Appendix

A. Evolution on the Timescale of Thought and Action: Darwinian Approaches to Language, Planning and Consciousness, and Some Lessons from Paleoclimate About How to Speed Up Evolution

Leader: William H. Calvin, [1]
Professor, University of Washington
June 2, 1998

Introduction

Questions about consciousness and intelligence arise in the context of both the information and biological revolutions. To understand these phenomena, this meeting began a two-part inquiry into the human brain. This summary was drafted by Richard Schum.

The goal of the study groups on the brain was to understand the its physical possibilities and limitations. In addition, we examined the physical origins of conscious thought and intelligence within the brain. This information is needed to put in context any technological enhancements to the human brain that might be researched and attempted. In addition, this information is key to understanding whether computers or networks of computers can ever achieve a kind of conscious intelligence.

As the information revolution unfolds, understanding the interaction between the biological brain and silicon-based intelligence is becoming an imperative. Electronics research and manufacturing continue to shrink computer chips (and hence computers) and increase computational speed. This suggests a future in which a minuscule computer with blinding speed and considerable memory would be available to nearly everyone on the planet. Recent research using biological and atomic-level processes to conduct computation suggests that the computers of the future may use, or be compatible with, biological material.

[1]For additional information on this subject, please visit http://WilliamCalvin.com

Some have called these developments "bioinformatics" or biological processing, although the developments are so new that language to describe it is lagging.

Humans already use computers as tools. It is not hard to imagine a day when the technology will enable people to implant a tiny computation machine into the human body or, alternatively, to use human biological processes to do computerlike computations. This raises a number of questions, such as whether such developments might be desirable and to what extent such developments might affect individual and group human behavior. Could such developments be the next step in human evolution?

Although computers are far from "intelligent" in the sense that we think of intelligence today, a number of people are working on creating intelligent systems, such as neural nets. Some would suggest that trying to create a computer modeled on the human brain is misguided and doomed to failure. Even so, this research, as well as research on the quantum basis for consciousness (a subject that will be discussed in the October 22, 1998, study group), has preoccupied a number of computer scientists and brain researchers.

To understand the future of mind, brain, and computing, and to question whether a human carbon-silicon brain is possible or even desirable, we need to understand the nature of the brain, how it evolved, and what challenges lead the brain to develop new capacities. In addition, it is important to consider the capacity of the brain—and possible alterations to this capacity—within the context of challenges to and impacts on human behavior, public policy, and the social order.

William Calvin: Evolution on the Timescale of Thought and Action

To help initiate discussion of mind, brain, and computing, the study group leaders invited Professor William H. Calvin, Ph.D., to present his research into the human brain and evolution. Professor Calvin is a theoretical neurophysiologist at the University of Washington in Seattle. He is the author of nine books including *The Cerebral Code* (MIT Press, 1996); *How Brains Think* (Science Masters 1996); and, with the neurosurgeon George A. Ojemann, *Conversations with Neil's Brain* (Addison-Wesley, 1994).

Calvin's research interests include the recurrent excitatory circuitry of cerebral cortex used for split-second versions of the Darwinian bootstrapping of quality, the fourfold enlargement of the hominid brain during the ice ages, and the brain reorganization for language and planning. He has long been following the paleoclimate and oceanographic research on the abrupt climate changes of the ice ages, hoping to find a connection to the "big-brain problem." He has an amateur interest in prehistoric astronomy and the associated archaeology.

He recently returned from a stay at the Rockefeller Foundation's study center in Bellagio, Italy, collaborating with linguist Derek Bickerton on their forthcoming book about the evolution of syntax, *Lingua ex Machina: Reconciling Darwin and Chomsky with the Human Brain*. His presentation centered on four themes: (1) levels of organization in the human brain, (2) using a Darwinian approach to understanding brain development, (3) the biological basis for a "Darwin Machine," and (4) how knowledge about the brain might be used to enhance brain functions. This paper summarizes Calvin's presentation.

Levels of Organization in the Human Brain

Artificial intelligence (AI) approaches to machine intelligence have traditionally been limited to the cost and capability of the computer. Originally, most AI work was analytical. The limited computational abilities of existing computers put a premium on efficiency and limited the ability of researchers to incorporate redundancy into their models. Recent advances in information technology are enabling the introduction of more randomness and redundancy into AI design. In an age of inexpensive computing and networking capabilities, AI is no longer a novelty—it has real applications. As a result, we may be able to model how the brain uses Darwinian processes—the same approach responsible for such biological adaptations as species evolution and immune response—to make decisions on the timescale associated with thought and action.

While computational speed is important, the key factor lies in levels of organization. Consider the following four levels: fleece, yarn, cloth, and clothes. Each is more highly organized and is the product of the one that came before it. For example, fabrics are woven of yarn, and yarn is spun. Each state is transiently stable and reflects a ratchetlike characteristic that prevents backsliding, or disorganization into its former state. Each level is causally decoupled from the next, so one can weave fabric without knowing how to spin yarn or make clothing with it. A whole set of techniques and body of knowledge exist within each level.

The organization of science reflects this kind of approach. For example, chemistry, the study of chemical bonds, can proceed pretty well without understanding anything about atomic spectra or about the Kreb cycle. Nevertheless, it certainly helps culturally for chemists to know something about atomic physics and biochemistry, even though they could function pretty well without them. Similarly, some of the highest functions of the nervous system, like consciousness, may in fact constitute another level of organization with rules of its own.

Within the neurosciences, there are probably a dozen levels. Some researchers say that memory arises from a synaptic change, while others argue that it is the nerve circuit or even gene expression that changes. Ironically, they are probably all right. This multidimensionality of explanation is what happens when a branch of science spans a number of levels of organization.

Within the nervous system, individual neurons produce impulse trains that can collectively affect the level of spatiotemporal patterns in the brain. The representation of a particular memory is not a result of a single cell firing, but rather the firing pattern within a cell committee (a Hebbian cell assembly, as it is more properly known). Thus, the key to memory functions appears to be pattern recognition and increased cell organization, not the behavior of the individual cells. By way of analogy, consider a computer screen. The behavior of each pixel—whether lit or unlit—has no meaning in and of itself. The meaning is derived from the pattern that is created, not from the individual constituents.

Larger assemblies that go beyond Hebbian-sized groupings also exist and are probably on the order of several square centimeters. These may represent objects, actions, relationships, analogies, or sentences. Composed of individual elements that are about 0.5 mm in size, these “hexagonal mosaics” compete against each other and attract additional members, each adopting the spatiotemporal pattern of its neighbors. This process of quality shaping via ad hoc assemblies continues until a winner emerges.

With various territories competing simultaneously with one another for the limited space on the association cortex, winners and losers are determined in a kind of “playoff.” The winning pattern becomes the conscious focus of the mind, and the “losers” become secondary or subconscious thoughts. A succession of focus occurs when the content of consciousness shifts and a new pattern prevails. This explains how the right birthday present for your spouse might suddenly pop into your head in the middle of a meeting.

A major theme of competitions—whether conscious, subconscious or unconscious—is the search for hidden patterns. In their first four years of life, children go through at least four major stages of discovery in identifying hidden patterns in their environment. During the first year, infants discover some three dozen speech sounds, known as phonemes (for example, ba, da, ca), and create standard categories for variants. They then discover unique patterns in strings of phonemes, known as words, at the rate of about six new words every day.

Between 18 and 36 months of age, toddlers discover structural relationships between words and syntax. The transition can often be very rapid, sometimes within one or two weeks. Finally, children discover Aristotle's rule about narratives—that they properly have a beginning, a middle, and an end—and start demanding appropriate endings to their bedtime stories. Of course, the search for meaning does not end there. As adults, we are constantly trying to make sense of our experiences and discern meaningful patterns in our actions, perceptions, and environment. Indeed, most of the tasks of consciousness are aimed at coping with novel situations, finding suitable patterns amid confusion, and creating new choices. Standard responses to ordinary situations do not require the attention of conscious thought.

Considering these factors, it appears that consciousness operates on different levels. If consciousness is defined as the highest current level of thought, it stands to reason that conscious thought operates at the level of objects and simple actions upon waking up in the morning. Forming relationships, like speaking in sentences, becomes possible only after a sufficient warm-up period or event, like morning coffee. Relations between relationships, like analogies, require even more time acclimation, like a double espresso. Given the ephemeral nature of consciousness, understanding how to improve the stability and duration of these levels is critical to building new ones. Such techniques could then be incorporated into our educational and training programs, enabling students to process information at higher levels of mental operation.

Using a Darwinian Approach to Understanding Brain Development

Given its well-deserved reputation for achieving quality on a timescale of species and antibodies, can the Darwinian process be used to improve the raw material of consciousness repeatedly beyond the incoherence of dreams? In principle, the problem is not whether it *can* be done—it can—it is whether coherence can be achieved quickly enough to be of use to our higher-level mental abilities, on the timescale of conversational replies.

The Darwinian process promotes coherence, or quality, through variation and selection over many generations. However, many aspects of selectionism are referred to as "Darwinian" when they are not truly Darwinian processes. For example, simple selective survival processes, such as leaf culling on trees, which results in a pattern, do not involve a Darwinian process. Neural connections in the brain also engage in the latter behavior by sprouting in abundant amounts, then culling into adult patterns. Take monkeys, for example: The axon count in the corpus callosum drops 70 percent drop between birth and the age of six months.

Similarly, there are a lot of connections from all parts of the cerebral cortex, even the visual cortex, down to the spinal cord at an early stage of development. By the time an individual reaches adulthood, only the motor strip and the premotor area still have direct connections. All the others have withdrawn. This sort of sprouting and culling is a major feature of the development of a timescale at the individual level. But it is not the recursive bootstrapping of quality that we associate with the reputation of Darwinian processes.

One should also not confuse change with Darwinism. Leaf locations can be modified by rotating a potted plant; climates vary, but so do skirt lengths. Alterations in quality or complexity are not associated with these changes, and successes are neither achieved nor repeated to achieve more success. Darwinian adaptations can be pyramided to achieve new levels; these cannot.

A Darwinian process has six essential characteristics:

- A pattern exists (e.g., genes).
- The pattern is copied or cloned.
- Variant patterns arise because of copying errors and recombinations.
- Populations of some of these variants compete against other populations for area (e.g., bluegrass and crabgrass).
- A multifaceted environment makes some of these variants more common than others (i.e., natural selection).
- The more successful variants serve as the most frequent centers for further variation, and future generations spread out to nearby regions to repeat this process (that is, Darwin’s Inheritance Principle).

All six characteristics are required to affect the recursive bootstrapping of quality. Having first five features, without the sixth, results in nothing more than population drift or random jumps from one barrier solution space to another.

So, can Darwinian processes be accelerated so that coherence, or quality, can be achieved quickly enough to be of use to our higher-level mental abilities? Once again, the answer is yes. Four known catalysts help speed the evolutionary process along:

- Systematic recombination: Variation is introduced systematically (for example, bacterial conjugation or sex).

- Fluctuating climates: Climate changes result in more-severe selection and more-frequent culling and, thus, in more-frequent opportunities for variation during expansion.
- Island biogeography, resource scarcity, geographical barriers and climate factors: These can separate a population into isolated subdivisions that discourage migration and promote inbreeding. Repeated separation and unification, or “pumping,” increase diversity and create variants capable of living further out on the habitat’s margins (e.g., the San Juan Islands, which are now surrounded by ocean but were hilltops connected by a broad valley during an ice age).
- Empty niches to fill: These can be due to the extinction of entire subpopulations; pioneers from other subpopulations will rediscover the vacated region and its replenished resources, and a population boom will result; an absence of competition for several generations results, giving rare variants that would otherwise perish a chance to survive. Once established, they may be able to survive future threats to their existence.

Attempts to duplicate the evolutionary process go well beyond the notions of connections and artificial neural networks. The best effort yet is Holland’s genetic algorithm that includes the six essential characteristics of the Darwinian process and one catalyst, systematic recombination. Even more promising, however, is the concept of a Darwinian machine that can incorporate these six characteristics and all four catalysts with stabilizing levels of organization. This concept will likely serve as the basis for intelligent machines of the future.

The Biological Basis for a “Darwin Machine”

While the brain contains the necessary circuitry for a Darwinian process, it has not yet been determined whether one actually occurs, much less how often or in which areas. Such a process would take place in a group of nerve cells in the cerebral cortex. As shown in Figure A.1, each neuron is a treelike structure that contains some 10,000 inputs, called synapses, and about the same number of outputs, called axons, that branch out to connect to other cells.

The pyramidal neurons are the excitatory neurons of cerebral cortex. Figure A.2 shows the axon of each of the three neurons spreading sideways in the superficial layer for a few millimeters in each direction. These cells are arranged in a pattern that is capable of a Darwinian process; those located in the deep layer pyramid do have such a pattern.

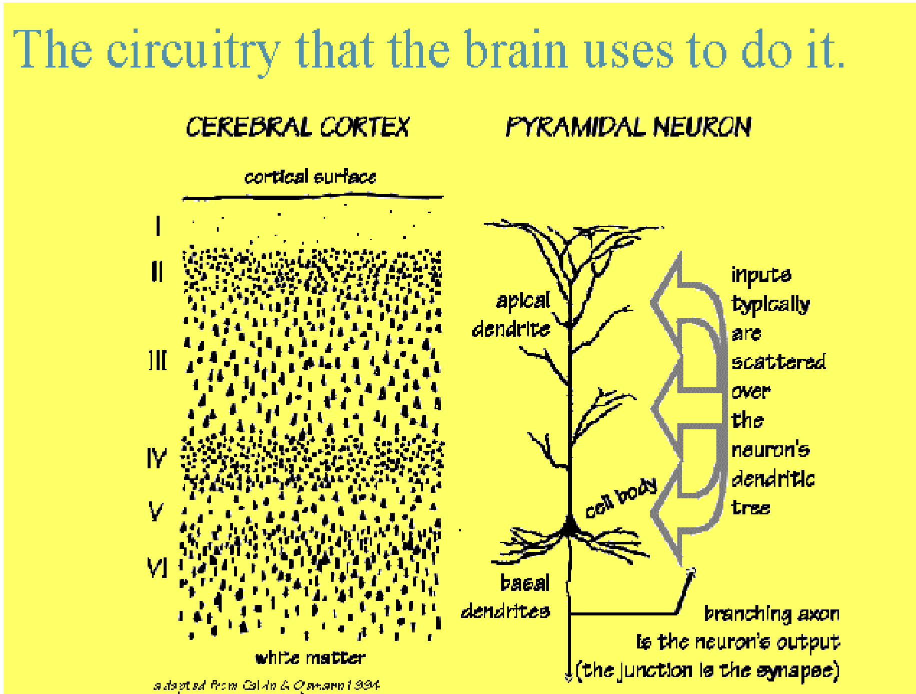

Figure A.1—The Brain's Circuitry

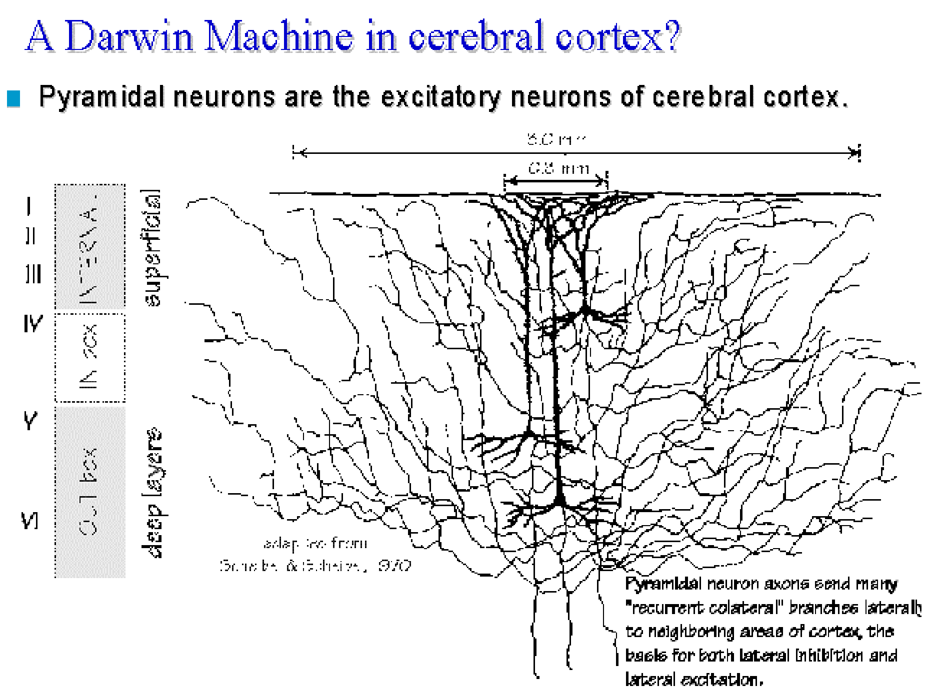

Figure A.2—Three Neurons and Their Axons

The cerebral cortex can be thought of as a series of six layers, as shown in Figure A.3. The deep layers tend to act as a sort of outbox, sending axons down to the spinal cord, thalamus, or other destination. The middle layers tend to receive inputs from the thalamus. Finally, the upper layers tend to act like an internal or interoffice mailbox, sending their outputs to other parts of cortex, either immediately to the side or down through the white matter.

Unlike their deep counterparts, the superficial layers exhibit a patterning of their axons, as shown in Figure A.4. The axon tends to go about a 0.5 mm before any output occurs. Figure A.5 is a drawing of a neuron and a cluster of synapses from the superficial layers. These clusters are formed by the overlap of axons terminating near their immediate neighbors, creating a sort of annular ring that surrounds the area at which an input is received. In general, such clusters will occur every 0.5 mm for a distance of 3 or 4 mm in three dimensions, though a local metric may dictate slight variations (for example, 0.65 or 0.85 mm, depending upon the exact location in the cortex). This “express train” arrangement allows outputs to skip intermediate junctions in their path.

Although this may look like a very simple-minded pattern, it is exactly what is required for Darwinian processing. When cells talk to one another, they tend to synchronize with each other. This is true of most excitable systems, and there are

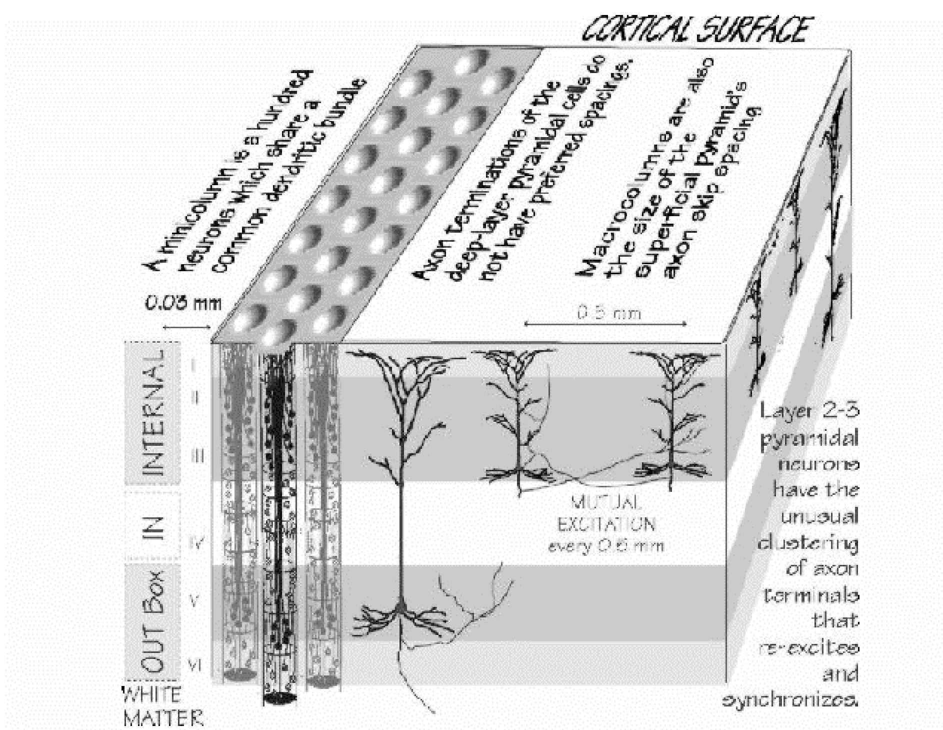

Figure A.3—The Layers of the Cerebral Cortex

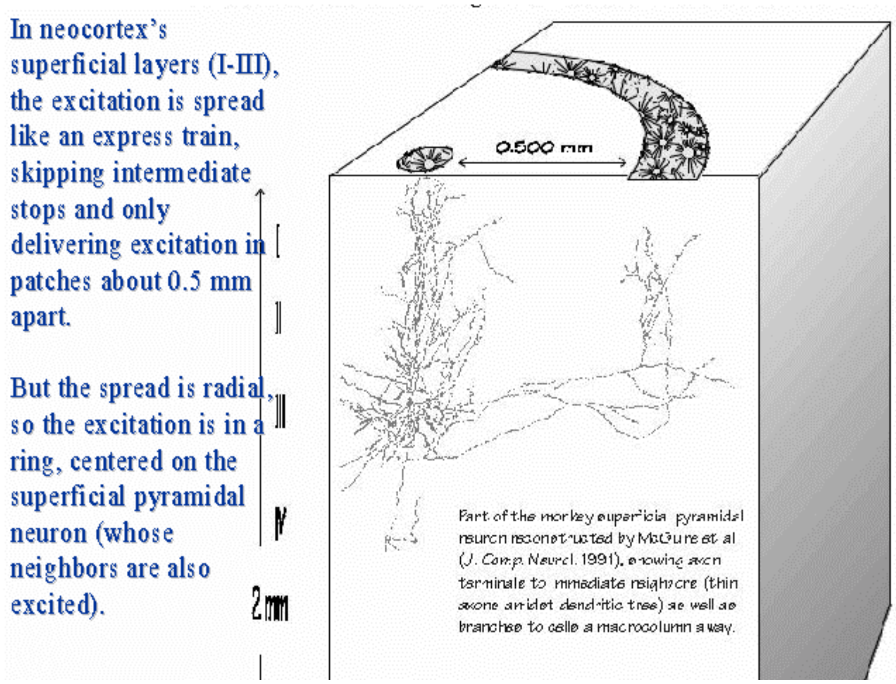

Figure A.4—Patterning in the Superficial Layers

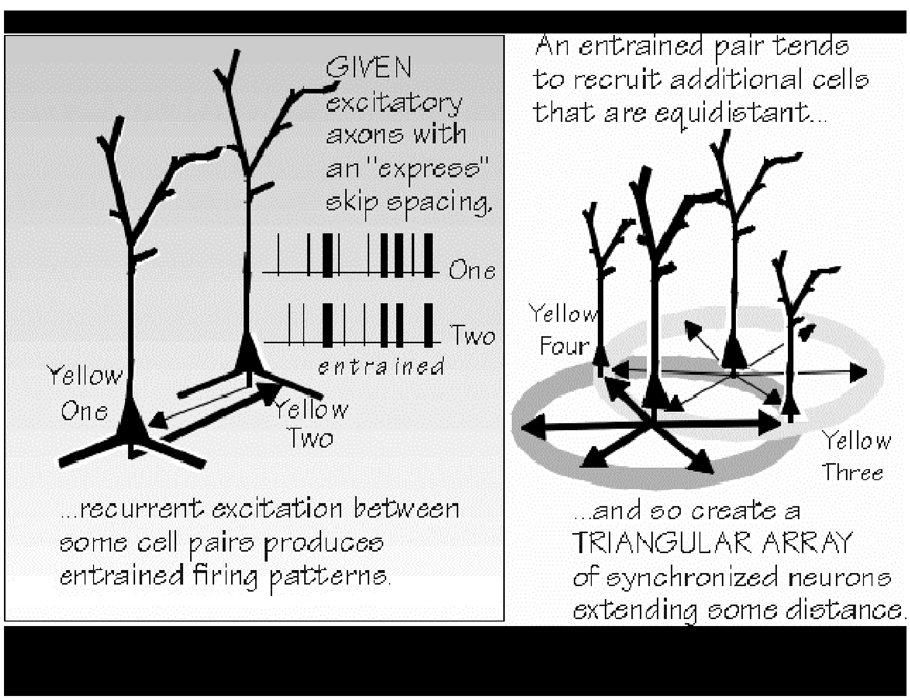

Figure A.5—A Superficial Layer Neuron and Its Synapses

many examples of this process, known as *entrainment*, in nature. Pendulum clocks on the same shelf synchronize their tics after about 30 minutes or so, and linear relaxation oscillators do it even more quickly. Neural cells are able to synchronize their firings so efficiently because they receive input simultaneously as a result of the overlapping ringlike structures of axons that, taken together, form a triangular array. Cells in this recruitment arrangement tend to fire together.

Many of these triangular arrays exist in the cerebral cortex—one might be sensitive to the color of an object, another might be sensitive to its shape. To avoid redundancy, the largest possible number of arrays is probably limited to a few hundred. Together, these arrays form hexagons in the “mosaic.” The pattern suggests that the spatiotemporal firing within each hexagon is a complete descriptive set, akin to a little musical pattern lasting only a couple of hundred milliseconds.

However, the spatiotemporal firing patterns in each descriptive set are not the only representation of a thought or action in the brain. Synaptic connectivity, the weightings that help maintain these firing patterns, enables the brain to remember these patterns. For example, the spinal cord has the ability to produce a number of different spatiotemporal patterns, called the “gates of locomotion.” This term is used to describe the manner and order in which a leg’s muscles fire to bring it forward. While each type of movement (e.g., walking, trotting, running) has its own spatiotemporal pattern, the same cells are used to do them all—it is just a matter of the initial conditions.

These two levels of representation—a short-term spatiotemporal pattern that is needed to effect thought or action and a long-term spatial pattern that is needed to store it. A good analogy might be a phonograph record whose spatial-only pattern of grooves is able to recreate the spatiotemporal pattern of music and speech. A consequence of this arrangement is that the triangular arrays do not always fire patterns, but rather compete for territory in the cortex. For example, some “undecided” areas (arrays) of the mosaic may receive input from two or more surrounding areas, each with a different firing pattern—say apples and bananas, as in Figure A.6. If the undecided area resonates better (due to a memory imprint) with apples than bananas, it will likely begin firing “apples.” Thus, success in cloning is subject to the extant memories of an environment. This competition provides the mechanism by which decisions are made and ambiguity is resolved. “Winning” spatial temporal patterns are responsible for any motor function output.

With this mechanism in mind, the essential Darwinian characteristics can now be evaluated in the context of the brain:

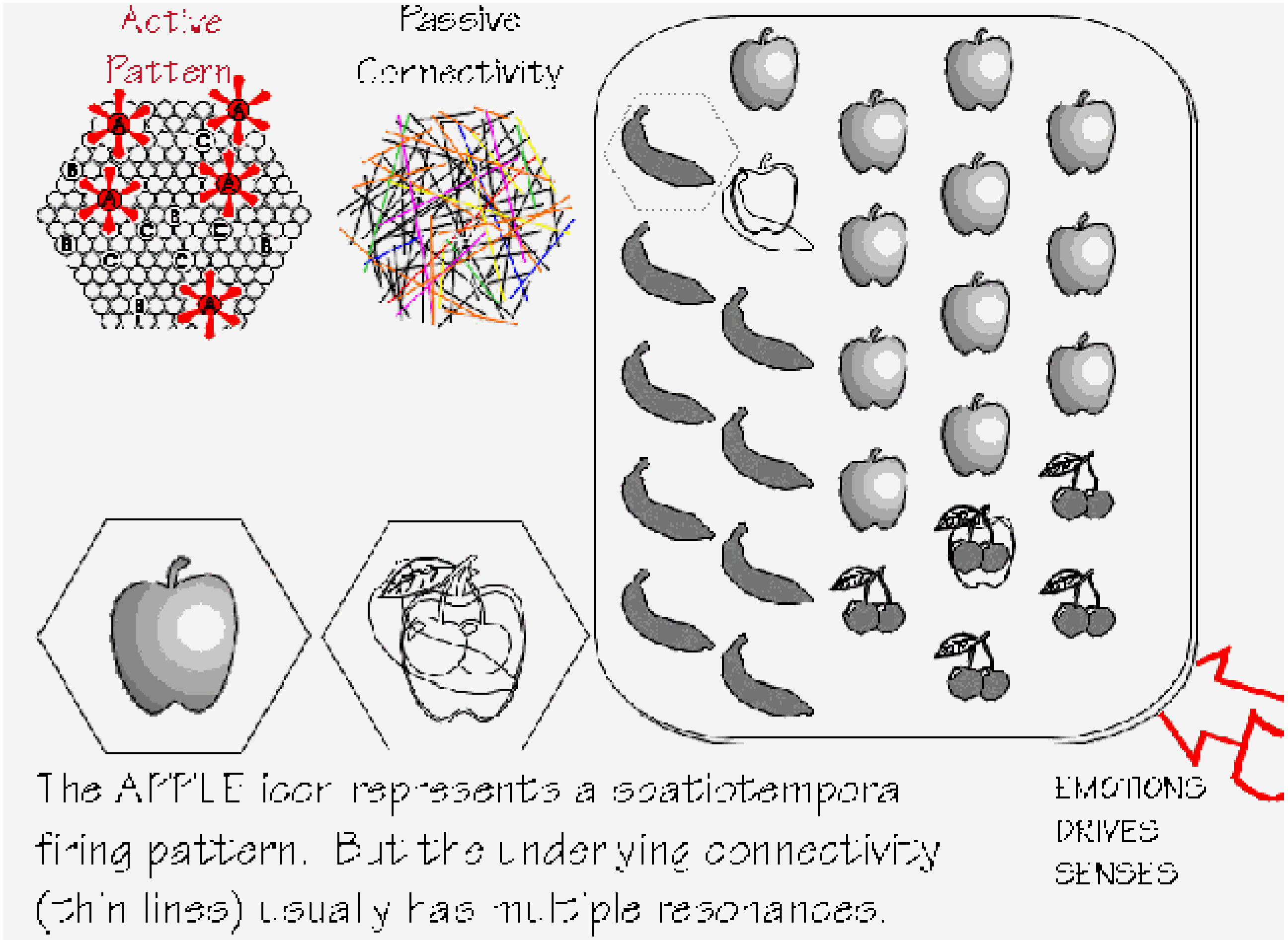

Figure A.6—Undecided Cells Receiving Input from Two Areas, in this Case, Representing Apples and Bananas

- A 0.5 mm hexagonal spatiotemporal pattern exists.
- The pattern is copied by recurrent excitation and entrainment of triangular arrays.
- Variant patterns arise when these triangular arrays escape conforming neighbors.
- Populations of these variants (hexagonal mosaics) compete for space.
- A multifaceted environment of sensory inputs and memorized resonances makes some variants more common.
- The more successful variants are the most frequent centers for further variation.

The evidence suggests that the neocortex could, in fact, be a Darwinian machine. It has all six essentials and all four optional catalysts, and it produces in advance the spatiotemporal firing patterns needed for converting thought into action. Success and quality are biased by real-time sensory inputs, the environment, and the memorized features of previous environments resulting from synaptic connectivity.

However, there are a lot of ways in which this mechanism can operate. While the brain circuit outlined above is fairly common and appears to be capable of performing all of these actions, it is not known how much time the different parts of the brain spend engaged in this activity. Is it used only during development to lay down the cortical structure or all the time in all areas of the cortex? It is likely somewhere in between. What is known is that these express-train connections exist in all of the common varieties of lab animals with one exception: the rat.

Another thing that is unresolved is whether the brain engages in anything fancier than the Darwinian process. While a more sophisticated process would appear to be unnecessary to perform these kinds of activities, it cannot be determined whether the brain uses this circuit in the exact manner described above until more precise observations from the neocortex are recorded. This technology is around the corner but is not yet available.

“Enhancing” Brain Function

Some glimpses of how we might improve this process emerge from this Darwinian view of how the brain could operate on this timescale. Understanding how a system works often makes it easier to improve its performance. For example, some higher aspects of intelligence—speed of learning, speed of operation, number of thoughts that can be held simultaneously in the mind—may be able to be addressed, toward the goal of eventually stabilizing higher, more-abstract levels of consciousness.

The concept of cerebral circuitry and Darwinian processing was first envisioned in the 1930s, but it is the technology of the 1990s that makes it possible to duplicate. Using a high-speed, hybrid computer, it appears feasible to emulate—not reverse engineer—this process. While it can be accomplished as a straight digital simulation, a more-natural fit involves the use of a hybrid digital-analog output device that would digitally copy spatiotemporal codes but record in analog spatial-only resonances. These resonances are what gives the circuitry its interesting properties. In the future, researchers will invent a number of circuits to undertake this task of bootstrapping quality through a series of generations. At that point, the notion of a Darwinian machine that thinks like a human will become a reality.

Finally, there is the issue of how to speed things up. A learning process that takes days to produce results is of little value to most brain functions. In the absence of further information, it is impossible to detail this process. However, some lessons about how to speed up evolutionary processes can be drawn from other Darwinian mechanisms. A primary factor appears to be "windows of opportunity" in behavior—what the French call *avoir l'esprit de l'escalier* [to have the spirit of the staircase]: thinking of the right reply on the stairway after leaving the party. From the perspective of the brain, the timescale associated with an evolutionary process must be a few seconds or less.

To illustrate the catalytic factors at work in this process, consider what happened to our ancestors over the past few million years. About 20 years ago, it was discovered that brain size started increasing some 2.5 million years ago, just around the same time that stone toolmaking became prominent. It turns out that it was during this period that the australopithecine branched off into the *homo* genus, as shown in Figure A.7. The question thus arises as to what was happening back then that could have caused all this to occur and continue? The most likely answer is the onset of the ice ages.

The Role of Climate in Developing the "Big Brain"

During the last 15 years, researchers have concluded that the ice ages were characterized by abrupt climate changes on a number of different timescales, as shown in Figure A.8. What is very obvious are the temperature fluctuations throughout the ages. For example, some 15,000 years ago, at a time when ice sheets covered the northern hemisphere, the temperature abruptly rose to almost

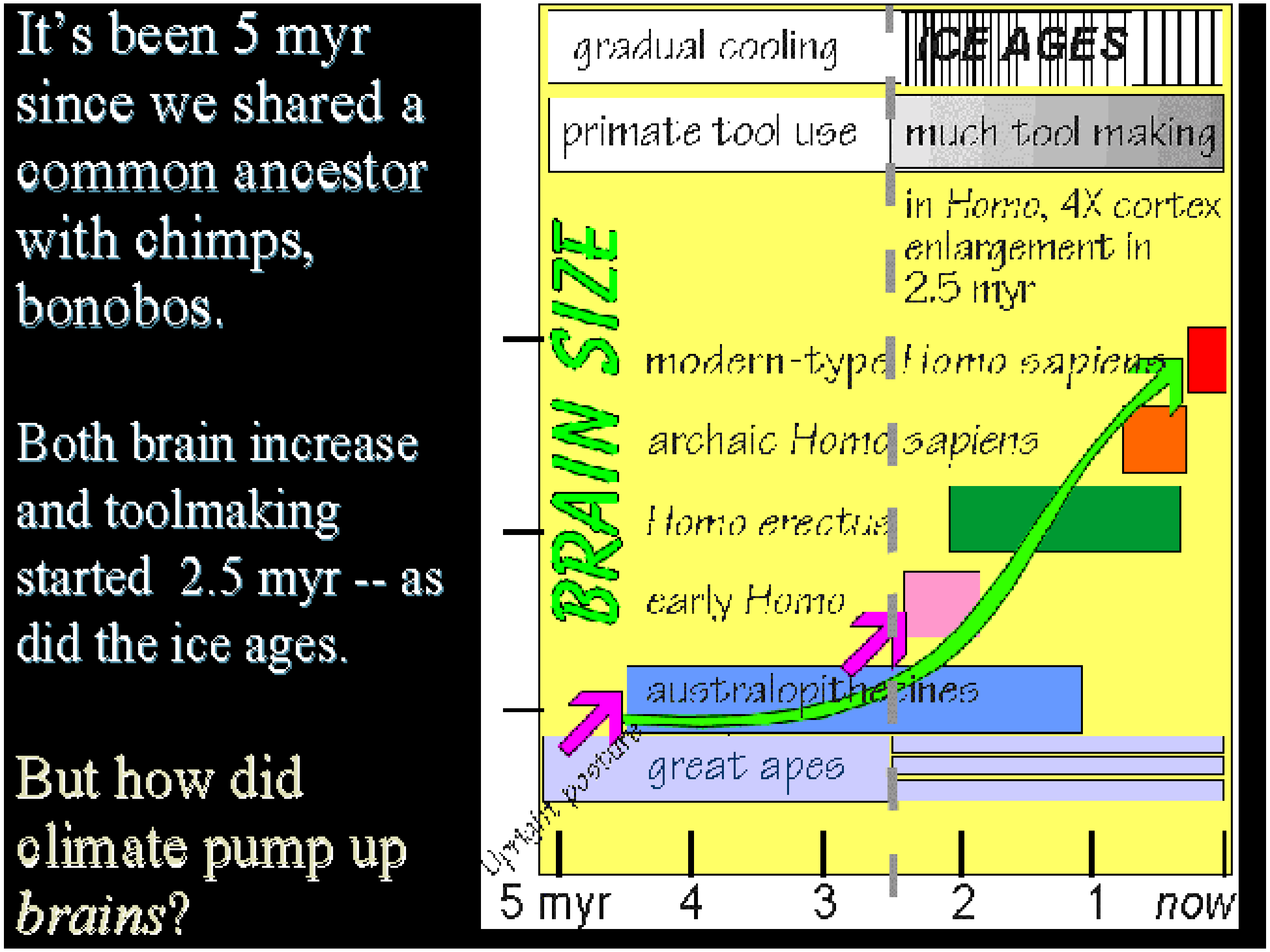

Figure A.7—The Emergence of the Genus *Homo*

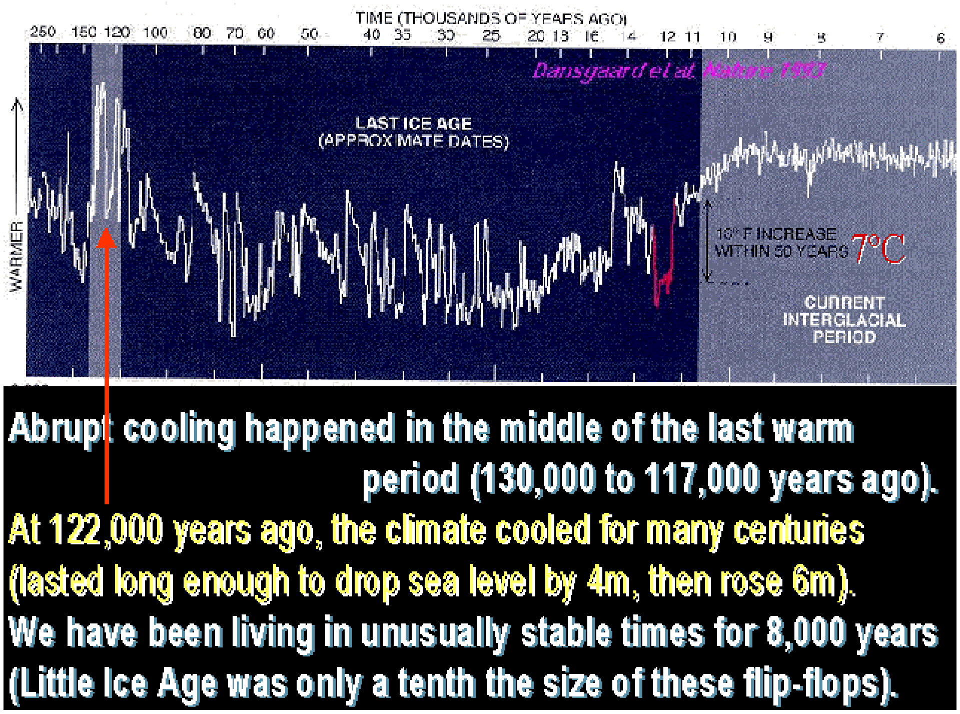

Figure A.8—Abrupt Climate Changes During the Last Ice Age

modern levels, despite all the ice. Some 2,000 years later, the temperature abruptly cooled and then warmed back up just as suddenly. One can see the same thing happening even further back on a very compressed timescale. Approximately 122,000 years ago, an abrupt cooling occurred in the middle of the last warm period. It lasted long enough for the sea level to drop some 4 m before warming back up to raise the sea level by about 6 m. About 8,000 years ago, there was also a brief period of moderate cooling.

The explanation for this climatic behavior is as complicated as it is lengthy. What is apparent, however, is that the climate has two stable states—a warm state and a cold state—and it flips between them based on the nature of the ocean currents, as shown in Figure A.9. The consequences of this transition are extreme. It is equivalent to jacking up (or ratcheting down) the entire landscape into a new climatic zone. Contrary to popular myth, however, it is not the magnitude of a cooling that threatens hominids, but its velocity, once the magnitude becomes large enough to effect the mix of plants and prey. That is, the process happens so quickly—within a human generation—that there is not enough time for biological adaptations to take place. The timescale is critical to survival.

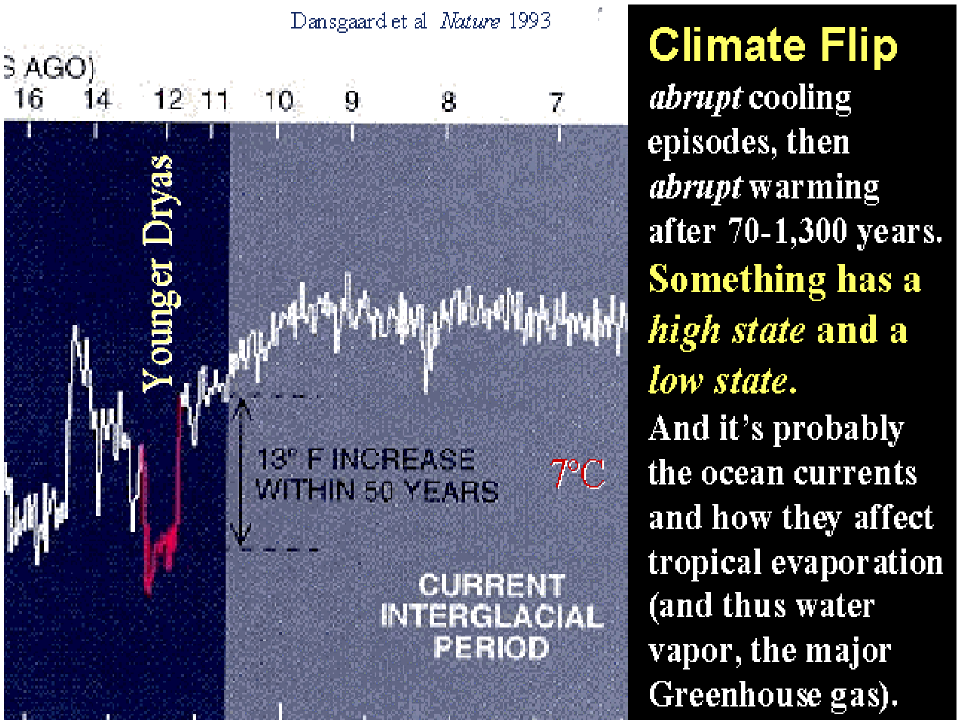

Figure A.9—Two Stable States, Cool and Warm

If, for example, the global temperature were to ramp down over a period of 500 years, life would be able to adapt to its new environment. A gradual change in vegetation would occur, emphasizing colder-weather species, like those normally found in higher altitudes, and hominids would likely learn to cope with new challenges. A stepwise cooling over a period of just 10 to 20 years, however, would pose a real threat to the survival of many species, including humans. Reduced rainfall would cause forests to dry up and burn off, leaving grass as the major food resource for at least a couple hundred years. To survive, then, an animal must either be able to eat grass or eat an animal that eats grass until plant secessions allow the ecosystem to advance past this monoculture to a forest more suited for these temperatures. The historical record indicates that our ancestors were subjected to many of these transitions over a period of thousands of years.

Conclusion: The Timescale of Thought and Action

Now transpose these lessons to the timescale of thought and action. The significance of timescale in the evolutionary process indicates that periods of monoculture are important to the neocortex. It stands to reason then that narrowly focused activities, such as concentrating, meditating or sleeping, will likely result in a downsizing and fragmenting of hexagonal mosaics into regional populations as cortical excitability fluctuates. Because these "climate" fluctuations occur rapidly, they "pump" the other three Darwinian catalysts—systematic recombination, island biogeography, and empty niches for new populations. Thus, it appears that one can control the speed at which change occurs by affecting the noise level in the neocortex.

Since climatic change occurs on various timescales, ranging from the millennia of ice ages to abrupt phenomena, such as el Niño. So, too, would we expect the brain's cortical "climate" to operate on various timescales (if an electroencephalogram is any acceptable measure of the brain's excitability). This process would involve repeatedly reducing and expanding to select the types of cells that are most capable of surviving bottleneck conditions. In addition to these quantum fluctuations, the neocortex is engaged in many parallel processes involving lots of territory, enabling it to maintain independent branches in a "playoff" system of alternatives. To further complicate the matter, different hexagonal arrays represent the different levels of organization and consciousness. As a result, a slow Darwinian process, such as forming a mental agenda, could bias a faster Darwinian process, such as thought and action, thereby skewing the results.

These characteristics are what one can expect from a forthcoming wave of Darwinian technologies. Of course, ethical questions must be considered. If a Darwinian circuit can be replicated in the artificial intelligence of a machine, it ought to be able to do what the Darwinian process is famous for elsewhere: shape up quality. While such a machine would have novel processing and problem-solving capabilities, it would not necessarily be considered "conscious." But as enhancements are made and versatility increases, society will face some very serious issues. For example, what if these technologies are able to work faster than humans? Is it possible to reach a point where all but the most intelligent people can be replaced by these devices? Then what happens when Moore's Law, some 18 months later, makes even those persons obsolete? Theoretically, there is no upper limit on processing speed if enough resources are available. These are the implications and dangers associated with building intelligence into machines.