

Neuroinformatics as Explanatory Neuroscience

RODNEY DOUGLAS, MISHA MAHOWALD, AND KEVAN MARTIN

WHAT IS NEUROINFORMATICS?

There are two points of view about the meaning of neuroinformatics, which we may write neuro-Informatics and Neuro-informatics, to reflect their different emphases. The proponents of neuro-Informatics hold that it is the application of conventional informatics to the domain of neuroscience. By contrast, the proponents of Neuro-informatics hold that it studies information processing by nervous systems. There is a very significant conceptual difference between these two views, which arise very naturally out of two contrary views of science.

The first view, neuro-Informatics, arises out of the philosophy that science is description, and so the major task in modern science is to accumulate and catalogue data. Thus, neuro-Informaticians look to informatics as a maturing information technology based on general purpose computing principles. For the proponents of this goal, informatics is a tool to aid neuroscience. The aid it gives is to catalogue and manipulate neuroscientific data. The hidden assumption is that scientific data are absolute, and that once we have enough data, we will inevitably be able to answer the hard questions.

The second view, Neuro-informatics, arises out of the philosophy that science is explanation, and so the major task is to extract predictive principles. Neuro-informaticians take the view that nervous systems are probably qualitatively different from the general purpose computing principles that have dominated the past few decades. Reasons for anticipating these differences are not hard to find. Indeed, many of them were pointed out by von Neumann, the very inventor of general purpose computers. In the view of Neuro-informaticians, resources should be focused on the substantive problem of neuroscience: What is the nature of computation in biological nervous systems?

Our research, at the Institute of Neuroinformatics in Zurich, follows the latter point of view. More specifically, we aim to cast the neural computational processes in an electronic medium, using analog very large scale integration (aVLSI) technology. Carver Mead (1989) introduced the term *neuromorphic engineering* for this new approach based on the design and fabrication of artificial neural systems, such as vision systems, head-eye systems, and roving robots, whose architecture and design principles are based on those of biological

nervous systems (Douglas *et al.*, 1995). Neuromorphic systems try to emulate the organization and function of biological nervous systems—they are a method of exploring the principles of neural computation from the vantage points of both neuroscience on the one hand and engineering and computer science on the other. Implicit in neuromorphic engineering is the hypothesis that neural computation may be qualitatively different from classical computers and computation.

The enormous success of digital technology and general purpose computers in performing abstract tasks bred confidence that neural computation could be simply captured by those tools. In fact, general purpose computers have been quite unsuccessful in performing autonomously tasks that require any degree of sophisticated sensorimotor interaction with the real world. Even rather primitive biological nervous systems are able to extract meaningful information from a noisy world in real time, but artificial systems still lag far behind such performance.

THE BIOLOGICAL ROOT

Our particular research goal has been to explore the principles of biological computation as they present themselves in primary visual cortex. For over a decade Kevan Martin and I have focused on the detailed structure and function of microcircuits in visual cortex. We think that the cortex is like the retina, in that the anatomical circuitry is configured in the *z* (or thickness) dimension to perform some specific signal processing operations, but that essentially the same operations are then tiled across the two-dimensional plane of cortex. Therefore we are focusing on the inter- and intralaminar connections within a small volume of cortex (say, 2 mm on a side). We do this by a combination of extracellular and intracellular recording, intracellular labeling, tract-tracing techniques, and three-dimensional reconstruction of neurons at electron microscopy and light microscopy level.

We have used these data as a basis for exploring by simulation methods the behavior of single cortical neurons, small circuits of neurons. This aspect of our work has been done largely in collaboration with Christof Koch at Caltech. In those projects we have used a variety of simulation and mathematical software ex-

ecuted on standard good performance workstations. During this phase of our work it quickly became clear that, while simulation was a useful tool for examining particular quantitative questions about neuronal function, it was much less useful for reaching explanatory principles because we were in danger of replacing the complexity of the nervous system with another complexity, that of the detailed computer simulations. The durations of the simulations were approaching those of the original experiments, i.e., months.

THE ELECTRONIC ROOT

At the time that we were exploring the neocortical circuitry, Carver Mead and Misha Mahowald, also at Caltech, were building neuromorphic aVLSI retinas. One intriguing feature of their approach is that the aVLSI circuits are not built by exact description. Instead, one needs to understand roughly the nature of the operation occurring between nodes. For example, knowing that one node has an exponential dependence on another is crucial—but knowing the exact parameter values for the relationship is much less relevant because they cannot be accurately designed or set. Moreover, it was clear that to build ultralarge systems in practice, it would be necessary to find methods that were simple, were fault tolerant, and required only low precision. Their view resonated with what we were learning about the biological systems. Despite the careful statistics quoted in the literature, the large variations in performance of neurons are as compelling as their similarities. It seems that the cortex is not designed in the highly specified style of a Pentium, but rather depends on local adaptive mechanisms to draw its components into calibration with respect to one another.

SILICON RETINA

The silicon retina is a good example of a neuromorphic analog VLSI. Sensors are the nervous system's starting point for interaction with the world. If these source data are incomplete, or incorrectly transmitted to later processing stages, the errors cannot be recovered, and the animal will act inefficiently on false premises. Classical machine vision processing postpones the processing problem, collecting and transmitting as much low-level source data as possible. This strategy makes strong demands on the dynamic range of the sensors and the communication bandwidth of the transmission channel to the later processing stages. This is particularly true of vision because natural visual scenes usually have dynamic ranges of three orders of magnitude during steady illumination, and the dynamic range of the same scene viewed from sunny afternoon to dusk may vary by seven orders of

magnitude. These ranges exceed the limits of current digital imaging technology.

Biological systems have evolved an alternative approach to sensing. Rather than transmitting all possible information to later stages of processing, they transmit only that which is salient. By computing high-order invariants close to the sensors the bandwidth of communication to subsequent processors is reduced to a minimum.

The degree of invariance achieved by early sensory processing is a compromise that is made differently by various modalities and species. This compromise is forced by considerations of space and processing strategy. The space available for additional circuitry restricts the sophistication of signal processing that can be accomplished near to a planar array of sensors, such as the retina. Invariance of response implies that sensory information has been discarded. The question of what to discard cannot be context sensitive if invariants are hardwired into early sensory processing.

The silicon emulates the outer plexiform layer of the retina of the mudpuppy (Mahowald and Mead, 1989; Mead and Mahowald, 1988). Its components represent the photoreceptor, horizontal cell, and bipolar cell layers of the retina. The photoreceptor transduces light into a voltage that is logarithmic in the intensity of the stimulus. The synaptic interactions between the cell types are implemented in analog circuits. The photoreceptors drive the horizontal cells (a noninverting synapse) via a transconductance. The gap junction connections between horizontal cells are emulated by a resistive network. Thus, the voltage at each node of the horizontal cell network represents a spatially weighted average of the photoreceptor inputs to the network. As in the biological retina, the electrotonic properties give rise to an exponentially decreasing spatial receptive field in the horizontal cell network. The antagonistic center-surround receptive field of the bipolar cell is implemented with a differential amplifier that is driven positively by the photoreceptors, but inhibited by the horizontal cell output.

Unlike CCD cameras, the silicon retina reports contrast rather than absolute brightness and so is able to see comparable detail in shaded and bright areas of the same scene. This contrast encoding is an essential property of retinas that arises out of the center-surround receptive field of the bipolar cell, which computes the difference of the logarithmic photoreceptor and horizontal cell outputs. By using the local average of the horizontal cells as a reference signal, the redundant features of the image are suppressed while novel features are enhanced. This occurs because large areas of uniform luminance produce only weak visual signals. In these regions the output from any single photoreceptor is canceled by the spatial average signal from the horizontal cell network. Novel luminance

features, like edges, evoke strong retinal signals because the receptors on either side of the edge receive significantly different luminance. A similar principle applies in the time domain, where the relatively slow temporal response of the horizontal cell network enhances the visual system's response to changing images.

The functional interpretation of processing by the outer plexiform layer is that it prevents the output of the retina from saturating over a change of several orders of magnitude in background illumination, while allowing the retina to report reliably the small contrast differences in a typical scene. The encoded retinal output still contains sufficient information to support fully general visual processing in the brain, but it can be transmitted at a much lower bandwidth than can a full gray-scale image.

It is lateral inhibition that removes redundant information caused by correlation of luminance across the image (Barlow, 1981). However, lateral inhibition exacerbates another form of image redundancy, noise. Noise is a form of redundancy because it does not supply additional information about the world, so bandwidth is wasted by transmitting it. In the retina, noise arises from photon fluctuations at low light levels. But, it can also arise from the photoreceptors themselves, as dynamic thermal noise and as static miscalibration of photoreceptors' gains and operating points.

Similar to the biological case, in the aVLSI circuits, static mismatches between transistors was an intractable problem in early silicon retinas. These mismatches cause the photoreceptors to give a different response to identical inputs, and the mismatches could be incorrectly interpreted by postretinal processing as properties of the scene. Later generations of silicon retinas suppress these intrinsic sources of noise by using adaptive elements and resistive coupling between the receptors (Boahen and Andreou, 1992; Dowlings, 1987). For example, the new retinas have feedback from the horizontal cells to the photoreceptors (Skrzypek, 1991; Yau and Baylor, 1989). Mahowald (1992) proposed that this feedback plays a role in dark adaptation by shifting the operating point of the high-gain retina to match the average ambient light level, thus allowing the photoreceptor to respond over a large input range with higher gain than in solely feed-forward retinas. This adaptive mechanism has been incorporated successfully into silicon retinas (Mead, 1989; Mahowald, 1991).

FUTURE CHALLENGES

We have learned much from the construction of the silicon retina. We are now extending these concepts to the more complicated cortical problems. Unlike retina, the connectivity of cortex is not so dominated by

nearest neighbor connections. So, to construct neuro-morphic cortical circuits, we had first to develop a general purpose spiking neuron in aVLSI (Mahowald and Douglas, 1991) and a method of assembling large numbers of such neurons into neuronal networks. On the biological side, we are quantifying the connectivity of neuronal circuits in visual cortex (Ahmed *et al.*, 1994), and exploring the computational processes that they support (Douglas *et al.*, 1995).

Fabrication of large cortical networks has forced us to confront once again the problems first appreciated by von Neumann (1958), at the time that he was considering the design of the first electronic computers. He was faced with memory and computational resources that were small compared to the size of the computational problems that needed to be solved by those machines. Von Neumann solved this problem by choosing an architecture based on serial processing. His strategy was to store the problem in a way that permitted components to be extracted serially and assembled transiently in the working memory of the computer. The limited computational resources could then act on this small chunk of the problem, and store intermediate results, which would interface with the next chunk of the problem to be assembled, and so on. This style of computation not only implies a particular physical organization, but also a particular method of encoding the problem. The encoding must be general, so that the many different aspects of the problem can be encoded on the same physical components.

Von Neumann was quick to point out that his solution was not the style of computation in the brain. In the brain all parts of the computation are present simultaneously and compute interactively in a fine-grained parallelism. Unlike general purpose computers, biological computation is profoundly distributed. Every part of the nervous system combines computation directly with the memory elements required to support it—computation and memory are colocalized. Moreover, the result of the computation at any location in the biological computer seems not to be of a generally encoded form. Instead it is essentially symbolic—the signal expresses its meaning in terms of the specific physical operations that the circuit performs. We think that the understanding of key conceptual issues such as these constitutes the critical challenge of neuroinformatics.

REFERENCES

- Ahmed, B., Anderson, J. C., Douglas, R. J., Martin, K. A. C., and Nelson, C. 1994. Polynuclear innervation of spiny stellate neurons in cat visual cortex. *J. Comp. Neurol.* **341**:39–49.
- Atick, J., and Redlich, A. 1990. Towards a theory of early visual processing. *Neural Comput.* **2**:308–320.
- Atick, J., and Redlich, A. 1992. What does the retina know about natural scenes? *Neural Comput.* **4**:196–210.

- Barlow, H. B. 1981. Critical factors limiting the design of the eye and visual cortex: The Ferrier lecture. *Proc. R. Soc. London B* **212**:1–34.
- Boahen, K., and Andreou, A. 1992. A contrast sensitive silicon retina with reciprocal synapses. In *Advances in Neural Information Processing Systems*, Vol. 4, pp. 764–772. Morgan Kaufmann, San Mateo, CA.
- Douglas, R., Mahowald, M., and Mead, C. 1995. Neuromorphic analog VLSI. *Annu. Rev. Neurosci.* **8**:225–281.
- Douglas, R. J., Koch, C., Mahowald, M. A., Martin, K. A. C., and Suarez, H. 1995. Recurrent excitation in neocortical circuits. *Science* **269**:981–985.
- Dowling, J. 1987. *The Retina: An Approachable Part of the Brain*. Harvard Univ. Press, Cambridge, MA.
- Mahowald, M. 1992. VLSI analogs of neuronal visual processing: A synthesis of form and function. Ph.D. thesis, Department of Computation and Neural Systems, California Institute of Technology, Pasadena.
- Mahowald, M. A. 1991. Silicon retina with adaptive photoreceptors. In *Proceedings of SPIE International Symposium on Optical Engineering and Photonics in Aerospace Sensing*, Orlando, FL, April.
- Mahowald, M., and Douglas, R. 1991. A silicon neuron. *Nature* **354**:515–518.
- Mahowald, M., and Mead, C. 1989. Silicon retina. In *Analog VLSI and Neural Systems* (C. Mead, Ed.), pp. 257–278. Addison–Wesley, Reading, MA.
- Mead, C. 1989. *Analog VLSI and Neural Systems*. Addison–Wesley, Reading, MA.
- Mead, C. 1990. Neuromorphic electronic systems. *Proc. IEEE* **78**:1629–1636.
- Mead, C. A., and Mahowald, M. A. 1988. A silicon model of early visual processing. *Neural Networks* **1**:91–97.
- Skrzypek, J. 1991. Light sensitivity in cones is affected by the feedback from horizontal cells. Technical report, Department of Computer Science, University of California, Los Angeles.
- von Neumann, J. 1958. *The Computer and the Brain*. Yale Univ. Press, New Haven, CT.
- Yau, K., and Baylor, D. 1989. Cyclic GMP-activated conductance of retinal photoreceptor cells. *Annu. Rev. Neurosci.* **12**:289–327.