Are Neural Nets Like the Human Brain?

Not necessarily, though they are capable of astonishing feats. Only a close inspection of the brain will tell for sure

CAN NEURAL NETS TELL US how the brain works? Not necessarily, says Francis Crick of the Salk Institute, adding a hefty dose of skepticism to the euphoria surrounding this new field. While these sophisticated models are capable of astonishing feats and brainlike computation, says Crick in the 12 January issue of *Nature*, they should not be confused with the real thing.

"No one actually claims this is how the brain works," Crick told *Science*, "but there is a tacit assumption that it might be. There is a tendency to believe. Some people know perfectly well this is not how the brain works. Others know but tend to forget. And some don't realize how fully unbrainlike it is." Crick chides his colleagues for failing to look to the brain itself for answers.

Neural nets are computer models inspired by the brain's own hardware. They consist of processors, or "units," that share many of the properties of neurons. (Indeed, they were called neurons until Crick beat his colleagues into submission.) Each unit receives inputs, both excitatory and inhibitory, from a number of other units and, if the strength of the signal exceeds a given threshold, sends signals to other units.

As in the brain, the action, so to speak, is in the connections among these units. The secret of how the brain works will not be found in a single neuron, Crick points out, but in how groups of neurons interact. So, too, in these models, where the properties of the network arise from the overall pattern of the interactions among units. Each of the many connections or synapses among units has its own strength, or weight—essentially, a multiplier—that can be adjusted as the net tackles new tasks.

Not surprisingly, discussions of neural nets can easily bog down in the details of synaptic weights and their adjustment, but the underlying idea is that concepts and memories are embedded in the connections among neurons in the brain, and, by analogy, among the units in the model.

One goal of this work with neural nets is to learn how the brain works and eventually to understand more complex processes like learning and memory. The other goal is simply to use these nets as tools to carry out complex analytical tasks.

Perhaps their most striking attribute is that neural nets are capable of a form of learning. An example is NETtalk, a model that can learn to pronounce English. It was developed by Terrence Sejnowski, who just left Johns Hopkins University to join the Salk Institute, and Charles Rosenberg of Stanford University.

NETtalk receives English text, processes it, and then sends a message, or output, to a machine that makes sounds. At first, writes Crick, the machine babbles like a baby, but as training progresses it begins to speak more intelligibly. At the end, it can produce passable speech 90% of the time. What is surprising is that the network does not have rules of pronunciation wired into it but rather picks them up along the way by example, much as a child would.

"Some people know this is not how the brain works. Others know but tend to forget."

NETtalk is just one application of a recent algorithm, called back-prop, that is at the center of the current excitement about neural networks. Back-prop, short for back propagation of errors, refers to how the network corrects mistakes. It is known as supervised learning because there is a "teacher," in this case the computer modeler, who can tell the system what to do.

When the network makes a mistake—say, when an output unit sends the wrong sound to the machine in NETtalk—the "teacher" sends a signal to the output unit telling it to make a hard "c" rather than a soft "c." That correction message is then "back propagated" through the entire network. The net then makes a small adjustment in the weight of all the connections, so that the next time it receives the same input, the output is closer to the correct one. As this process is repeated, the error is minimized.

Similar back-prop models are being used to deduce the structure of a protein from its amino acid sequence, says Crick, and to distinguish shape from shading in images.

What is different in NETtalk and other recent nets is not just the error correction procedure but the three-layered architecture it is used in, which more closely approximates the structure of the brain than did the earlier nets. Early neural nets consisted of two layers of units, with a single set of connections between them. This new one has three layers: an input layer, an output layer, and a middle, or "hidden," layer.

"That is what all the excitement is about—the layer in between," says Sejnowski. "The brain has sensory receptors for input and motor neurons for output. But what makes life interesting is the interneurons in between, which allow animals to have complex behaviors." In neural nets as well, it is the hidden layer that holds the key to more complex computation.

All of which is very exciting, Crick says, but it is probably not how the brain works. The obvious problem with back-prop, says Crick, is that feedback messages would have to go down the axon the wrong way. And it implies the existence of a "teacher," which in the brain would presumably be another set of neurons. "Such a set of neurons, if they exist, should have novel properties and would be worth looking for, but there is no sign of back-prop advocates clamoring at the doors of neuroscientists, begging them to search for such neurons," says Crick, getting to the crux of his criticism.

"We just don't know how the brain works," says Crick, or if it uses something like back-prop. His major gripe is that neural net researchers seem surprisingly disinclined to find out.

Instead of looking to neuroscience, he says, modelers—many of whom he suspects of being "frustrated mathematicians"—seem all too intent on developing lofty general principles to explain human information processing. But nature may not work by grand principles, cautions Crick. "Evolution is a tinkerer. It is opportunistic; anything will do as long as it works." Instead of grand principles, it may prefer a series of slick tricks, says Crick. "Only a close inspection of the gadgetry will tell us."

Sejnowski is not so quick to dismiss either lofty principles or admittedly oversimplified models if they can shed some light on the brain, which he considers "one of the biggest challenges left in science. We need more frustrated mathematicians and more people who are testing ideas."

He points out that "neural nets are powerful tools for studying complex systems," applicable to many areas, from economics to designing new computers. And in those other applications, as Crick readily admits, Crick's criticisms do not apply.

"Even if the brain makes no use of back-prop, there is no reason why we should not use it as a mathematical tool," says Sejnowski. "Nature does not have calculus, but we use it to understand nature."

The goal of much neural net research is to understand complex human performance, such as how people learn to play the flute. "Even the simplest models are providing insights into how that learning occurs," says Sejnowski. "So even if back-prop has nothing to do with it in the brain, the fact that we can create models with similar performance means you can use it to understand psychology."

"This approach allows me to study aspects of the mind that can't be touched from the neuroscientist's approach," agrees David E. Rumelhart of Stanford University, a mathematical psychologist and one of the leading modelers in the field. The danger in that approach, says Crick, is that without a few reality checks in the brain, such investigations may yield exquisite theories that have no correlation with reality, like phlogiston, the early explanation of fire.

Rumelhart concedes that some neural net researchers may not be paying enough attention to neuroscience in developing their models. "Crick is nudging people like me, saying you can do better on the brain end. He is probably right."

Crick's closing plea is, "Why not look inside the brain, both to get new ideas and test existing ones?" Such work is getting under way, if belatedly. The delay was not just because modelers and neuroscientists have been wearing blinders, says Sejnowski, but because the tools simply did not exist to test some of these ideas experimentally. Sejnowski is setting up a computational neurobiology lab at the Salk Institute with just that goal: to test some of the theoretical predictions in the nervous system.

History gives some grounds for hope. In 1949 Donald Hebb predicted a synaptic mechanism that would explain learning and memory. Neuroscientists have recently discovered a mechanism—the NMDA receptor—that behaves just that way. In his article, Crick calls on neural net researchers to develop models that embody the principle of the NMDA receptor.

Sejnowski and his colleagues have just uncovered another synaptic mechanism, known as long-term depression, that also seems to be involved in memory. "These are two examples of abstract ideas being tested in parts of the brain. They are harbingers of the progress that can be made once models and the experimental work come together."

■ LESLIE ROBERTS

Inbreeding Costs Swamp Benefits

Inbreeding can have a variety of important genetic consequences, good and bad. For instance, a female that mates with a relative benefits because the offspring have additional copies of her genes (those she shares with the relative): and the closer the relationship, the greater the potential genetic benefit. However, the closest possible mating pairs—parent-offspring and sibling-sibling—are apparently rather rare in nature, an observation that is usually explained by the disadvantages of close matings. These disadvantages, collectively known as inbreeding depression, include reduced viability and fecundity of offspring. Biologists are therefore very interested in the effects of inbreeding, for genetic theory and in the practical consequences of maintaining captive populations. Until recently, however, there were very few good data on the effects of inbreeding in populations, but a study by Katherine Ralls and Jonathan Ballou of the National Zoological Park, Washington, D.C., and Alan Templeton of Washington University has now provided some.

Two striking observations resulted from a survey of juvenile survival from parent-offspring and sibling-sibling matings in 40 captive populations belonging to 38 species. First, in all but four of the populations there was some reduction in juvenile survival. Second, the variation in reduced survival was great, ranging from just a few percent to 100% in one case. The average reduction was 33%. This latter figure just happens to match the potential genetic advantage of parent-offspring and sibling-sibling matings, and so it might seem that the costs and benefits of such matings are finely balanced.

Ralls and her colleagues note, however, that the figure for costs is probably too low. For instance, it measures only juvenile viability, but does not include increased embryonic death or survival of juveniles to maturity; nor does it take into account reduced fecundity or increased susceptibility to disease, both of which are known consequences of inbreeding. These extra costs would swamp the 33% genetic advantages of close inbreeding, and thus account for its virtual absence in natural populations.

The disadvantages of inbreeding are assumed to result from the expression of deleterious recessive genes that occur in double doses. (In large, outbreeding populations these genes will usually occur only in single doses, and thus be masked.) If there is considerable variation among populations in the extent of such genes, then this could account for some of the variation in the severity of inbreeding depression observed by Ralls and her colleagues. A reduction in incidence of deleterious recessive genes can occur, for instance, if a population successfully goes through a bottleneck: the population crashes to a few individuals, genetic variance is greatly reduced, and deleterious recessives might be quickly lost through high mortality. There is considerable debate about the genetic effects of population bottlenecks, but it is clear that many populations would become extinct under such circumstances, an issue of particular interest to conservationists who must maintain small populations, either in natural habitats or in captivity.

■ ROGER LEWIN

ADDITIONAL READING

K. Ralls *et al.*, "Estimates of the cost of inbreeding in mammals," *Conserv. Biol.* 2, 185 (1988).

Roger Lewin

No inbreeding here. *Animals in natural populations avoid close genetic matings.*