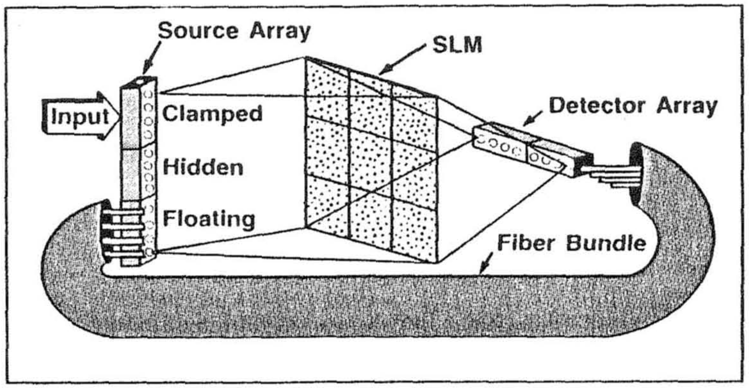
R. Colin Johnson, "Optics And Neural Nets: Marriage Of Convenience,"
Electronic Engineering Times, January 18. 1988.

Optics And Neural Nets: Marriage Of Convenience

By R. Colin Johnson

LOS ANGELES — Last week's O-E/LASE '88 conference here saw the marriage of optical computing and neural nets. The couple was hitched at the Neural Network Models for Optical Computing portion of the Jan. 10 to 15 show, sponsored by the Society of Photometric and Instrumentation Engineers.

Both technologies were given a boost from academia and industry. A team of University of Washington researchers demonstrated how their neural architecture could learn in a single pass what it takes others hundreds of passes to do. And a Lockheed research group described how to control mirror perturbation in sensitive "listening" instruments.



An optical alternating projection neural network accepts input to clamp some nodes, electronically generates a non-linear combination of the input at its hidden nodes and sends the result around an optical feedback loop through a spatial light modulator to its floating nodes.

Optical technology has for decades been a solution looking for problem. Though many simultaneous signals can be passed through any given node in an optical network without scrambling them, engineers have been hard pressed to capitalize on that characteristic.

The same problem exists in traditional computer technology. Tens, hundreds or even thousands of parallel processors have been fabricated on chips. The biggest trouble isn't with forging the hardware links for all these nodes but writing the software that harnesses them.

Optical technologies promise even more parallel processing, but what to do with it? If traditional parallel processors can't take full advantage of their nodes, then what profit is in having orders of magnitude more nodes?

Enter neural networks, whose main purpose is to simulate the manner in which billions of analog processing nodes (nerve cells or neurons) are connected in the brain. There, each node simultaneously evaluates the state of thousands or even tens of thousands of incoming messages from its neighbors. After processing, the node then sends on a single message to thousands of other nodes.

Optical technology is perfect for the massive number of connections needed for neural nets, since light beams can pass though each other without interacting. And they can be passed through the light-sensitive media separating each neural plane, which is where the strength of connections between neurons is stored. The network is usually "programmed" by altering the light-sensitive material separating the planes.

Cal Tech researcher Demetri Psaltis has demonstrated several neural-net prototypes over the last few years, most recently at the IEEE conference on Neural Information Processing Systems—Natural and Synthetic, held in Denver. In some of these systems a so-called volume hologram separated the planes and could be altered in real time by the actual flow of light-encoded information among its nodes. Such systems, when perfected, should be able to learn the tasks assigned to them by example, rather than depend upon explicit programming.

Most of the current systems take many presentations of a data set to learn it, since they are based on neural network architectures such as the back-propagation network. One paper at O-E/LASE, though, described a neural architecture for optical technology that took but a single presentation for any particular set of data to be learned.

It also was claimed to be very fast, since its passive optical feedback used only guided or free-space propagation. Other systems rely on the intervention of slow optical devices, such as phase conjugators, or even slower electronics.

The University of Washington professor Robert Marks II gave the presentation on his collaborative work with professor Les Atlas and assistants Seho Oh and Kwan Cheung. The architecture he described is called an alternating projection neural network (APNN). In it, a collection of nodes is divided into those whose states are fixed and those whose states are termed "floating."

The fixed-state nodes are ei-

(Continued on Page 42)

Optics, Marks' specialty, will be used to 'show' images to the computer and to manipulate the data internally. "At the front end of the computer, where you gather the data," Marks explains, "there might be an array of photo-detectors that would detect the image. Internal manipulation of the data that is conventionally done electronically would be done using light instead of electrons. It's obviously faster; you can't get much faster than light."

More than just a search for speed is involved in modeling the internal architecture of a neural network. The hundreds of electronic connections required between the neurons, using a conventional computer, would be impossible due to interference, but using photons rather than electrons eliminates that interference. The basic artificial neural network consists of many nodes or neurons that do very simple operations, and in some models, every neuron is connected to every other neuron. Using conventional connections would require the impossible: electrons going through electrons. Marks describes the advantage of using optics: "If you do it optically, photons can go through photons. Light can go through itself, so using light gives you the nice ability to have the natural physics for intense interconnections of the nodes or neurons."

One technology available with the neural network is parallel rather than serial processing. "One neuron doesn't have to wait for what another neuron does; they all kind of do their own thing and come out with a really neat answer."

Reaching "a really neat answer" in neural network parlance is called converging, and Atlas and Marks' APNN outperforms previous thermodynamic models of neural networks in accomplishing convergence efficiently and consistently. The thermodynamic models use an energy reduction approach which Marks says, "doesn't prove uniqueness of convergence, that is, one time the neural network converges to one thing, and another time it converges to something else. So in that sense it's a relatively poor model." Marks elaborates, "Our model of the APNN draws upon a wealth of mathematical theory, including projection onto convex sets, which is a recent field of interest and analysis from which we've been able to borrow."

Besides convergence, the ability of a neural network to generalize is a requirement of any efficient classification network. Marks describes generalization between the two modeling systems, "It's easy to train a classifier to respond to training data. What's important, however, is how it responds to new data. Can it recognize a totally new bush?" A disadvantage of the conventional neural network is that determining how it will respond can only be done empirically. "You actually have to expose it to the new material and see if it responds correctly. However, with the APNN, the math is so well developed that we can predict the manner in which the network generalizes, and we can write down math equations that show whether and in what manner the network generalizes to other than the training data."

The ability to generalize to new data or environments is a problem that conventional computers respond to poorly. Even the recent developments in artificial intelligence, such as expert systems, have this problem. "Neural networks offer the theoretical potential to control and design the specifics of generalization," according to Atlas. "However large amounts of data from many real-world environments are needed to test and refine this theory."

Training a network by example requires incredible amounts of time to pass through the data, and the problem with conventional neural nets is that they can forget the earliest data by the time they are exposed to the final data. This forgetting requires repetitive passes through the training data. However, repetitive passes are not required for the APNN, because it has an elephant-quality memory. It never forgets. A single pass through the training data is sufficient.

Improved memory within the actual computer architecture is another advantage of the APNN. The associative memory capability of the artificial neural network could allow the APNN to identify a black and white picture (similar to a digitized picture) of the Mona Lisa, given only her smile. "We have a matrix of neurons," explains Marks, "that can take on gray levels. In this matrix every neuron is connected to every other neuron, and each neuron can assume a value that relates to a gray level. So, having been given a picture of the Mona Lisa, the gray levels of that picture are imposed on the neurons and the information is stored in the interconnects," (these interconnects correspond to the synapses that connect the neurons in the biological brain) "and remarkably, if the network is then given only the Mona Lisa's smile, the APNN could then extrapolate the entire face of the Mona Lisa."

The future of the APNN, is being extended to some real world applications: A *speaker-independent* system of speech recognition is being developed by Atlas and his team of graduate students. Using a large data base containing many words from many speakers, the team plans to have a demonstration system ready in two years. In order to make the system commercially acceptable, it is necessary to keep the rate of recognition errors to a minimum. It is also essential that the remaining errors be as "natural" as possible. "Human voice interaction is not error-free either," Atlas explains. "A key problem with conventional recognizers is that their errors are not at all like natural human errors. We feel that the APNN has the potential to behave as a human does, which would include the errors that naturally occur in human speech recognition." Other applications of the APNN include efficient routing of computer links and an automatic system to identify irregularities in electrocardiograms (EKG's).

Funding for Atlas and Marks' APNN comes from a variety of sources: The National Science Foundation, The Office of Naval Research, Physio Control Corp. and the Washington Technology Center. Although a considerable amount of research remains to be done, based on the available funding and the incredibly high level of interest in the field, Marks and Atlas are optimistic that neural network computers will be commercially available in the near future.

Optics And Neural Nets: Marriage Of Convenience

(Continued from Page 41)

ther the input nodes or hidden nodes that are set according to a non-linear function of the input. While the hidden nodes have no effect on the learning of the network, they have a profound affect on the ability of the network to generalize.

Unlike traditional semiconductor memories that can recall only that with which they have been programmed, neural networks often possess the peculiar ability to generalize from their data set and come up with accurate responses to queries with which they have not been specifically trained.

For instance: Suppose an associative neural network is taught the cosine of each whole degree angle from 1 degree to 360 degrees. A traditional memory would not know how to respond to a request for 28½ degrees, but a neural network would generalize on its knowledge and come up with a reasonable answer. It turns out the number of specific data packets (vectors in this case) that can be stored in the APNN is on the order of the number of fixed nodes it contains.

The floating nodes have the most interesting behavior. They take on a value that is the sum of their inputs from the other nodes. The inputs to each node are multiplied by a value stored in a passive, planar spatial-light modulator of the kind developed at Stanford University (Palo Alto, Calif.) in the late 1970s. By providing feedback with fiber optics, a loop can be formed from the floating nodes into the spatial light modulator and then back into the floating nodes. This feedback loop converges on the "answer," which is then read by other devices.

Also at the conference, researcher Robert Smithson (of Lockheed Missile and Space Corp.) used neural networks to control mirror perturbation for sensitive listening instruments.

Lockheed allocated over $330,000 in 1987 toward developing an analog neural network, largely under Smithson's guidance (see Dec. 14, Page 51). The result was an LSI programmable-interconnection chip fabricated by Siliconix Inc. (Palo Alto, Calif.). It will be used by Lockheed to build feedback-style neural networks such as the energy-minimization nets originated by professor John Hopfield at Cal Tech. The chip is basically a crossbar switch with adjustable resistor values, called weights, at each connection. Smithson's segmented active mirror for solar observations demonstrated that neural networks can be used for real-time control. Since light beams are deformed by turbulence in the atmosphere, a neural network can be used to earn about those deformations and compensate by controlling the mirror. Smithson offered a tutorial on his techniques at the conference. His paper addressed the general area of applying neural network concepts to adaptive control.

In active mirror-control applications, Smithson's team has built both feed-forward and feedback prototypes. The feedback networks, of the Hopfield type, have also been developed for target classification. Such energy-minimizing feedback networks may produce the first workable neural network applications, especially for adaptive control systems. But Smithson cautions that the applications in which feedback works best is when the system is asked to make small perturbations from a known solution, as when interpreting signals that have been slightly altered by atmospheric conditions.

Smithson's project for feed-forward active mirror control incorporated learning capabilities. By adapting to changing atmospheric conditions, it used self-programming for different mirrors and wave-front sensors. Analog hardware operating at 10 kHz to 100 kHz should be relatively easy to build. The main restriction on network is the lack of architectural definition.

Currently, Smithson is studying the convergence and stability criteria to make the circuits more reliable. That involves looking in detail at the energy surfaces produced and the circuit dynamics.