Annals Hist Comput (1989) 11:197–201

The Computer and the Brain Revisited

TERRENCE J. SEJNOWSKI

Terrence Sejnowski assesses von Neumann's contribution of mathematical and computational tools for the development of computational neuroscience. He surveys the progress that has been made in this field since von Neumann's death and outlines the difficulties that remain.

Categories and Subject Descriptors: K.2 **[Computer Milieux]:** *History of Computing—hardware, people, systems, theory. H.1.2* **[Information Systems]:** *Models and Principles—Systems and Information Theory; H.1.2* **[Information Systems]:** *User/Machine Systems—Human information processing.*

I first read John von Neumann's book *The Computer and the Brain* in the summer of 1970, while studying for the general examination for doctoral candidacy in physics at Princeton. Ever since then, I have been thinking about the issues von Neumann raised in his book. Rereading the book recently has highlighted the progress that has been made on trying to understand information-processing in the brain, as well as the difficulties that remain.

When von Neumann wrote the manuscript for the Silliman Lectures at Yale in 1956, the general character of electrical transmission and communication between neurons had just recently been elucidated through the seminal work of Alan Hodgkin and Andrew Huxley on the squid giant axon, and Bernard Katz on the frog neuromuscular junction. The all-or-none nature of the action potential had suggested analogies with binary gates in digital computers (McCulloch and Pitts 1943), but the analog nature of neural integration was just beginning to be fully appreciated. Typically, the accuracy of numerical calculation in a modern digital computer is 8 to 16 significant figures. But in a neuron, signaling by means of the average firing rate has at best one or two significant figures of accuracy. We still do not understand how information is represented in the brain in such a way that low accuracy is not a problem.

Von Neumann recognized that the reliance of the brain on analog-signal processing had far-reaching significance for the style of computation that the brain was capable of supporting. He pointed out that the logical depth of a calculation, for example, can be very great for a digital computer that retains high accuracy at each step in the calculation; but for an analog system like the brain, the compounding of errors causes severe problems after only a few steps. Much of the work in artificial intelligence depends on the efficient use of a sequential, symbol-processing architecture, and on tree searches that have great logical depth. The model of computation based on logic that led to sequential architecture also served as a model for human reasoning in cognitive science (Newell and Simon 1976). The recent availability of digital computers with parallel architectures makes apparent the extent to which cognitive science and artificial intelligence have been shaped by hardware that is based on sequential symbol processing.

It is somewhat ironic that the sequential architecture of digital computers is generally called von Neumann architecture. Von Neumann was well aware of the need for a broader science of computation and contributed not just to the development of sequential architecture but also to the foundations of cellular automata (von Neumann 1963)—an early parallel architecture that has only recently been exploited (Wolfram 1983). It is also apparent from *The Computer and the Brain* that von Neumann was skeptical about sequential architecture as a model for how the brain works.

The use of memory in digital computers to store both sequences of instructions and data was a breakthrough to which von Neumann made major contributions. Memory is one of the central themes in his book. Very little was known at the time concerning memory mechanisms at the cellular level or the locations where memories are stored in the brain. "We are as ignorant," he stated, "of its nature and position as were the Greeks, who suspected the location of the mind in the diaphragm." Today we have a much better knowledge of the condition for neural plasticity in many different areas of the brain. In the hippocampus, for example, a form of plasticity called long-term potentiation has been found that results in changes that can last for many days. Moreover, the plasticity in some parts of the hippocampus seems to depend on a learning rule first proposed by Hebb (1949), and molecular mechanisms are being identified that control this plasticity (Brown et al. 1989). However, the organization of knowledge in networks of neurons is still a mystery (Sejnowski and Tesauro 1989).

Information in digital computers is stored at locations that can be individually addressed. In humans, information is organized in a complex web of structured associations, so that memory is accessed through content. Von Neumann calculated an upper bound for how much information could be stored in the brain. Assuming that a trace of all information impinging on our sensory receptors is stored, he arrived at a capacity of about 10^{20} bits, which is probably a vast overestimate of what we actually retain. A lower bound on memory capacity, based on how much information can be consciously recalled, is 10^9 bits (Landauer 1986). This is, however, only one type of memory, the declarative memory that we have for facts and events (Squire 1987). Other memory systems, such as motor learning and procedural knowledge, appear to be organized separately from the memory system for facts. It is more difficult to quantify the capacity of nondeclarative memory sytems because they are not easy to dissect into components. One approach to this problem is to estimate the total amount of information that can be stored at synapses, the specialized contacts between neurons that are used for signaling.

A synapse is activated when an action potential—an impulse emanating from a neuron—invades a synaptic terminal, a chemical neurotransmitter is released, and a signal is communicated to the postsynaptic cell. The electrical response produced in the postsynaptic cell can be excitatory or inhibitory; that is, it can bring the cell either closer to or farther away from the threshold for initiating an action potential. Thousands of synapses converge on a single neuron. The process of integrating information, of collecting signals from thousands of sources, involves filtering processes and nonlinear decisions that we are just beginning to understand at the molecular level. Von Neumann made many numerical estimates of memory capacity, neural power dissipation, and time scales in his book, but, curiously, he did not estimate the number of synapses in the brain. Since it would be physically impossible to count every synapse in the brain, estimates are based on sampling techniques. When I first looked in the literature for this number in 1970, I found estimates of around 10^{13} synapses, but this now appears to be an underestimate. The best current estimate is 10^{14} synapses, and the number may continue to change as automated anatomical methods continue to improve. If we assume that synapses are sites of information storage, then we can make a rough estimate for the total information stored in the brain. A single synapse can store only a few bits of information in the form of a coupling strength. Thus, a rough estimate for the information stored in our brain is around 10^{14} bits. Compared to the estimate of 10^9 bits of information that is consciously available to us, our nondeclarative memory could be larger by as much as 10^5.

One of the most difficult problems facing us is that of trying to understand thought processes through understanding the brain (Sejnowski and Churchland 1989). Here too, von Neumann stated a very clear position: "Thus, logics and mathematics in the central nervous system, when viewed as languages, must structurally be essentially different from those languages to which our common experience refers." This statement reflects a lifetime of thinking about computation and mathematics. It was perhaps his last thought on this issue and was in many ways deeply prophetic of current research on brain models as the substrate for cognition (Churchland 1986; Rumelhart and McClelland 1986). We are still very far from knowing what the hidden language of the central nervous system might be. Given what we now know about the structure of the brain, is it possible to begin to study the hidden language of brain systems?

A new field is developing called computational neuroscience; its goal is to understand how the brain represents and processes information (Sejnowski et al. 1988). One of the principal techniques used is the modeling of brain function at many different levels of investigation, from the molecular to the systems levels. The digital computer provides for the first time, enough computing power to explore the complexity of the brain by simulating massively parallel models of brain structures at the level of neurons and synapses. At present, it is possible to model only small parts of the brain system; however, computing power is rapidly increasing, and we may eventually learn some of the design principles of the brain.

In 1970, inspired by von Neumann's numerical estimates of brain capacities, I attempted to estimate the computing power needed to simulate a brain. Without knowing how the brain is designed, it is not possible to provide a definitive answer; however, a rough lower bound is possible, and here, very briefly, is the answer I found. The first thing I did was to define an elementary operation in the brain as a single synaptic event. In a digital computer an elementary operation constitutes a set of instructions corresponding to the simulated update at a synapse (typically, a few instructions such as load, add, and multiply). In Figure 1, the logarithm of the number of elementary operations per second that the largest existing digital computer could accomplish is plotted as a function of time. In the 1950s, the earliest digital computers based on vacuum tubes could do about 10,000 operations per second. In 1970, when I made my first estimate, the largest computer available was a computer that could perform about one million operations per second. Given that there are on the order of 10^{14} synapses in the brain, and that they are being activated on the average of about 10 per second, one can estimate that the brain is performing, at minimum, on the order of 10^{15} operations per second. This is a lower bound, because we know that very sophisticated analog processing occurs within a neuron. If we optimistically extrapolate the straight line in Figure 1, which is an empirical fit to the data, we find that it crosses the minimum processing capabilities of the human brain sometime around 2010. A more thorough analysis, which takes into account other important factors, such as the communication bandwidth, leads to the same conclusion (Waltz 1988).

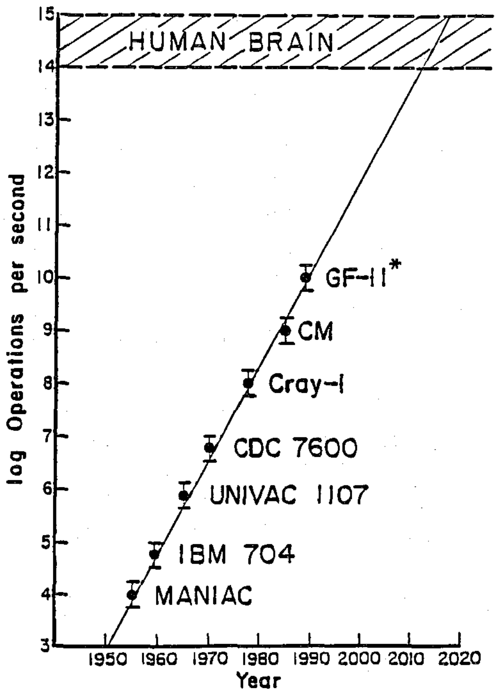

Figure 1. The logarithm of the number of elementary operations per second by the largest digital computer plotted as a function of time.

The straight line in Figure 1 represents an exponential increase in processing power that is a cascade of many different technologies, starting from the vacuum tube, to discrete transistors, to integrated circuits. The most recent generation of computers, such as the Connection Machine, are highly parallel architectures. Although my extrapolation seemed somewhat dubious in the 1970s, and a bit discouraging because so much more computing power was needed, the processing power has continued to march up the curve. It is now halfway between 1970 and 2010 and computers exist that can process a billion operations per second. However, there is still no guarantee that we will continue to move up the curve. New technology is now being developed that may make it possible to achieve rates of computation that we only dream about today. For example, analog very large scale integration (VLSI) technology will make an enormous difference in our ability to process sensory information in parallel. The accuracy and dynamic range of analog processing in silicon is low, as it is in the brain, but the speedup from massively parallel processing in wafer-scale integration has the potential to produce spectacular results (Mead 1988).

It is apparent in Figure 1 that we are also far from performing simulations on a scale large enough to test our ideas of how the brain overcomes problems of low accuracy and shallow logical depth. Some progress has already been made using simplified models of neural networks which can be used to explore issues, such as how information about single items or relationships can be represented in a distributed fashion over many synapses and neurons (Hinton and Anderson 1981; Rumelhart and McClelland 1986; Hopfield and Tank 1986). These models demonstrate that difficult problems can be solved with relatively simple network architectures, and that the performance of a network model is surprisingly immune to damage and noise. However, they do not prove that the brain solves these problems in the same way; the next step is to scale up these simulations to more complex networks with more realistic assumptions. It is very likely that we are still missing a crucial insight into the information codes used by the brain to represent abstractions and symbolic relationships. There are already hints that spatially and temporally coherent bursts of action potentials at high frequencies may carry such a code (Brown et al. 1989; Crick 1984; von der Malsburg 1987; Sejnowski 1976). Von Neumann called this the problem of the "short code"; this is the problem of finding a representation in the brain sufficiently powerful to allow the brain to imitate the behavior of another computing system.

We owe much to von Neumann for the mathematical and computational tools that we now bring to bear on problems of the mind and brain. The next few decades should prove to be very exciting ones for computational neuroscience. By 2010 we should have enough computing power to simulate large brain systems, and many of the problems raised by von Neumann in *The Computer and the Brain* should become amenable to experimental investigation and modeling studies. It is difficult to predict what methods for studying complex systems may be needed before an understanding of cognition is achieved, and what information will be found with these methods. The traditional analytic techniques in mathematics may not be sufficiently powerful for the exploration of nonlinear brain models. It may even be possible that all existing traditional techniques based on the manipulation of symbols will be inadequate. The last words in von Neumann's book will also be ours: "However, the above remarks about reliability and logical and arithmetical depth prove that whatever the system is, it cannot fail to differ considerably from what we consciously and explicitly consider mathematics."

Acknowledgments

I am grateful for many stimulating conversations with Drs. Patricia Churchland, Francis Crick, and Christof Koch.

References

Brown, T. H., A. H. Ganong, E. W. Kairiss, C. L. Keenan, and S. R. Kelso. 1989. "Long-term Potentiation in Two Synaptic Systems of the Hippocampal Brain Slice." In J. H. Bryne and W. O. Berry (eds.), *Neural Models of Plasticity,* New York, Academic Press, in press.

Churchland, P. S. 1986. *Neurophilosophy: Toward a Unified Science of the Mind/Brain,* Cambridge, Massachusetts, The MIT Press.

Crick, F. 1984. "Function of the Thalamic Reticular Complex: The Searchlight Hypothesis." *Proceedings of the National Academy of Sciences USA* 81, pp. 4586–4590.

Hebb, D. O. 1949. *Organization of Behavior,* New York, John Wiley & Sons.

Hinton, G. E., and J. A. Anderson. 1981. *Parallel Models of Associative Memory,* Hillsdale, New Jersey, Lawrence Erlbaum Associates.

Hopfield, J. J., and D. W. Tank. 1986. "Computing with Neural Circuits: A Model." *Science* 233, pp. 625–633.

Landauer, T. K. 1986. "How Much Do People Remember? Some Estimates of the Quantity of Learned Information in Long-Term Memory." *Cognitive Science* 10, pp. 477–493.

McCulloch, W. S., and W. Pitts. 1943. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics* 5, pp. 115–133.

Mead, C. 1988. *Analog VLSI and Neural Systems,* Reading, Massachusetts, Addison-Wesley.

Newell, A., and H. A. Simon. 1976. "Computer Science as Empirical Enquiry." *CACM* 19, pp. 113–126.

Rumelhart, D. E., and J. L. McClelland. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* Cambridge, Massachusetts, The MIT Press.

Sejnowski, T. J. 1976. "On Global Properties of Neuronal Interaction." *Biological Cybernetics* 22, pp. 85–95.

Sejnowski, T. J., C. Koch, and P. S. Churchland. 1988.

"Computational Neuroscience." Science 241, pp. 1299–1306.

Sejnowski, T. J., and P. S. Churchland. 1989. "Brain and Cognition." In M. I. Posner (ed.), *Foundations of Cognitive Science,* Cambridge, Massachusetts, The MIT Press, in press.

Sejnowski, T. J., and G. Tesauro. 1989. "Building Network Learning Algorithms from Hebbian Synapses." In J. L. McGaugh, N. M. Weinberger, and G. Lynch (eds.), *Brain Organization and Memory: Cells, Systems, and Circuits,* New York, Oxford University Press, in press.

Squire, L. R. 1987. *Memory and the Brain,* New York, Oxford University Press.

Von der Malsburg, C. 1987. "Synaptic Plasticity as Basis of Brain Organization." In J.-P. Changeux and M. Konishi (eds.), *The Neural and Molecular Bases of Learning,* New York, John Wiley & Sons, pp. 411–432.

Von Neumann, J. 1958. *The Computer and the Brain,* New Haven, Yale University Press.

Von Neumann, J. 1963. "The General and Logical Theory of Automata." In A. H. Taub (ed.), *John von Neumann: Collected Works,* New York, Pergamon Press, pp. 5, 288.

Waltz, D. H. 1988. "The Prospects for Building Truly Intelligent Machines." *Daedalus* 1, pp. 191–212.

Wolfram, S. 1983. "Statistical Mechanics of Cellular Automata." *Reviews of Modern Physics* 55, pp. 601–604.