Lloyd Watts, “Visualizing Complexity in the Brain”, in *Computational Intelligence: The Experts Speak*, edited by D. Fogel and C. Robinson, IEEE Press/Wiley, 2003, pp. 45-56.

Visualizing Complexity in the Brain

Lloyd Watts

How does the human brain work? The challenge presented by this question has motivated countless philosophers and scientists throughout history to study the brain and the nature of intelligence, in search of the organizing principles of human thought and perception. Yet, despite the enormous funding of the neurosciences and tremendous advances in technology in the latter half of the twentieth century, it appears that there is not yet a meaningful consensus on the organizing principles of brain function.

There are good reasons to believe that we are at a turning point, and that it will be possible within the next two decades to formulate a meaningful understanding of brain function. This optimistic view is based on several measurable trends, and a simple observation which has been proven repeatedly in the history of science:

> *Scientific advances are enabled by a technology advance that allows us to see what we have not been able to see before.*

Examples of enabling technologies include the invention of the compound microscope in the 1860's, the invention of neural staining techniques by Golgi in the 1880's, the invention of stroboscopic illumination, each of which enabled new scientific discoveries. In these cases, the technology breakthroughs that enabled something new to be seen had to do with seeing something *smaller*, *more translucent, or faster-vibrating* than could be seen before. But for understanding the brain in the early twenty-first century, we need to be able to see something *more complex* than we have been able to see before, which is a fundamentally different kind of problem. We already have a good understanding of the behavior of individual synapses, neurons, axons and dendrites. The interesting question is now: *How do we understand and visualize the complexity of the brain?*

The technology advance that allows us to see the complexity we could not see before is provided by the recent strong advances in computing and graphics technology. The availability of inexpensive computers with multi-gigahertz processors, gigabytes of on-board memory, hundreds of gigabytes of disk storage, and powerful graphics rendering chips provide an unprecedented platform for real-time brain modeling and visualization that simply did not exist even a few years ago. Kurzweil has projected Moore's Law out to 2020 and beyond (Kurzweil, 1999), and has concluded that we will have computers with sufficient memory and computing capacity to simulate major brain functions within the next twenty years. The question is: will we have the right algorithms to run on these phenomenal machines to simulate brain function and achieve brain-like performance?

I believe that the way to create a brain-like intelligence is to build a real-time working model system, accurate in sufficient detail to express the essence of each computation that is being performed, and verify its correct operation against measurements of the real system. The model must run in real-time so that we will be forced to deal with inconvenient and complex real-world inputs that we might not otherwise think to present to it. The model must operate at sufficient resolution to be comparable to the real system, so that we build the right intuitions about what information is represented at each stage. Following Mead, the model development necessarily begins at the boundaries of the system (i.e., the sensors) where the real system is well-understood, and then can advance into the less-understood regions (Mead, 1989). The model of the independent or well-understood parts of the system can be used to gain insights into dependent or less-understood parts of the system, since in many cases the well-understood parts provide inputs to the less-understood parts. In this way, the model can contribute fundamentally to our advancing understanding of the system, rather than simply mirroring the existing understanding. In the context of such great complexity, it is possible that the only practical way to understand the real system is to build a working model, from the sensors inward, building on our newly-enabled ability to *visualize the complexity of the system* as we advance into it. Such an approach could be called *reverse-engineering the brain*.

Note that I am not advocating a blind copying of structures whose purpose we don't understand, like the legendary Icarus who naïvely attempted to build wings out of feathers and wax. Rather, I am advocating that we respect the complexity and richness that is already well-understood at low levels, before proceeding to higher levels. Once information is thrown away, it can never be recovered. One of the powerful principles that is emerging about the operation of the brain is that it extracts and makes use of all the information in the signal.

Mead has pointed out that the constraints we impose on a problem have a powerful influence on our approach and the form of the solution that will result (Mead, 1989). Therefore, it is very important at the beginning of a project of this magnitude to be clear about the choice and prioritization of those constraints. I have had success so far using the following prioritization of constraints:

1. High resolution representations, verifiable against the biology.
2. Real-time operation and visualization of the results.
3. Fast design turnaround time.

A notable and perhaps surprising omission in this approach is any constraint on implementation technology. Allen advised an early focus on the algorithms, while remaining flexible on the implementation technology, since the project was likely to take many years and the implementation technology changes so fast (Allen, 1999). Since this work began in 1989, the algorithms have been implemented in several different technologies, appropriate for the questions under investigation and resources available at the time: analog VLSI, field-programmable gate arrays (FPGAs), batch-mode software on desktop machines displayed in real-time as QuickTime movies, and real-time software on a networked supercomputer. Other implementation technologies are imminent as the project verges on commercial deployment. The only constant is the need to get the algorithms right, in real-time, as soon as possible.

The ambition of reverse-engineering the brain in verifiable detail may appear overwhelming and unrealizable, but it is possible to make a sound argument that the goal is attainable within 20 years. Substantial progress has already been demonstrated on a significant subsystem with this approach. The issues to be addressed are:

1. **Neuroscience knowledge:** Do we know enough about the brain to begin building an artificial one?
2. **Computing technology:** Do we have a computational medium in which to prototype a design that can express the richness of the computations done in a real brain, such that the model could really inform the study of the real system?
3. **Non-technical issues:** Many experts will be required to contribute to the effort. Why should they help? Who will pay for the monumental effort? What is the economic model for funding the work?

As in all major endeavors, timing is everything. The approach can only succeed if all necessary ingredients are present and can contribute synergistically. Since 1950, advances in neuroscience knowledge and computing technology have led many workers to speculate that brain-like intelligence and performance was just around the corner, only to discover that the system was more complex, interconnected and robust than had been previously appreciated, and that far more computing horsepower was required than was available at the time. The bold attempts and subsequent disappointments in each decade since 1950 have all lead to a well-justified skepticism in both the scientific and investment community as to whether it will be possible to build a working intelligent machine. In fact, the previous attempts were based on overly simple models implemented on the inadequate machines of the day, which, in hindsight, did not have the necessary ingredients for success. I believe that the lesson to be learned from the previous disappointments was: the brain is much more complex than we would like to admit, and we need correspondingly complex models and serious computing horsepower, properly utilized, to build a robust, working system.

Fortunately, neuroscience knowledge and computing technology have both advanced dramatically in the last decade, as has our respect for the required complexity. Unusually favorable conditions exist for the period 2000-2020, whereas they most certainly have not existed in any previous time in history. For the remainder of this chapter, I will discuss the reasons for this optimistic view.

Neuroscience Knowledge

Does anyone know enough about the brain, in 2002, to be able to build an artificial one? No. There are still many unanswered questions about brain function, and there is no single individual who understands the brain in its entirety to the level that he or she could build a working model. But does the neuroscience community, in 2002, *collectively* know enough that we could *begin* modeling the brain in the well-understood regions? Absolutely. I will describe ongoing modeling efforts in the auditory pathway, shown in Figure 1, as an example brain subsystem in which significant progress has already been made, and as an example of the kind of complexity that is evident in the real brain.

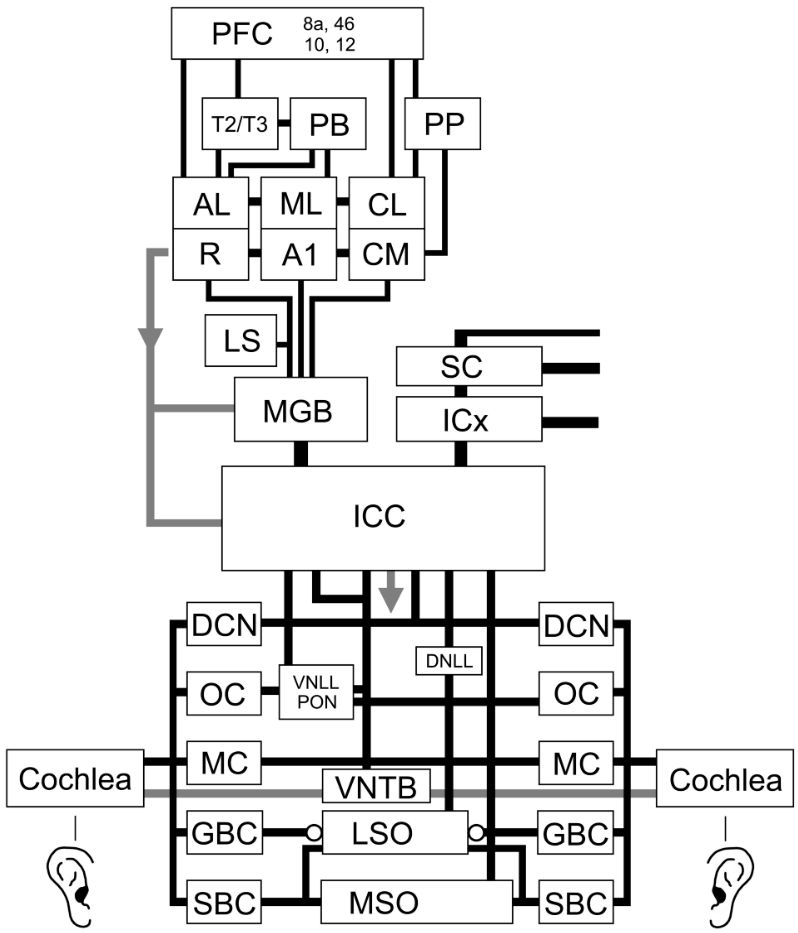

Figure 1: Auditory Pathway (highly simplified). Adapted from Young 1998, Oertel 2002, Casseday et al. 2002, LeDoux 1997, and Rauschecker and Tian 2000.

In the auditory system, the middle ear and cochlea are now well-understood, after about 100 years of research (the cochlea could not have been called well-understood even 10 years ago, due to controversy over the role played by the outer hair cells). It is now possible to build a real-time, high-resolution working model of the cochlea that accounts for its spectral sensitivity, temporal responses, nonlinear frequency-dependent amplitude compression and gain control, masking, and other subtle features, directly verifiable against biological and psychophysical data. This model can then be used in our investigation and modeling of the next layer in the system: the cochlear nucleus, which

Shepherd has described as one of the "best understood regions of the brain" (Shepherd, 1998). The basic circuit of the cochlear nucleus is given by Young (Young, 1998), describing in detail the essential cell types responsible for detecting spectral energy, broadband transients, fine timing in spectral channels, enhancing sensitivity to temporal envelope in spectral channels, and spectral edges and notches, all while adjusting gain for optimum sensitivity within the limited dynamic range of the spiking neural code. Again, it is possible to build a real-time, high-resolution working model of the cochlear nucleus that receives its inputs from the working model of the cochlea, and can be verified against real measurements from cochlear nucleus neurons.

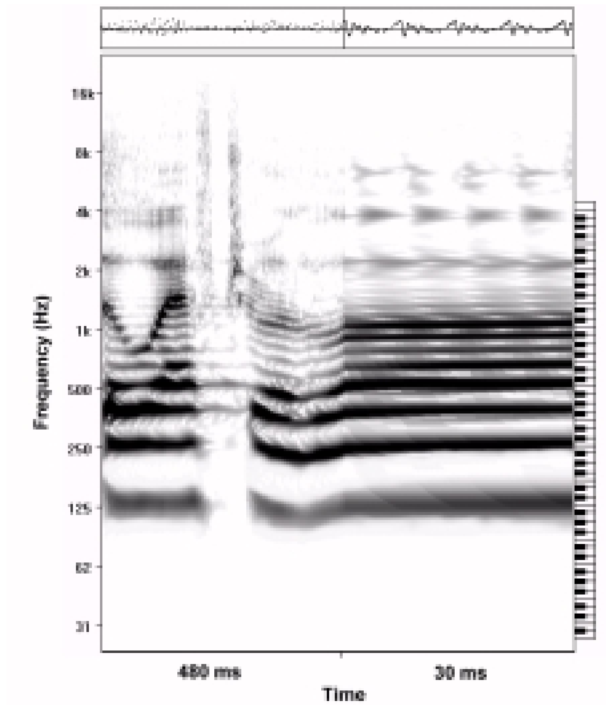

Figure 2: Multipolar Cell response to speech (male speaker in the middle of the utterance "so after a lot of thought").

An output of such an auditory model is shown in Figure 2, for a complex speech input. This shows the output of the ensemble of T-multipolar cells (labeled MC in Figure 1), which receive inputs from the cochlea / auditory nerve, and extract and encode the spectral energy in a way that is stable as sound level changes (Young, 1998). In Figure 2, the fundamental frequency of the speaker's voice can be seen at approximately 125 Hz, along with many harmonics at the integer multiples of the fundamental frequency. Strong bands of resonant energy can also been seen at around 500 Hz, 1 kHz, 2 kHz, and 4 kHz – these are the formant frequencies (resonances of the vocal tract), that correspond to and encode the different vowel qualities. The figure includes two time-scales to allow fine response detail to be seen in the right half of the display, with context over a longer period shown in the left half of the display. The four strong pulses at 4 kHz in the right half of the display correspond to the glottal pulse periods at about 8ms, giving a periodicity cue that corresponds to the pitch frequency at 125 Hz. Also evident in the display are two short bursts of high-frequency noise, corresponding to the letters *f* and *t* in the word "after". Figure 2 represents a snapshot of a working, high-resolution, real-time system – it is not possible on a printed page to convey the complexity of the animated output that responds in real time, synchronized with the input sounds.

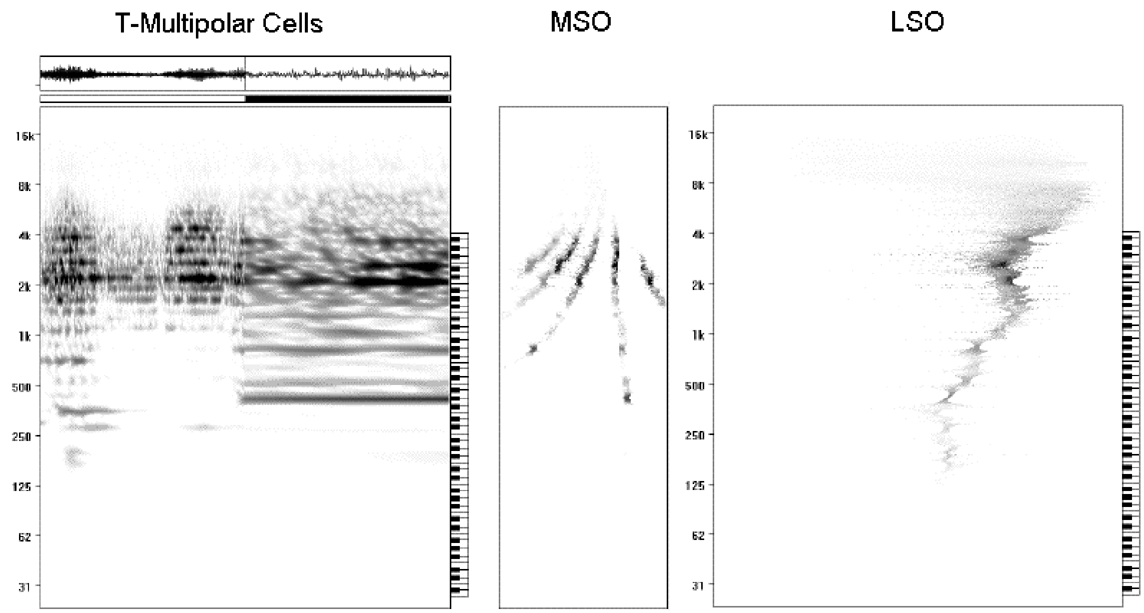

Figure 3: Multipolar Cell, MSO, and LSO response to a moving sound source coming from the right side.

Figure 3 shows a snapshot of the dynamic output of the auditory model for a moving sound source, showing the spectral energy representation computed by the multipolar cells in the cochlear nucleus, interaural time difference (ITD) representation computed by the medial superior olive (MSO) (Yin, 2002) and the interaural level difference (ILD) representation computed by the lateral superior olive (LSO) and further normalized and refined by the inferior colliculus (IC) (Casseday et al., 2002). These images indicate the complexity of the representations that are computed in the auditory brainstem, organized by the inferior colliculus, and conveyed to the cortex via the thalamus. If we are to have any hope of understanding what is happening in the cortex, we must first have a good model of all of the low-level representations that serve as its inputs.

In 2002, there is good understanding of the representations up to the Inferior Colliculus. However, there is no simple answer to the question "What is the Inferior Colliculus doing?" – it is doing everything! It aggregates, normalizes and organizes all the ascending representations from lower centers, computes new representations, and modifies them all with descending information from the cortex (Casseday et al., 2002). This does not mean that we will be unable to answer the question, it just means that the answer will be necessarily very complex, and we will need a way to express that complexity and verify that our understanding of it is correct, before being able to advance in a detailed way into the auditory cortex.

In 2002, in the absence of a conclusive model of the inferior colliculus, does this mean we must wait another decade before proceeding? Not at all. Bregman has elucidated the psychoacoustic principles by which sounds are grouped to form perceptions of objects in auditory scenes (Bregman, 1990). Many research groups around the world, including Sheffield and MIT, have been building models that account for various parts of the scene analysis machinery. What is important is that we link the high-level psychoacoustic model to the low-level neurophysiological model, so as to reap the benefit of all the information that is extracted by the lower levels of the real brain.

One example of the opportunity that exists is in the area of speech recognition. Present-day speech recognizers are capable of 90-98% word recognition accuracy, depending on vocabulary size, number of users, noise levels and other factors. It has taken forty years and thousands of talented engineers to build speech recognition systems this good. And yet, these systems still perform very poorly compared to humans (Lippmann, 1997). These systems use a kind of engineered psychoacoustic model (phoneme classification and Viterbi search through a space defined by a Hidden Markov Model) to account for the human cortical recognition process, while relying on a very low-resolution front-end (128-point Fast Fourier Transform (FFT), smoothed and orthogonalized to create a 13-point cepstrum, updated every 10 milliseconds) to provide the spectral information to the phoneme classifier (Rabiner and Juang, 1993).

At the level of the FFT, the amount of information represented is 64 spectral coefficients, 100 times/second, or 6,400 values/second. By comparison, the human cochlea represents the spectrum with 30,000 auditory nerve fibers, representing the outputs of approximately 3,000 inner hair cells (10:1 for better signal representation). The cochlea produces outputs with 6 microsecond temporal resolution (Yin, 2002), with 3,000 unique outputs represented, organized on an (approximately) logarithmic frequency scale. Accounting for the fact that low-frequency channels do not carry as much information as high-frequency channels, we estimate the bandwidth on the auditory nerve to be about 10 million values/second, approximately 1,500 times as much information as the frame-based FFT, as used in a conventional speech recognizer. (The expansion in bandwidth from the raw audio waveform at, say, 44,100 samples/second, to the auditory nerve representation at 10 million values/second may seem surprising – the auditory nerve representation is a high-bandwidth redundant version of the signal from which the higher centers can extract the necessary information). The higher temporal and spectral resolution is vital for performing auditory stream analysis to extract speech from background sounds and to detect fine timing distinctions used in distinguishing confusable consonants – precisely the two areas in which existing speech recognizers fail. Speech recognition is a prime example of a relatively powerful psychoacoustic model of a cortical process (Hidden Markov Models and Viterbi Search) being starved by a poor model of a subcortical process, to the detriment of overall system performance.

The preceding argument suggests that it should be possible to improve speech recognition with an auditory model front-end that replaces the FFT/filterbank/cepstral front-end. To date, all such attempts have failed to produce any significant improvement, so that, again, there is great skepticism in both the scientific and investment community about the potential for progress in this area. Hermansky has surveyed the previous attempts and offered several possible explanations for the disappointing results, including a wise warning of the dangers of blindly copying what is known about the brain without understanding the underlying principles (Hermansky, 1998). I agree that it would be pointless to simply bolt a cochlear model onto a Hidden Markov Model. The biological system does a tremendous amount of feature extraction between the cochlea and the cortex, and within the cortex, as shown in Figure 1. It is not surprising that early naïve attempts to connect a cochlea to the cortex have failed. Success will be achieved when we extract all the robust features used by the human system to process speech, and feed them correctly into an appropriate cortical model. It is not possible to extract the necessary features if vital information is thrown away at the first processing step. The cues for distinguishing the confusable consonants are well-known (Miller and Nicely, 1955, Ladefoged 2001), and the high-resolution cochlea model supports the extraction of all of those cues, whereas the frame-based FFT method fundamentally does not. In addition, the cochlea model provides all the information used by the human system to separate the speech from other sounds, whereas the frame-based FFT method again does not. I expect to see improved speech recognizers, based on powerful auditory models, commercially deployed by 2005-6, with vastly better performance than present-day systems both in quiet conditions and in noisy, reverberant environments.

The previous discussion has focused on the auditory system, and built the case that the neuroscience community already knows an overwhelming amount about how it works. Everything we know about the auditory system suggests that it is far more complex than any present-day engineering model. The same case can be made for the visual system. The visual system has been mapped out in detail in the macaque monkey (Van Essen and Gallant, 1994), resulting in a convenient high-level block diagram from the retina up to complex cortical areas responsible for recognizing faces, supported by hundreds of studies of the various cell-types and their connectivity. Douglas and Martin have studied the neural structure of the neocortex, and provided a canonical microcircuit diagram for conceptualizing its operation (Douglas and Martin, 1998). Calvin has offered a rather speculative theory of how the cortex might work, useful at least in encouraging us to begin thinking about a connection between the cellular connectivity of the cortex and its high-level functions (Calvin, 1996). Churchland and Sejnowski have examined many aspects of neurobiology, including learning and memory (Churchland and Sejnowski, 1992). The problem is not that we know too little about the brain. The problem is that what we do know is overwhelming, and we need a fundamentally new way to think about it, visualize it, collaborate on it, and consolidate it into a working engineering model.

Computing Technology

Do we have computers powerful enough, in 2002, to be able to build an artificial brain? No – at least most of us don't; there are some research labs that are making impressive progress, however. In 2002, IBM is the current record-holder with a 7.2×10^{12} operations/second machine (Top500 website, 2002). Present-day desktop computers are capable of performing about 4×10^{9} operations/second. Kurzweil has estimated that we will need a computer capable of performing 10^{16} operations/second, and that these machines will not be available, at reasonable cost, until about 2020 (Kurzweil, 1999).

But that is the estimated computing performance needed to simulate the *entire* human brain. As described in the previous section, we don't need to do that yet – we only need to simulate the well-understood parts, well enough to let us branch into a new less-understood part. The available computing power has recently become adequate for that purpose. In my own work, I have found that a modest amount of computing hardware is sufficient to simulate the major functions of the auditory pathway up to the inferior colliculus, with efficient algorithms. Rodney Brooks stated in a 1999 colloquium at Stanford University that, in his work on building anthropomorphic motor systems, "a paradigm shift has recently occurred – computer performance is no longer a limiting factor. We are limited by our knowledge of what to build." (Brooks, 1999)

Prior to 1999, we knew enough about the brain to have been able to build a little of it, but the necessary computing power was not available. This meant, in practice, that researchers struggled with feeble computers to painstakingly simulate the operation of a few neurons, and did not really learn anything they didn't already know. Now, the computing power is great enough that we can implement a detailed model of a subsystem that we understand, and learn something *new* about it by watching it run in response to a complex, real-world stimulus. This new ability is related to Kurzweil's Law of Accelerating Returns (Kurzweil, 2000), and it will add an important new element to neuroscience research. The only requirement will be that the model be realistic, that is, verifiable against the neurobiology. Otherwise, we will get a fast answer about some system we made up, not a fast answer about the brain. This is why priority #1 must be: High resolution representations, verifiable against the biology.

Computing power, in operations/second, is not the only factor in the paradigm shift. How will we verify our high resolution representations? The only way I can see is to make high-resolution animated images of them. In addition to requiring a lot of computing power, this also requires fast, high-resolution graphics rendering capability. Even as recently as 1998, real-time graphics rendering of high-resolution brain simulations was simply not possible on affordable machines. In 2002, it is not only possible, it is inexpensive.

Non-Technical Issues

With a wealth of neuroscience knowledge available with which to begin, and all the computing power we can really use, what else do we need to undertake the program of reverse engineering the brain?

The first major issue is the need for direct collaboration with qualified neuroscientists. In the first section of this chapter, I stated that the neuroscience community *collectively* knows enough about the brain that we could begin a detailed modeling process. But their collective knowledge won't help us build a working model. Someone has to distill their knowledge into a concise form that can be efficiently implemented in a machine. This requires an active, long-term collaboration between the modeler and many neuroscientists – it is simply not possible to learn what the system is doing by reading their papers. There are just too many papers. And the only way to verify that the model is right is to show it to the neuroscientists who measured the real system, and keep changing it until the neuroscientists agree that it is right.

The other major issues relate to funding. Who will pay for this effort? What is the economic model for funding the work? Are there commercial applications that can justify an investment model?

There are several ways that this kind of project can be funded. The choice of funding model and funding source depends on the particular characteristics of the problems being addressed and the people involved, and could range anywhere from an academic lab funded by a government agency, a corporate research lab, or a startup funded by angel investors or venture capitalists, if commercial applications can be developed on the appropriate time-scale. I have found that the single biggest obstacle in funding this kind of project is the widespread skepticism in both the scientific and investment communities, after so many decades of high hopes and deep disappointments. I have been very fortunate to find scientific advisors and visionary investors who have taken the time to understand the promise of the approach, and by contributing to the effort, are causing it to succeed.

Conclusions

At about the turn of the twenty-first century, we passed a detectable turning point in both neuroscience knowledge and computing power. For the first time in history, we (collectively) know enough about our own brains, and have developed such advanced computing technology, that we can now seriously undertake the construction of a verifiable, real-time, high-resolution model of significant parts of our own intelligence. The ability to visualize the staggering complexity as we develop and verify the working model will be a necessary element in this ongoing program.

References

Allen, P. (1999), Interval Research internal project review, personal communication.

Bregman, A. (1990), *Auditory Scene Analysis*, MIT Press.

Brooks, R. (1999), Stanford Engineering Department Colloquium, personal communication.

Calvin, W. (1998), *The Cerebral Code*, MIT Press.

Casseday, J., Fremouw, T., Covey, E. (2002), in Oertel, D., Fay, R., and Popper, A., ed., *Integrative Functions in the Mammalian Auditory Pathway*, Springer-Verlag, New York, pp. 238-318.

Churchland, P., and Sejnowski, T. (1992), *The Computational Brain*, MIT Press.

Douglas, R., and Martin, K. (1998), in Shepherd, G., ed., *The Synaptic Organization of the Brain*, fourth edition, Oxford University Press, pp 459-510.

Hermansky, H. (1998), "Should recognizers have ears?", *Speech Communication*, vol. 25, num. 3-27.

Kurzweil, R. (1999), *The Age of Spiritual Machines*, Penguin.

Kurzweil, R. (2000), *The Singularity is Near*, book precis.

Ladefoged, P. (2001), *A course in phonetics*, Harcourt Brace, 4th edition.

LeDoux, J. (1997), *The Emotional Brain*, Simon and Schuster, New York.

Lippman, R.P. (1997), "Speech recognition by machines and humans", *Speech Communication*, vol. 22, no. 1, pp 1-15.

Mead, C. (1989), *Analog VLSI and Neural Systems*, Addison-Wesley.

Miller, G., and Nicely, P. (1955), "An analysis of perceptual confusions among some English consonants", *Journal of the Acoustical Society of America*, Vol. 27, no. 2, March.

Oertel, D. (2002), in Oertel, D., Fay, R., and Popper, A., ed., *Integrative Functions in the Mammalian Auditory Pathway*, Springer-Verlag, New York, pp. 1-5.

Rabiner, L., and Juang, B. (1993), *Fundamentals of Speech Recognition*, Prentice Hall, New Jersey.

Rauschecker, J., and Tian, B. (2000), "Mechanisms and streams for processing of "what" and "where" in auditory cortex", *Proceedings of the National Academy of Sciences*, vol. 97, no. 22, 11800-11806.

Shepherd, G. (1998), *The Synaptic Organization of the Brain*, fourth edition, Oxford University Press, p. vi.

Top 500 supercomputers website (2002), http://www.top500.org/ .

Van Essen, D., and Gallant, J. (1994), "Neural Mechanisms of Form and Motion Processing in the Primate Visual System", *Neuron*, Vol. 13, 1-10.

Yin, T. (2002), in Oertel, D., Fay, R., and Popper, A., ed., *Integrative Functions in the Mammalian Auditory Pathway*, Springer-Verlag, New York, pp. 99-159.

Young, E. (1998), in Shepherd, G., ed., *The Synaptic Organization of the Brain*, fourth edition, Oxford University Press, pp. 121-158.