Measuring and Reconstructing the Brain at the Synaptic Scale: Towards a Biofidelic Human Brain *in Silico*

Joshua Vogelstein[1], Carey E. Priebe[1], Randal Burns[2], R. Jacob Vogelstein[3] , Jeff Lichtman[4]

[1] Dept. Applied Math & Stats, [2] Dept. of Computer Science, Johns Hopkins University
[3] National Security Technology Department, Johns Hopkins University Applied Physics Lab
[4] Dept. of Molecular & Cellular Biology, Harvard University

In the 100 years since Waldeyer-Hartz posited the Neuron Doctrine [1] and Golgi [2], Nissl [3], and Ramon y Cajal [4] published the first illustrations of individual neurons and their connections, experts in fields ranging from neuroscience to computer science have formulated detailed models of the brain to explain complex cognitive processes, create machines with artificial intelligence, prescribe psychoactive drugs to correct mental illness, and serve myriad other purposes. Amazingly, despite the fact that a fundamental assumption in all of these models is that the *connections* between neurons in the brain imbue it with its computational power, to-date no one has measured the complete connectivity of even a single vertebrate neuron! Therefore, previous large-scale brain simulations and emulations necessarily lack the potentially required detailed knowledge of neural connectivity [5]. Today, however, a combination of new technologies provides the opportunity to radically change this state of affairs by creating large-scale three-dimensional images of real brains at nanometer resolutions, sufficient to observe neurons, synapses, and even individual vesicles. In this poster, we present our work on measuring and reconstructing the brain at the synaptic scale and provide a roadmap from today's state of the art to tomorrow's biofidelic human brain *in silico*.

The key enabling technologies underlying the current and upcoming revolution in human brain emulation include: (i) high-resolution, high-throughout nanoscopy, (ii) automated machine annotation of three-dimensional images, (iii) exabyte-scale data management, and (iv) massively parallel chip design. Collectively, these techniques will generate biologically accurate neural network emulations, a reversely engineered brain. To make sense of these networks, though, we will also need advanced mathematics in graph theory and statistical inference to elucidate the recurring network motifs and identify the fundamental computing elements in the brain. We discuss the developments of each of these fields briefly below, along with our contributions.

The most powerful approach to studying the ultrastructure of the brain is serial electron microscopy. Technically however this method has been tedious and error prone and only one serious attempt – the decade long effort to get the connectional map of the ~300 neurons of a small worm was actually competed (White et al., 1986). A number of groups have thus begun to consider approaches to make this process more automatic. One notable advce was Denk and Horstmann's serial blockface approach which permitted large three-dimensional tissue nanostructure, with resolution sub 20 nm in two dimensions, and sub 50 nm in the third [6]. For a number of technical reasons related to a desire to image more quickly, at higher resolution, and without losing the opportunity to image the same data set at multiple resolutions we have approached this problem in a somewhat different way. Our current system combines a new Thin Section Scanning Electron Microscope (TSSEM) with an Automatic Tape-collecting UltraMicrotome (ATUM) [7]. Together, these tools allow us to collect image data at a resolution of 3x3x25 nm^3. Although imaging a whole human brain today is impractical, in development parallel beam electron microscopes and several new imaging modifications suggest that extraordinary imaging speeds (1-10 Gpixel/second) are within reach. At these speeds imaging a cortical column (approximately 1 mm^3), presumed to be fundamental building block for cortical computations, will be possible in several days and a rodent brain in less than 2 years, especially if the white matter tracks are imaged at lower resolution than the synaptic neuropil.

Once the image data is collected, it must be stored in a manner to facilitate image processing. This is a non-trivial problem, as the three-dimensional image of an entire human brain will require 3.3

exabytes of data. Converting the raw images into information about neurons and synapses will require non-local information outside of the field of view of a single image plane, because neural processes are often complex, tree structured, three-dimensional objects, sometimes spanning across both hemispheres. Furthermore, their relationships may only be evident after the entire structure is known. To address this issue, we are designing database structures specifically tuned for efficient non-local queuing of image data and simultaneous access by hundreds or thousands of processes to ensure that the machine annotation effort can be effectively parallelized.

Because manual annotation of images comprising a cortical column or entire brain would likely take many decades, a machine annotation strategy is necessary to extract the desired information. We have recently made a number of promising advances in machine annotation—including three-dimensional segmentation and tracking of neural processes [8]— but manual interaction is still necessary. Improvements in the less often reported but critically important stages of image pre-processing to align, stitch and stack the volumetric images are also necessary.

Once the data is collected, stored efficiently, and annotated, dedicated software and potentially hardware will be used to simulate the human brain. While the jury is still out on whether large-scale brain simulations are best implemented on dedicated chips, several of the current state-of-the-art large-scale brain simulations utilize massively parallel chips. In our work towards this goal, we have designed dynamically reconfigurable silicon chips, such that connectivity can be programmed into the chip, as more information comes online [10].

Although there is much work yet to be done, we are hopeful that neuroscience is on the verge of a revolution, poised for the first time to see actual biological neural networks and perhaps take the first steps to understand brain function at a fundamental level. This information will probably affect many fields of scientific inquiry and lead to countless new technologies including the first machines with true artificial intelligence, powerful neuromimetic computing systems, and the first biofidelic human brain emulation. Much like the sequencing the human genome allowed for rapid and unforeseen progress, becoming a central tool in genomics research, determining the complete wiring diagram of a mammalian brain could have a similar impact. At a minimum, it will provide constraints on the kinds of networks typical of brain circuits, upon which neuroanatomists, systems neuroscientists, and cognitive scientists can build new theories.

[1] von Waldeyer-Hartz, H. W. G. Ueber einige neuere Forschungen im Gebiete der Anatomie des Centralnervensystems. *Deutsche medicinische Wochenschrift,* 1891.
[2] Golgi, C. On the structure of the brain grey matter. *Gazzetta Medica Italiana*, 1873, 6: 244-246.
[3] Nissl, F. Ueber eine neue Untersuchungsmethode des Centralorgans zur Feststellung der Localisation der Nervenzellen. *Neurologisches Centralblatt*, 1894, 13: 507-508.
[4] Ramón y Cajal, Santiago. Comparative Study of the Sensory Areas of the Human Cortex. Clark University, 1899.
[5] de Garis, H. and Shuo, C. and Goertzel, B. and Ruiting, L. A world survey of artificial brain projects Part I: Large-scale brain simulations. *Neurocomputing*, in press.
[6] Denk, W. and Horstmann, H. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure . *PLoS Biology*, 2004, 2:e329.
[7] Hayworth, K. J., Kasthuri, N., Schalek, R. and Lichtman, J. W. 2006. Automating the Collection of Ultrathin Serial Sections for Large Volume TEM Reconstructions. *Microscopy and Microanalysis,* 2006, 12: 86-87
[8] Jain. V., Seung. H.S., and Turaga. S.C. Machines that learn to segment images: a crucial technology for connectomics. *Current Opinion in Neurobiology*, in press.
[9] Boahen, K. Neurogrid: emulating a million neurons in the cortex. *Engineering in Medicine and Biology Society, 2006.*
[10] R. J. Vogelstein, U. Mallik, J. T. Vogelstein, G. Cauwenberghs. Dynamically Reconfigurable Silicon Array of Spiking Neurons with Conductance-Based Synapses. *IEEE Transactions on Neural Networks*, 2007, 18(1):253-265.