ELECTRON IMAGING TECHNOLOGY FOR WHOLE BRAIN NEURAL CIRCUIT MAPPING

KENNETH J. HAYWORTH

Center for Brain Science, Harvard University,
52 Oxford St. Northwest Bldg. 347.20, Cambridge MA 02138, USA
kenneth.hayworth@gmail.com

The goal of uploading a human mind into a computer is far beyond today’s technology. But exactly how far? Here I review our best cognitive and neuroscience model of the mind and show that it is well suited to provide a framework to answer this question. The model suggests that our unique “software” is mainly digital in nature and is stored redundantly in the brain’s synaptic connectivity matrix (i.e., our Connectome) in a way that should allow a copy to be successfully simulated. I review the resolution necessary for extracting this Connectome and conclude that today’s FIBSEM technique already meets this requirement. I then sketch out a process capable of reducing a chemically-fixed, plastic-embedded brain into a set of tapes containing 20 x 20 micron tissue pillars optimally sized for automated FIBSEM imaging, and show how these tapes could be distributed among a large number of imaging machines to accomplish the task of extracting a Connectome. The scale of such an endeavor makes it impractical, but a version of this scheme utilizing a reduced number of imaging machines would allow for the creation of a “Connectome Observatory” — an important tool for neuroscience and a key milestone for mind uploading.

Keywords: Connectome; mind uploading; FIBSEM.

1. Introduction

On May 25, 1961 John F. Kennedy delivered his now famous speech to a joint session of congress stating “…I believe that this nation should commit itself to achieving the goal, before this decade is out, of landing a man on the moon and returning him safely to the earth”. This was an incredibly ambitious goal especially given that the only American space flight prior had been Alan Shepard’s sub-orbital flight aboard a Redstone rocket less than a month before the speech. However Kennedy’s science advisors were reasonably sure that the difficulty lay mainly in great engineering challenges — much more powerful versions of the liquid fueled Redstone rocket, more sophisticated control and guidance electronics — and did not require radically new science or technologies to achieve. Today we may be in a similar situation relative to the goal of uploading a human mind into a computer. It is unlikely that radically new science and technologies are needed to achieve the goal of “landing a
man in cyberspace and returning him safely to consciousness”. I will argue that the key technology that needs to be developed to make mind uploading a reality is the ability to create a 3D volume electron microscope image of an entire chemically-fixed and plastic-embedded brain at sufficient resolution to allow the tracing of synaptic connections between all neurons. As in the Apollo program, such a feat represents an enormous engineering challenge but the basic technology for 3D electron imaging of brain tissue at this resolution exists today, and I will lay out one possible path for how this imaging technique could be robustly scaled to the size of an entire human brain.

In this paper I first review today’s most well supported cognitive and neuroscience model of the human cognitive architecture (ACT-R), and show how it implies that a successful mind uploading could be accomplished based only on a static map of all neural connections in a brain (along with sufficient structural information on synapse size, location, cell morphology, etc. to allow correlation with functional experiments performed on other brains). Such a static map of all neural connections has been termed the “Connectome” [Seung, 2012] and a new field called “Connectomics” has sprung up over the last few years [Eisenstein, 2009] developing automated electron microscopy techniques for the 3D imaging of brain tissue at resolutions sufficient to reliably map such connections over long distances.

I review one of these techniques in particular, the Focused Ion Beam Scanning Electron Microscopy (FIBSEM) technique [Knott et al., 2008], which has demonstrated the ability to image tissue at resolutions approaching $5 \times 5 \times 10$ nm. The finest axonal and dendritic process in the brain can shrink down to $\sim 40$ nm in diameter and researchers using traditional electron microscopy techniques based on 50 nm thick serial sections have found it difficult to reliably trace such processes over long distances using such thick sections. In contrast, the new FIBSEM images have such high resolution that all neuronal processes should be reliably traceable even using computer automated algorithms. Furthermore, FIBSEM’s use of a focused ion beam (as opposed to traditional physical sectioning with a diamond knife) gives it the potential to achieve the reliability and large-scale automation that would be necessary for a human mind uploading attempt. In these regards, FIBSEM represents a truly enabling technology for mind uploading. Continuing with the rocket science analogy, we might say that today’s FIBSEM resolution demonstration is equivalent to the first demonstration of a liquid fueled rocket engine achieving a specific impulse theoretically sufficient to achieve orbital velocity.

A fundamental physical limitation of the FIBSEM ablation process is that this high resolution can only be obtained for very small samples on the order of 20 microns across. To overcome this limitation I have developed a technique [Hayworth et al., 2010] which can section large blocks of plastic-embedded brain tissue into 20 micron thick strips optimally sized for high-resolution FIBSEM imaging. This thick sectioning procedure results in such high-quality surfaces that the finest neuronal processes can be traced from strip to strip. This lossless thick-sectioning technique opens up the possibility of applying the high resolution of FIBSEM to arbitrarily large tissue volumes (potentially even whole brains), and allows for the use of massively
parallel imaging by spreading the thick strips across multiple FIBSEM imaging machines.

Sectioning an entire human brain into 20 micron thick strips with 100% reliability will be a difficult engineering challenge, but one that appears quite achievable. Below I discuss the technical challenges in preserving an entire human brain at the ultrastructure level suitable for electron microscopic imaging. I then sketch out an industrial process which could reduce such a chemically-fixed and plastic-embedded brain into a set of tapes containing 20×20 micron tissue pillars optimally sized for automated high-resolution FIBSEM imaging. Finally I discuss a set of technical milestones (which I believe are reachable in the next few years) which would demonstrate all of the crucial imaging steps necessary for uploading a human mind.

2. Evidence that We are Our Connectome

A central thesis of this paper is that an electron microscope volume image of an entire chemically-fixed and plastic-embedded human brain (with sufficient resolution to allow the tracing of synaptic connections between all neurons) would contain sufficient information to fully encode the unique memories, skills, and personality of that particular individual. Further, given this static image (and a thorough understanding of the functioning of individual neuron types gleaned from side experiments not performed on this particular brain) it is hypothesized that a computer simulation could be constructed that when turned on would behave so similarly to the original that close friends and relatives would be hard pressed to distinguish (from behavior alone) this simulation from the original. Finally, and crucially, it is hypothesized that this simulation would exhibit first-person conscious experiences very similar to those experienced by the original person. This is, after all, what is meant by the phrase “mind uploading”. There are many levels at which one could disagree with this thesis so I will briefly cover the logic behind it before moving on to the technical discussion of how to obtain such a high-resolution volume image of an entire brain.

The most fundamental assumption on which this thesis depends is that the human mind and its conscious experiences are purely a computational phenomenon. We are broadly similar to an intelligent robot. Our body is analogous to the robot’s mechanical body. Our brain is analogous to the robot’s computer hardware including the physical hard drives which contain its control software. Our mind is analogous to the computational processes that result by the running of this control software on the robot. Our first-person conscious experiences are analogous to the particular stream of tokened data structures that the robot uses to model its own perceptions, goals, and actions as part of its self-model (used as a higher-order guide for reasoning about and planning its future behaviors). What would “mind uploading” mean for such a robot? It would simply mean that the robot’s software is copied to another computer

1Sebastian Seung in his book “Connectome” [2012] has introduced the phrase “We are our Connectome” to tersely encapsulate the idea that a sufficiently high-resolution map of brain connectivity would capture everything that is unique about an individual. I am adapting his phraseology here.
controlling another robot body. This of course would result in a similar behaving robot with a similar tokening of internal data structures. Although in principle there may be some deep flaw in this analogy between a human and a mechanical robot, it is an analogy that rests at the core of all of our research in cognitive science, neuroscience, and even biology itself.

If we accept this general “computational theory of mind” assumption there is still the question of whether the human mind’s computational process is even remotely similar to what researchers in the Artificial Intelligence (AI) community would recognize. Decades of research in AI have produced incredibly sophisticated computational systems capable of all manner of intelligent behavior. These results provide the bedrock for understanding in general how a physical machine for manipulating symbols can produce a wide variety of intelligent goal-directed behaviors. However, the algorithmic structure of most of these AI systems is completely incompatible with what is known about the computational limitations of the human brain. An often cited example of this is the 100-step constraint [Feldman and Ballard, 1982]. Since computational results are signaled between neurons at timescales no faster than a few milliseconds, human behaviors like visual object recognition (which can be completed within a few hundreds of milliseconds) must not be being implemented by algorithms requiring more than 100 sequential computational steps. It is clear that the vast majority of symbolic AI algorithms for vision, problem solving, speech recognition, etc., violate this and other known nervous system constraints. These ubiquitous violations are sometimes taken to imply that the human brain should not be viewed as a sequential symbol processor at any level [Fodor and Pylyshyn, 1988], and that therefore the core insights from AI research are irrelevant to how the biological brain functions. If true, this would mean that the results about how to create and manipulate structured symbolic representations and use these as the basis for goal-directed decisions would be immaterial to our understanding of the human mind. And in turn, this would raise serious questions about what we even mean by “copying” the software of the brain and ensuring that the simulated mind produces a “similar tokening of internal data structures”. In a digital system like a computer we can be very clear about what constitutes a successful copy of an algorithm, for an arbitrarily complex analog system this is less clear.

In fact however, the most empirically successful and sophisticated models of the human cognitive architecture stem directly from symbolic AI research and are fundamentally digital at their uppermost algorithmic level. Starting in the 1970’s the cognitive psychologist and AI researcher Allen Newell looked very seriously at the constraints that biological neurons put on the types of algorithms that AI researchers were proposing at the time. His answer to these constraints was to suggest that the brain actually uses a production system framework to carry on its symbolic computations [Newell, 1973]. In a production system there is a structured-symbolic working memory representation that holds all information about the current state of the program. At each step of a computation, this working memory representation is pattern-matched against a set of “production rules”. Each production rule is an
“if-then” statement whose “if”-clause looks for particular (syntactically specific) patterns in the current contents of working memory and if a match is detected its “then”-clause “fires” producing a specific change to the contents of working memory. Unlike a program written in a traditional software language like C, a production system program consists of an unordered set of possibly thousands of production rules which are all pattern matched against working memory simultaneously and in parallel at each step of computation. This was the key insight since even though the brain’s circuitry is too slow to support algorithms with tens of thousands of **sequential** steps, it seems perfectly tailored to allow the **parallel** matching of tens of thousands of patterns simultaneously.

Newell and colleagues showed that such a neurologically-plausible production system framework could support AI algorithms for general problem solving and that a host of complex human skills and learning behaviors could be modeled within the framework [Newell, 1990]. The results of this research have guided the field of cognitive science ever since and today’s most sophisticated and empirically supported model of the human cognitive architecture — the ACT-R model — is a direct decedent of this production system research.

Simply put, the ACT-R model is the most comprehensive theory available today for how the human mind and brain function at an algorithmic level [Anderson et al., 2004; Anderson, 2007]. It has been used in literally hundreds of cognitive science publications from dozens of laboratories to model behaviors ranging from simple stimulus response, visual perception and attention, and accuracy and timing of memory recall, to natural language understanding, problem solving and decision making, and the learning of complex skills [Anderson, 2011]. It would be impossible to fully describe the ACT-R model here but a basic understanding of what it says about the composition of the human cognitive architecture is crucial to any intelligent discussion about the possibility of human mind uploading.

The ACT-R model proposes that the human brain is a production system as Newell suggested. Cortical systems like the ventral visual stream process raw sensory information into patterns of neural activity over specific high-level sensory cortical areas (e.g., the lateral occipital complex) which act like buffers in a global working memory representation of the environment. Because of the highly processed nature of these sensory representations they act as symbols (e.g., I see a coffee cup) rather than collections of basic features. Other cortical areas (e.g., certain parietal lobe and frontal areas) also act as buffers holding symbols in the global working memory workspace, but these buffers are filled as a result of internal computations as opposed to sensory input. There are also global working memory buffers which connect to cortical and subcortical motor control systems. The high-level symbols in these buffers represent primitive motor actions like “pick up the object” or “saccade to the red object” whose contents are eventually interpreted by downstream motor centers to produce finely-timed sensory-motor behaviors.

This network of cortical buffers is the global working memory workspace for the brain and it is looked at by pattern matching circuits within the basal ganglia. The
ACT-R model of the basal ganglia postulates that it encodes the full complement of the “production rules” that we have acquired over our lifetime. The basal ganglia act as a straightforward pattern associative memory, and if a pattern is recognized (i.e., if a production fires) then it activates an associated set of fibers projecting from the basal ganglia (and other structures) back to the cortex to overwrite or route patterned activity between cortical buffers [Stocco et al., 2010]. Every hundred milliseconds or so the basal ganglia perform such a match-and-route operation, and the sequential result of many of these operations explains all of our actions from the most basic learned reactions (e.g., hit the brakes when you see a red light) to the most sophisticated (e.g., reasoning through a new scientific theory).

Another central part of the ACT-R theory is the presence of a dedicated declarative memory storage and retrieval system. The hippocampus and temporal lobe structures can on-demand produce a snapshot of part of the global working memory buffers (via a form of Hebbian learning) producing a new episodic or declarative memory “chunk”. As part of any production firing, the production system (i.e., basal ganglia switching circuits) can request a retrieval of a declarative memory chunk based on presenting a partial pattern (e.g., “2 + 2 = ?” would retrieve the declarative memory chunk “2 + 2 = 4”).

Learning in ACT-R consists of forming new procedural memory productions in the basal ganglia and forming new declarative memories in the hippocampus and temporal lobe. There is also an analog component to this learning which can strengthen or weaken particular procedural and declarative memories based on experience, thus making them more or less likely to match or be retrieved.

That is the 10,000 foot overview of ACT-R — the most detailed and empirically supported model of the human cognitive architecture we have available to us today. The key thing for us to note is that the ACT-R model of the brain suggests that our “software” is mainly digital. Working memory buffers hold discrete symbols (modeled as stable attractor patterns in the cortical neural buffers). Declarative memories are also to be thought of as containing discrete symbols (which in this case would be modeled as stable attractor states in the hippocampus and temporal lobe). And each production rule is also a discrete entity (in this case an associative pattern memory stored in the circuits of the basal ganglia along with specific routing circuits in the cortex).

This means that, according to ACT-R theory, we are justified in using our analogy that mind uploading is a matter of copying the mostly-digital program associated with an individual brain to another computational substrate. We can go further and precisely specify what this program consists of — it consists of the unique collection of production rules, declarative memory chunks, and perceptual and motor memories that the human has accumulated over their lifetime.

Furthermore, the ACT-R theory actually suggests neural substrates for each of these, allowing us to evaluate the difficulty of recovering these unique components of identity from a static brain scan. Neural models of how stable attractor states are stored in the hippocampus and cortex suggest that such attractor states are robustly and redundantly stored in the specific pattern and strength of synaptic connections...
between neurons [Anastasio, 2010]. Just as one can read off the full complement of stable attractors from a Hopfield network [Hopfield, 1982] if given its weight matrix (i.e., which neurons connect to which others and the strengths and signs of their connections), one should be able to do the same in a real biological network of neurons similarly constructed to store memories as stable attractor states.

These models imply that the full complement of discrete stable attractor states of a network of hippocampal or cortical cells should be determinable if it is known precisely which neurons synapse with each other and if their relative strengths can be determined. Can the relative strength of a synaptic connection be determined from an electron micrograph? Yes, electron micrograph studies of hippocampal synapses following the induction of memory via a theta-burst stimulation protocol show clear structural changes in synapses that are correlated with their strengthening [Bourne and Harris, 2011]. These results and others show that a synapse’s size and area of post synaptic density (both easily measurable in a $5 \times 5 \times 10$ nm volume electron image) are well correlated with its measured strength. Similarly, associative neural network models (of the type assumed present in the basal ganglia) suggest that production matching rules would again be robustly and redundantly stored in the specific pattern and strength of synaptic connections between neurons and should thus be determinable by electron micrographs.

In summary, our best model of the human cognitive architecture predicts that a particular human being’s unique “software” consists of a discrete set of production rules, declarative memory chunks, and perceptual and motor memories learned throughout their lifetime. These “atomic components of thought” [Anderson, 1998] are robustly and redundantly encoded across the synaptic connections of cortical and subcortical brain areas in a manner that should be retrievable from a static $5 \times 5 \times 10$ nm volume electron image. A successful mind uploading of a particular human being would consist of creating a computer simulation of the brain with just enough fidelity that the same discrete switching events (production firings events, declarative memory retrieval events, perceptual tokening events, etc.) occur that would have occurred in the original biological brain under similar circumstances. This would lead to similar external behavior as well as a similar tokening of internal data structures.

Would this simulation be conscious? And would it be the same consciousness as the original biological person? If the computational theory of mind is correct then the answer to both of these questions is yes. In the phenomenal self-model of consciousness [Metzinger, 2003] our first-person conscious experiences are a reflection of the particular stream of tokened data structures that our brain uses to model its own perceptions, goals, and actions. Using the ACT-R framework we can identify precisely where this phenomenal self-model resides in the brain. It is the collection of episodic and self-directed declarative memories that encode all of our knowledge about our own perceptions, goals, and actions, along with all of our production rules which create and operate on these self-directed declarative memories. When we are in the midst of a conscious perception it is these production rules which take the current stream of symbols in our global working memory workspace and uses them to create
additional episodic and self-directed declarative memory chunks incorporating them into our continuously evolving phenomenal self-model. A simulated brain which is based on the same declarative memories and production rules would “reboot” the same phenomenal self-model as the original biological brain, and would exhibit the same tokening of internal data structures (and thus modifications to the phenomenal self-model) that would have occurred in the original brain under similar circumstances. From this perspective continuity of the person’s stream of consciousness would have been restored.

3. High-Resolution Reconstruction of Neural Circuits

As just discussed, significant portions of the brain may be well characterized at the algorithmic level as symbol-containing buffers, pattern recognition circuitry, routing circuitry, etc., but at the implementation level the brain consists of hundreds of specialized regions each containing networks of millions of biologically complex neurons. And each brain region is composed of dozens of different neuronal types whose functional properties and connectivity have been finely tuned to ensure robust operation at the algorithmic level. One of the promises of the new field of Connectomics is that mapping the detailed connectivity of neurons at this implementation level will provide a comprehensive understanding of how the algorithmic level functioning is generated.

A typical cortical pyramidal neuron may have a cell body that is only 10 microns in width but will send out an incredibly elaborate dendritic arbor spanning a cubic millimeter in volume. The finest branches of this dendritic arbor are thousands of dendritic spines whose tiny necks can shrink down to just a few tens of nanometers in diameter. This neuron will also send out a long thin axon that in places will shrink down to just a few tens of nanometers in diameter, but will travel many millimeters before branching and arborizing to form synaptic contacts with distant neurons. Any random cubic millimeter of cortical tissue will contain tens of thousands of neurons with overlapping dendritic arbors and will receive tens of thousands of axonal projections whose arbors also overlap. And within this cubic millimeter of densely tangled jungle will be on the order of a few hundred million synaptic connections each with a tiny area of contact on the order of 500 nm in diameter [Shepherd, 1990]. Staining a single neuron (as in the Golgi method) can reveal the gross structure of that single neuron in a light micrograph, but to actually map the connectivity of the entire neural network — determining precisely which neurons synapse on which others — requires high-resolution volume electron microscopy on a scale that has never before been attempted. The field of Connectomics has been formed to do just that by developing new automated methods for obtaining enormous numbers of serial electron micrographs from a volume of neural tissue, and by developing new computer algorithms that can efficiently process the resulting terabytes of image data to automatically reconstruct each neuron’s dendritic and axonal arbors and find where they make synaptic contacts with other neurons.
As mentioned above, the diameter of neuronal processes routinely shrink to less than 100 nm; for example, dendritic spine necks and fine axons can shrink down to \(~40\) nm. With standard tissue fixation and embedding protocols the membranes of these tubular structures are made electron dense, hence to resolve the tubular nature of these neuronal processes one requires resolutions on the order of 10 nm or less. Today both the transmission electron microscope (TEM) and the scanning electron microscope (SEM) can easily achieve such resolution while providing a signal to noise ratio sufficient to trace the finest neuropil in osmium fixed, heavy metal stained tissue. However, until relatively recently only the TEM was used for tracing neuronal circuits. This reliance on TEM has been a major roadblock to attempts at large-scale automation since TEM requires that sections be physically cut thin enough (<100 nm) for electrons to pass through and that they be mounted on gossamer thin plastic films throughout the imaging process. If scanning, as opposed to transmission, electron microscopy could be used it would open up many more possibilities for robust automation since then at least the thin sections could be collected and imaged on a thick sturdy substrate [Hayworth, 2008], but before the widespread introduction of high-brightness field emission electron sources (as opposed to tungsten thermionic sources) one simply could not achieve SEM electron probes of small enough diameter (5 nm) and high enough current to allow quality imaging of neural tissue [Bogner et al., 2007; Joy, 1991].

In 2004, Denk and Horstman showed that this reliance on TEM for tracing neural circuits could in fact be overcome. They published a seminal paper demonstrating the SBFSEM (Serial Block Face Scanning Electron Microscopy) method which could robustly automate the process of obtaining a series of electron micrographs from a block of neural tissue [Denk and Horstman, 2004]. Using a high-brightness field emission SEM equipped with a low-energy backscatter electron detector they first showed that one could obtain high-resolution images directly from the face of the tissue block (thus eliminating the need to collect ultrathin sections). Then, following an original design from Leighton [1981], they built an ultramicrotome into the vacuum chamber of the SEM which would repeatedly scrape (using an extremely sharp diamond knife) 50 nm layers of material off the surface of the tissue block while the SEM was used to image each freshly revealed block face at high resolution. The result was a fully automated method to volume image a block of neural tissue. They and others have continued to refine this technique so that now it can now achieve 23 nm section thickness and have successfully applied the technique to a reconstruction of the direction selective circuitry of the mammalian retina [Briggman et al., 2011].

Both section thickness and lateral resolution are limited in the SBFSEM technique due to the fact that the physical scraping process for removing plastic-embedded block face material becomes unreliable if the material properties of the embedding plastic are degraded due to long exposure to the electron beam during imaging. This means that, in general, a finer lateral resolution will require a greater electron beam dose to the block face and will thus make scrapping a very thin layer of tissue from the block face less reliable. This physical scraping process also makes the technique
very sensitive to the embedding properties of the tissue block — something that could make application to very large tissue volumes difficult.

In 2008, Graham Knott et al. introduced the FIBSEM (Focused Ion Beam Scanning Electron Microscopy) technique for tracing neural circuits [Knott et al., 2008]. This technique is also known as Ion Abrasion Scanning Electron Microscopy [Heymann et al., 2009]. FIBSEM works similarly to SBFSEM, but instead of physically scraping a thin layer of material off the block face with a diamond knife, the material is instead ablated away using a focused beam of gallium ions. This change overcomes the lateral resolution limitations of the SBFSEM since the FIB ablation process is relatively insensitive to the material properties of the embedding plastic and therefore a much larger electron dose can be used during imaging. What’s more, ion beams can be tightly focused achieving spot sizes in the 10 nm range. Aligning the ion beam parallel to the block face so that it just grazes the surface allows reliable and rapid removal of extremely thin layers less than 10 nm in thickness. What actually limits the resolution of the FIBSEM (for optimally-sized blocks) is the depth of penetration of the imaging electrons into the surface of the block [Muller-Reichert et al., 2010] and by using very low voltages (<2 kV), along with sensitive low-energy backscatter electron detectors with energy filtering, recent reports [Knott et al., 2011] have demonstrated depth resolutions in the range of ~10 nm with lateral imaging resolutions of 5 × 5 nm. This FIBSEM voxel resolution of 5 × 5 × 10 nm is more than sufficient to reliably resolve the finest neuronal processes and synapses, and should be sufficient to allow for extremely reliable automated reconstruction algorithms [Lucchi et al., 2010]. Recently Knott has published a Journal of Visualized Experiments online video article covering the entire FIBSEM procedure from tissue processing to automated FIBSEM milling and imaging, demonstrating a final volume image of a piece of cortex imaged with 5 × 5 × 5 nm voxels [Knott et al., 2011]. It is stunning how clearly demarcated the neuronal processes and synapses are in the resulting FIBSEM volume video.

There is, however, a fundamental physical limitation of the FIB ablation process — this slice resolution can only be obtained for very small samples on the order of a few tens of microns across. One cause of this is that the electrostatic lenses of the FIB are able to tightly focus the beam of gallium ions only over a short distance (i.e., its depth of focus range), and outside of this range the beam’s shape and its milling efficiency changes. This means that to achieve uniform removal of 10 nm of surface material the block being milled should be only a few tens of microns wide in the direction of the ion beam.

A straightforward solution to this sample size limit is to simply subdivide a large tissue block into smaller sub-blocks of optimal (~20 micron) size for FIBSEM imaging. However, if the goal is to trace neural circuits across the original larger volume, this subdivision process must be such that absolutely no material is lost or damaged during subdivision so that all of the tiny 40 nm diameter neural processes remain traceable across the subdivided faces. This requirement means that such subdivision cannot be performed prior to plastic embedding (for example by using a tissue
vibrotome) since sectioning wet, soft, unembedded tissue will distort and destroy the fine neuronal processes surrounding the cut.

To produce tissue surfaces with nanometer-scale smoothness one typically must use a diamond knife polished to a radius of just a few nanometers, and must cut tissue that has been infiltrated both extracellularly and intracellularly with a plastic resin so that the whole of the tissue block is rigidified to withstand sectioning and to prevent cutting-induced distortion of fine cellular processes at the cut surfaces. Such diamond knife sectioning is used in the SBFSEM to scrape very thin layers off the block face, and in obtaining ultrathin sections for TEM imaging as described above. However, such ultrathin (<100 nm) sections would be very difficult to reliably obtain in large numbers and to image reliably in the FIBSEM. A more optimal sectioning thickness (for producing sub-blocks of tissue ready for high-resolution FIBSEM imaging) would seem to be closer to the maximum thickness allowed by the FIBSEM milling process itself (i.e., ~20 microns). Unfortunately, if one simply attempts to use a diamond knife (which typically must have a large edge angle) to section such a 20 micron thick slab, the differential cutting stresses on the section as it bends are enough to crumple and crack the tissue and severely damage the surfaces.

I recently performed a set of experiments to determine if there were cutting parameters which would allow for lossless 20 micron thick sectioning of plastic-embedded brain tissue. Adapting a trick from McGee-Russell [McGee-Russell et al., 1990], I used a slightly heated (~60°C) diamond knife to section plastic-embedded brain tissue blocks. This heating temporarily and locally softens the plastic during sectioning, reducing the local cutting stresses while maintaining overall block rigidity. I also found that I had to use an oil lubricant during the sectioning process to prevent surface damage due to the thick section scraping along the knife face. After a bit of trial and error I determined a set of parameters that could indeed reliably produce 20 micron thick brain sections with such high-quality surfaces that the finest neuronal processes could be traced from one face of a thick cut to its matching face. This was verified by taking electron micrographs of both faces and then mirror-reversing one micrograph and aligning it to the other. I also verified that the neural tissue within these thick cut sections was still intact by performing a serial section volume reconstruction of several neuronal processes across two such thick cut sections [Hayworth et al., 2010; Hayworth, in preparation].

In summary, the FIBSEM technique represents a truly enabling technology for mapping neural circuits with 100% reliability. It can achieve voxel resolutions of at least 5 x 5 x 10 nm, sufficient to allow straightforward algorithms for computer automated tracing of all neuronal processes and identification of all synaptic connections [Merchán-Pérez et al., 2009] along with the morphological parameters which are correlated with their strength (e.g., area of contact and size of post synaptic density). Furthermore, its use of a focused ion beam (as opposed to traditional physical sectioning with a diamond knife) gives it the potential to achieve the very high reliability levels needed for large-scale automation. A serious limitation of today’s FIBSEM systems is that they can achieve such high-resolution images only
over tiny volumes — typically on the order of a few tens of microns across. However, this limitation can be relatively easily overcome with the use of a lossless subdivision technique like the one described above. And the availability of such a lossless subdivision technique opens up the possibility of massively parallel FIBSEM imaging, something that would, in any case, be necessary in order to image any relatively large volume of tissue in a reasonable amount of time.

4. Whole Brain Ultrastructure Preservation and Plastic Embedding

Electron imaging of brain tissue at the resolutions necessary to trace neural circuits requires that the tissue’s membranes and other key subcellular structures be selectively stained with heavy metals (osmium, uranium, and lead typically). The dense nuclei of these heavy metal atoms deflect the electrons of the SEM’s beam back at the electron “backscatter” detector allowing the stained membranes to stand out against the unstained background of the cytoplasm. Electron imaging of brain tissue also requires that this tissue be structurally stabilized by removing all water from within the tissue and infiltrating every nook and cranny of extracellular and intracellular space with a curable plastic resin. This is not only to rigidify the tissue for sectioning (as described above), but also to stabilize the tissue so that it can withstand the intense radiation bombardment of the electron beam.

Today’s most typically used tissue preparation protocols for electron microscopy are only able to prepare volumes of less than 1 mm$^3$. An animal’s vascular system is perfused through the heart with a mixture of paraformaldehyde and glutaraldehyde. The paraformaldehyde quickly stops cellular degradation and, at a slightly slower rate, the glutaraldehyde provides stronger crosslinks to fix proteins in place. If performed carefully, such perfusion through the vascular system is able to quickly fix the entire nervous system of the animal since every cell is within a few tens of microns of a capillary. However, following this fixative perfusion step the brain of the animal is removed and a very small piece dissected to undergo the remaining tissue preparation steps — which are typically all performed by simple immersion in chemicals. These steps include immersion in osmium tetroxide (to fix membrane lipid molecules in place), immersion in heavy metal staining solutions (e.g., uranyl acetate), immersion in a graded series of alcohols (to remove water from the tissue), and finally immersion in a plastic resin dissolved in an organic solvent (to completely infiltrate the tissue with the heat-curable resin). Because these steps are performed by simple immersion (and thus diffusion) the process fails if attempted on blocks larger than 1 mm$^3$. The result is destruction of tissue ultrastructure and poor staining in the depths of the block.

Obviously such volume limitations must be overcome before human mind uploading can be attempted. Although it has never been demonstrated that a whole mammalian brain can be preserved at the ultrastructure level for electron microscopic imaging, there are many results that suggest that a protocol could be developed to do just that. Perfusion fixation with osmium tetroxide has been demonstrated on whole
brains [Palay et al., 1962], and serial vascular perfusion first with glutaraldehyde, followed by osmium tetroxide and uranyl acetate, and finally by an alcohol dehydration series has been demonstrated on whole organs [Bachofen et al., 1982; Oldmixon et al., 1985] demonstrating sufficient preservation to allow ultrastructure studies. Plastic infiltration of whole mouse brains has also been demonstrated [Mayerich et al., 2008], and there are recent reports of the development of a full ultrastructure fixation, staining, and embedding protocol for the mouse brain for use in the mapping of long distance axon trajectories via serial block face SEM [Mikula et al., 2011]. There is even a challenge prize being offered for the first demonstration of such ultrastructure preservation across an entire large mammalian brain [Hayworth, 2011]. Given these results and the fact that the bourgeoning field of Connectomics will require larger and larger volumes of high-quality prepared tissue, it is likely that a protocol to preserve and plastic-embed an entire human brain at the ultrastructure level will be perfected relatively soon.

5. Industrial-Scale Thick Sectioning for Random-Access, Massively Parallel FIBSEM Imaging

To review, the FIBSEM technique can already achieve voxel imaging resolutions sufficient to allow the computer automated tracing of neural circuits and the identification and strength estimation of synaptic connections. In order to apply this technique to a whole human brain, the brain must first be chemically fixed and plastic embedded. It must then be subdivided into sub-blocks of optimal size for FIBSEM imaging — each approximately 20 microns in width. Above, I described a method (using a heated, oil-lubricated diamond knife) which can losslessly subdivide plastic-embedded brain tissue into such 20 micron thick slabs, but it is an open question whether such a technique could be successfully applied to anything as large as an entire human brain. Here I will briefly sketch out an industrial-scale process that might accomplish such a feat. The general plan is to section a human brain into over 30 million “tissue pillars”, each 20 microns wide × 20 microns deep × 180 mm long. At each step in the sectioning process the tissue to be sectioned is first securely adhered to a collection tape prior to being sliced off by the heated diamond knife — in this way the tissue slices are always robustly controlled. As a result, the final tissue pillars are securely adhered to the edge of a sturdy tape ready for high-resolution FIBSEM imaging. This tape-collection methodology is loosely based on the automatic tape-collection technique I developed for collecting ultrathin sections for electron microscopic imaging [Hayworth and Hayworth, 2004; Hayworth, 2008].

A plastic-embedded human brain is first mounted on a gantry-style Thick-Strip Sectioning and Tape Collection Machine (Fig. 1(a)). This device has two diamond knives (both heated and oil-lubricated) which are used in tandem to section the brain into a tape containing 3 mm wide 20 micron thick tissue strips. First a heated “furrowing” knife is used to create a set of losslessly-cut furrows in the block face at
3 mm intervals (Fig. 1(b)). Each furrow goes a little deeper than 20 micron into the face of the plastic block. Following this, a heated platen is used to adhere (using a thin layer of heat-activated adhesive) a 3 mm wide collection tape to the block face, aligning this tape precisely between two furrows. Once adhered, the entire platen and tape reel assembly move forward while a second heated diamond knife slices off the 3 mm wide 20 micron deep strip of brain tissue immediately under the tape (Fig. 1(c)). This results in a brain strip (180 mm in length) being reeled up securely adhered to the collection tape. Figure 1(d) shows the relation between the furrow, tissue strip, and block face-adhered collection tape.
the entire 100 mm deep brain block is reduced to set of “thick-strip tapes” containing a total of 250,000 tissue strips.

These 20 micron thick tissue strips are still not in a form that will allow efficient high-resolution FIBSEM imaging. To allow this, they must undergo additional sectioning into $20 \times 20$ micron pillars. The thick-strip tape is first adhered to a semi-rigid backing tape using a Semi-rigid Backing Applicator machine (Fig. 2(a)). This is done to make it sturdy enough to orthogonally section in a reel-to-reel fashion. Then this tape is put into a Pillar Sectioning and Tape Collection Machine (Fig. 2(b)). It is within this machine that the original 3 mm wide strips are repeatedly sectioned in the orthogonal direction creating the final $20 \times 20$ micron $\times 180$ mm long tissue pillars, each securely adhered to the edge of a new collection tape.

This pillar tape is now in the optimal configuration to allow high-resolution random-access FIBSEM imaging. A particular pillar tape is mounted in a reel-to-reel FIBSEM system as shown in Fig. 3. The tissue pillar to be imaged is first reeled into position under the electron beam and clamped in place. The SEM then takes a high resolution ($5 \times 5$ nm) image of the top face of the tissue pillar over an area of approximately $20 \times 20$ microns. Following this, the side-directed focused ion beam is used to ablate away the top 10 nm of material that was just SEM imaged. This process of SEM imaging and FIB surface ablation is repeated 2000 times to image an approximately $20 \times 20 \times 20$ micron cube of tissue.²

Perfecting such an industrial-scale brain sectioning scheme would no doubt be expensive and difficult, but its engineering tolerances are actually quite reasonable. This is because the required sectioning thickness is 20 microns which is 2–3 orders of magnitude thicker than the sections used for serial TEM reconstruction or in the tape-collection for SEM process [Harris et al., 2006; Hayworth, 2008]. I would estimate that a system to reduce a whole brain into FIBSEM-imageable pillar tapes could be developed and built with mainly off-the-shelf components for around $10$ million.

It is the FIBSEM imaging requirements that actually represent the most difficult and expensive investment for any large-scale imaging project along these lines. Today’s commercial FIBSEM systems are very complex machines costing around $2$ million apiece, and their imaging rates are quite slow — on the order of 1 MHz typically. Also, current systems are not well suited for automation or for extended imaging times due to limited lifetime ion sources. To make practical large-scale imaging a custom redesign of the FIBSEM will be necessary. Optimizing the tissue staining, detector, and electron beam parameters should allow imaging rates of up to 10 MHz (a rate that has already been achieved in the tape-collection for SEM process). It should also be possible to redesign the system for quick setup of a tape pillar by directing the FIB beam at a scintillator detector during operation — thus providing closed loop control of the FIB milling parameters and milling rate during

²The actual volume imaged must be trapezoidal in shape, with a wider area milled off the top of the volume than the bottom. This is to prevent the milled sidewalls of the tissue from shadowing the areas to be imaged in lower slices. This is what is depicted in the figure.
imaging and setup. Such closed-loop control of milling should also greatly increase the reliability of the entire process.

Given such an optimized FIBSEM with an imaging rate of $\sim 10 \text{ MHz}$ (and accounting for setup and FIB milling time), a single $20 \times 20 \times 20$ micron cube could
be imaged every two hours. At this rate it would take almost 30 years for a full cubic millimeter volume to be imaged using a single machine. Clearly a single FIBSEM does not come close to meeting the requirements for imaging an entire $1 \times 10^6 \text{ mm}^3$ human brain. Even a fleet of 100 such machines operating over 10 years could only succeed in imaging 0.003% of the brain. However, such a fleet of 100 FIBSEM machines would make possible the creation of a “Human Connectome Observatory”, something of critical importance to the advancement of neuroscience, and something that I believe is the most crucial milestone on the road to true mind uploading.

6. A Human Connectome Observatory

The astronomical research community excels at the efficient pooling of research funds. No individual astronomer could afford to build the Keck Observatory
telescopes on the summit of Mauna Kea (~$200 million), or the Large Binocular Telescope in Arizona (> $100 million), or ESO’s Very Large Telescope Array in Chile (> $300 million). And they certainly could not afford to build the Spitzer Space Telescope (> $800 million), or the Chandra X-ray Space Telescope (> $1 billion), or the Hubble Space Telescope (> $3 billion). But any individual astronomical researcher can apply for observation time on any of these grand instruments at relatively reasonable rates. The fantastic gains in our understanding of the cosmos have in large part been a result of this strategy of upfront pooling of resources to build top-of-the-line instruments which are then shared by all.

The ability to section an entire human brain into a set of randomly accessible FIBSEM-imageable tapes would make possible a similar pooling of resources within the neuroscience community. The key is random-access. Because the brain has been reduced to a set of tapes, any 20 x 20 x 20 micron cube within the brain can be called up for imaging by: (1) Loading the proper tape into a FIBSEM machine, (2) Reeling the tape so that the desired tissue pillar and region is under the electron beam, (3) Raising the tissue pillar until the FIB scintillator detector indicates that FIB surface milling has begun, (4) Performing ~2000 FIB milling and SEM imaging steps on that pillar region until the entire 20 x 20 x 20 micron volume in that part of the pillar has been volume imaged at 5 x 5 x 10 nm voxel resolution. For roughly the cost of a top-of-the-line earth-based optical telescope observatory (~$200 million) a facility could be built containing a fleet of 100 FIBSEM machines coupled to robotic tape loaders which would allow any pillar tape to be loaded into any of the 100 FIBSEM machines. The result would be a “Human Connectome Observatory” in which the highest resolution imaging could be directed anywhere that is desired within a single human brain.

What could the neuroscience community do with such a Connectome Observatory? From the above, we assume that each FIBSEM can automatically load and image any desired 20 x 20 x 20 micron cube of tissue every two hours. This means that the entire 100 FIBSEM observatory would have a total imaging throughput of 50 cubes per hour, or equivalently it could image a volume of 1 mm³ every 100 days. Such an observatory could image 4.4 million cubes over a 10-year operational life, representing a total volume of 35 mm³. Again, the key is that these 4.4 million cubes can be distributed across the brain in any fashion desired. For example, a single tiny axon could be efficiently followed from its cell body in one region of the brain to its terminal synapses several centimeters away in another brain region by sequentially targeting imaging cubes only along its path. The Connectome Observatory could trace such fine axons at an average rate of about 1 mm/hour (i.e., a line of 50 cubes stretches 1 mm in length), and thus over its 10-year lifetime the Connectome Observatory could trace a total of ~90 meters of long distance axonal connections. It is more likely that the neuroscience community would collectively decide (in an open, proposal-driven process similar to the one used to govern time on astronomical observatories) to perform large dense reconstructions in key brain regions along with directed axon following to trace specific fibers between these dense reconstructions.
For example, key circuits of the human visual system could be mapped out by imaging a 0.5 mm$^3$ volume of the lateral geniculate nucleus (LGN), a 1 mm$^3$ columnar volume of the primary visual cortex (V1), and a 1 mm$^3$ columnar volume of extrastriate cortical area (V2), while strategically directing groups of imaging cubes along the axonal pathways between these regions. Such a strategy would produce fantastically detailed circuits maps of the individual regions (sufficient to determine all neuronal types and the details of their local circuit connections) while at the same time putting each of these dense reconstructions in context of each other by determining the origins of particular input axons and the destinations of particular outgoing axons.

Over a 10-year period the Connectome Observatory would allow researchers to perform large dense reconstructions (each 0.5 mm$^3$) in a total of 40 different brain regions while at the same time putting each of these dense reconstructions in context of each other and of the rest of the brain by allowing dozens of fine long-distance axons (totaling 40 meters worth) to be traced into and out of each of the dense reconstructions — all within the same brain. And, of course, the same infrastructure could be used to map neural circuits across the brains of other species (e.g., fly, mouse, cat, monkey).

Such a comprehensive and detailed understanding of a single brain’s circuitry is crucial to understanding how the brain works. It will allow us to test high-level algorithmic theories of brain function like ACT-R by literally looking for the brain-wide routing circuitry that it predicts. It will allow us to test theories of how memories are encoded by extracting the complete connection matrix of a volume of cortex or hippocampus and determining the stable attractor states of the network. It will allow us to test theories of visual feature extraction and binding by looking for the intra- and inter-regional circuits that particular visual theories predict. It will allow us to begin to perform test simulations of the whole nervous systems of small organisms (like Drosophila) along with functionally well-characterized regions of larger organisms (e.g., the mammalian retina).

7. “We Choose to Go to the Moon…”

I believe that the successful construction of a Human Connectome Observatory is the crucial milestone in the path to human mind uploading. It is also readily achievable. The expected scientific payoff of such a Connectome Observatory should easily justify a construction cost on the order of an earth-based astronomical observatory, especially given that this cost, and its scientific bounty, will be shared by the entire neuroscience community.

The Connectome Observatory’s scale is far smaller than what is needed for the ultimate goal of uploading a human mind, but it will demonstrate the brain preservation, large-scale lossless subdivision, and imaging technologies necessary for such an attempt. More important, it will enable a comprehensive vetting of our cognitive and neuroscience understanding of the brain in a way that will give us the confidence
to proceed to this ultimate scale. Just as Alan Shepard’s sub-orbital flight was a sufficient demonstration to justify proceeding with the full Apollo program, the Connectome Observatory should provide a sufficient demonstration that the goal of uploading a human mind is achievable.

Like Apollo, scaling up to a full brain will be an incredibly expensive and difficult engineering challenge. For example, a 10-year project for mapping a whole brain would require the equivalent of 3 million of the above FIBSEM imaging machines. To fit within an Apollo-size budget (over $100 billion in today’s dollars) mass manufacturing would have to bring each FIBSEM’s cost down to around $40,000 apiece. This is actually not unreasonable given that we are already seeing the introduction of high-resolution “solid state” SEMs whose electron optical columns are manufactured using the same photolithographic techniques used by the semiconductor industry [Rynne, 2009] — these SEMs are already priced in the $100,000 range. More likely such mass-producible electron columns will be combined with new technologies like multibeam SEMs which can achieve imaging rates in the 1000 MHz range by imaging with dozens of electron beams simultaneously [van Himbergen et al., 2007]. Like computer technology, electron imaging technology seems to be headed toward exponentially decreasing the cost of each megabyte of data acquired. Any large-scale investment in brain mapping is likely to greatly accelerate this trend and may, like the Human Genome Project, result in a situation where the second mind to be uploaded is done at a tiny fraction of the time and cost of the first.

References


