, input from other nerve cells and changes at synaptic connections combine to coordinate the rhythmic movement of the teeth.

Such evidence indicates that a small, related set of neurons can indeed orchestrate a variety of effects. Moreover, Marder adds, similar tasks can sometimes be performed by different neuronal arrangements in the same organism. The findings, she maintains, compare with neural networks that perform surprisingly complicated tasks by altering connections between processing units.

Some researchers, however, doubt that neural networks and neuroscience are a

match made in heaven. At the IEEE meeting, neurophysiologist Walter J. Freeman of the University of California, Berkeley, reiterated his argument that the brain's complexity eludes connectionist computers (SN: 1/23/88, p.58). "Brains rely on chaos to operate in the blooming, buzzing confusion of the environment, unlike neural networks," he said.

The low hum of background electrical activity in the brain reflects a "chaotic" process — in the mathematical sense — Freeman contends. What on first glance appears to be random noise is actually a flexible energy state from which massive numbers of neurons can be organized instantaneously to respond to new as well as familiar sensory information. Chaotic activity patterns have been observed in the olfactory and visual cortex of rabbits, he says.

Computer scientist Paul Smolensky of the University of Colorado in Boulder also questions whether there is — or will be — an intimate link between neuroscience and neural networks, but for reasons different from those voiced by Freeman. Mathematical considerations determine the ways in which people design connectionist machines, Smolensky maintains; the "loose correspondence" between neurons and processing units, as well as between synapses and network connections, will probably unravel as mathematical schemes to increase computing power become more sophisticated.

This aside, Smolensky suggests in the March BEHAVIORAL AND BRAIN SCIENCES that connectionist systems can serve as a bridge — one might even say a connection — between neuroscientific studies of brain cells and artificial intelligence (AI) investigations of language-based rules governing thought processes.

For the last 30 years, AI researchers have designed digital computer programs in which information is processed through operations on strings of arbitrary symbols. They hold that mental processes — memory, language use and production and problem solving, to name a few — are made up of a sequential series of formal rules often followed automatically. For instance, a speaker unconsciously interprets rules for language production and a scientist employs another set of rules when thinking about and gaining insight into a physics problem.

Connectionist systems may shed light on the mathematical rules followed by groups of neurons to generate the language-based rules of interest to AI researchers, Smolensky says. Neural networks engage in what he calls "statistical inference," a process more complicated than merely making associations between bits of information but less refined

than the so-called higher forms of mental function, such as logical reasoning.

Smolensky cautions that the media and some scientists have made outlandish claims about the potential of connectionism, although "it currently seems quite unknowable whether connectionist models can adequately solve the problems they face."

The same can be said of computational neuroscience, note Sejnowski and his colleagues. The field has yet to yield any successful large-scale theories of how cell circuits in the brain compute mental processes, although many investigators are optimistic such theories will emerge.

In contrast, some researchers see a dim future for neural networks, whether or not they contain a strong dose of biology.

"The connectionists have surely done something, but no one seems to be certain quite what," contends computer scientist Lawrence E. Hunter of Yale University. The neural networks described by Smolensky and by Sejnowski and his colleagues form an association between a stored piece of information and a new, similar pattern of information, Hunter says. But Hunter's definition of learning — the improvement of an organism's ability to achieve its goals on the basis of its experience — is a more complex business. In some cases, for instance, decisions are made to focus attention only on selected stimuli, which are then used to reevaluate goals.

At other times, learning occurs from a single experience, or novel explanations of a problem are suddenly generated. Connectionist networks are programming techniques for a limited type of memory, but cannot perform important learning tasks, Hunter holds.

Even if a neural network manages to produce intelligent behavior, argue other critics, it provides no understanding of the mind because its inner workings remain as inscrutable as those of the mind.

Philosopher Jerry A. Fodor of the City University of New York Graduate Center and psychologist Zenon W. Pylyshyn of the University of Western Ontario, in London, Ont., are among the most vociferous critics of connectionism. Most learning is a kind of theory construction, they write in Connections and Symbols (Pinker and Mehler, editors, MIT Press, 1988). Predictions about how the world works are made and evaluated against new experiences. Thus, the "statistical inference" of connectionist machines addresses, at best, a small part of mental functioning, they argue.

Fodor and Pylyshyn maintain there exists a "language of thought" — an argument first presented by Fodor more than a decade ago. In their view, thought

processes are made up of mental representations operating much as natural language does. Mental representations of new information and experiences are arranged according to specific rules that give them meaning and allow for the richness of thought.

In a simple example, the mental representation corresponding to the thought "John loves Fido" contains a series of interrelated concepts concerning each part of the thought, they say; relations between the concepts mark the difference between the thought "John loves Fido" and the thought "Fido loves John."

The tie between this type of thinking, which AI computers attempt to model, and the mind is more intimate than the tie between brain and mind, say Fodor and Pylyshyn. There is no reason to assume that higher mental functions, such as reasoning, correspond in any way to the structure of brain cells, they argue.

AI researchers Marvin Minsky and Seymour Papert of the Massachusetts Institute of Technology see a need to combine conventional digital computers with neural networks. "Maybe, since the brain is a hierarchy of systems, the best machine will be too," they write in Perceptrons (MIT Press, 1988). Such hybrid machines are indeed beginning to appear under the label of neurocomputers.

Minsky and Papert say neural networks are limited to solving "toy problems." A network model can, for instance, learn to recognize a particular cat, but it cannot use that experience to recognize cats in general.

They propose the brain is made up of many small neural networks, each of which performs a few simple, interrelated tasks. A serial system, much like an AI program, directs the activity of these small networks and puts the right ones together to create appropriate thoughts.

"I expect five to 10 major discoveries each year in neural networks," Minsky said at the IEEE meeting. But the lack of a general theory of brain function means "we still don't have a good way of characterizing what are good questions for neural networks to address."

Unfortunately, merely merging brain biology with computer models will not serve the right questions up on a silicon platter, remarks Walter Schneider of the University of Pittsburgh. Nor will it single-handedly get to the bottom of how people think.

"Neurophysiologists tell a story that if you can think of five ways that the brain can do something, it does it in all five, plus five you haven't thought of yet," Schneider says. "In the study of cognition we need to control our desire to have one answer, or one view, and work with multiple views." □