Reverse engineering the cognitive brain

Gert Cauwenberghs[1]

Department of Bioengineering and Institute for Neural Computation, University of California, San Diego, La Jolla, CA 92093-0412

One of the greatest aspirations of the human mind has been to realize machines that surpass its cognitive intelligence. The rapid expansion in computing power, about to exceed the equivalent of the human brain, has yet to produce such a machine. The article by Neftci et al. in PNAS (1) offers a refreshing and humbling reminder that the brain's cognition does not arise from exacting digital precision in high-performance computing, but rather emerges from an extremely efficient and resilient collective form of computation extending over very large ensembles of sluggish, imprecise, and unreliable analog components. This observation, first made by John von Neumann in his final opus (2), continues to challenge scientists and engineers several decades later in figuring and reproducing the mechanisms underlying brain-like forms of cognitive computing.

Related developments are currently unfolding in collaborative initiatives engaging scientists and engineers, on a grander scale, in advancing neuroscience toward understanding the brain. In parallel with the Human Brain Project in Europe, the Brain Research through Advancing Innovative Neurotechnologies Initiative promises groundbreaking advances in enabling tools for revolutionizing neuroscience by developing nanotechnology to probe brain function at greatly increased spatial and temporal detail (3). Engineers are poised to contribute even further in revolutionizing such developments in neuroscience. In this regard it is helpful to relate the inquisitive nature of science—analysis—to the constructive power of engineering, synthesis. Despite fantastic feats of neuroscience in the analysis of the inner workings of neural and synaptic machinery down to the molecular scale, extending the level of understanding to something as complex as the human brain, not to mention its cognitive function, requires the power of synthesis in bridging across scales of analysis. Synthesis of complex function through hierarchical modular assemblies of successively more abstract representations is the forte of systems engineering, and provides a foundation for systems neuroscience in the multiscale investigation of the central nervous system (4). In his 1990 manifesto that launched the field of neuromorphic systems engineering (5), Carver Mead makes a com-

pelling case for such analysis by synthesis in reverse engineering neural circuits in silicon, drawing isomorphic parallels between modules representing various levels of neural computation in the brain and their emulation in silicon electronics, down to the fundamental physical level of Boltzmann statistics in ionic transport across lipid membranes, and electronic transport across similar energy barriers in metal-oxide-semiconductor transistors in the subthreshold regime (6) (Fig. 1A). In addition to supporting advances in systems neuroscience, experiments in neural analysis by synthesis using silicon offer tremendous side benefits to the engineering of extremely low-power miniaturized devices. By emulating functional structure of their biological counterparts and approaching their energy efficiency in sensory processing and computing, these neuromorphic devices can operate more effectively and more naturally in their surroundings (7). Some examples of recent feats of neuromorphic systems engineering—just to name a few—include silicon retinae and cochleae seeing and hearing the world through our senses (8), silicon cortical models running at speeds greater than real time (9), and synapse arrays running cool at nominal energy efficiency on par with that of synaptic transmission in the human brain (10).

However, demonstration of machine intelligence at the level of human cognition has remained elusive to date. A deeper look into the complexity of cognition helps to shed some light on the apparent challenges (Fig. 1B). Cognition is considered here, for the sake of the argument, as decision making by a motivated agent acting in the context of a given environment, such as a chess player making moves on the board. The cognitive task complexity, accounting for all possible states of the environment reachable by the agent, suffers from exponential scaling in the depth onto breadth of decision making, and quickly grows to astronomical proportions for any but relatively simple tasks. Efficient tree-search algorithms, capable of exhausting the search space using variants on Bellman's principle of optimality in dynamic programming (11), are capable of tackling relatively complex problems, such as games like chess, but at a significant cost by expending computing resources that scale
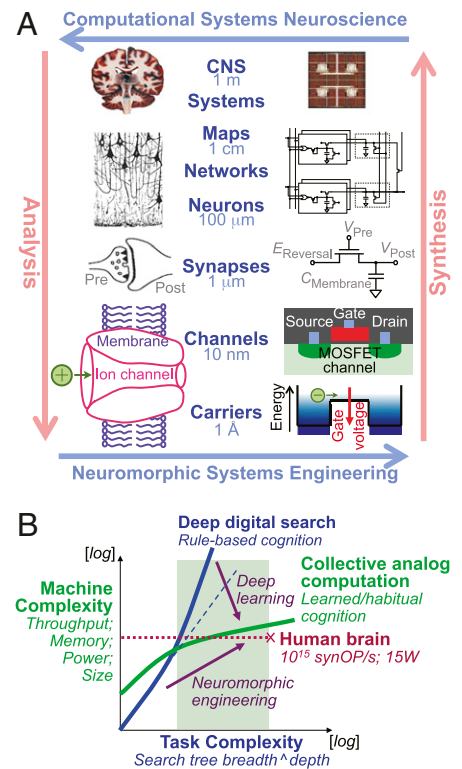
Fig. 1. (A) Multiscale levels of investigation in analysis of the central nervous system (adapted from ref. 4) and corresponding neuromorphic synthesis of highly efficient silicon cognitive microsystems. Boltzmann statistics of ionic and electronic channel transport provide isomorphic physical foundations. (B) Conceptual scaling of machine complexity with task complexity for a digital rule-based cognitive agent performing symbolic deep search, and a neuromorphic cognitive agent performing analog collective computation acquired through deep learning, targeting human cognitive performance. The shaded region indicates the desirable regime of high task complexity and lower machine complexity where the neuromorphic cognitive agent is expected to outperform symbolic digital alternatives.

almost linearly with task complexity. Such linear scaling is a fundamental limitation when having a need to sample large portions of the state space, a consequence of exact symbolic reasoning in search. In contrast, our brains execute such essentially sequential logic operations with substantial difficulty because of the need to dynamically instantiate a heap of nested working memory (2, 12). As such, it is not surprising that

computers easily outperform humans, not to mention their neuromorphic avatars, in tasks involving deep search in unstructured data, such as traversing a maze for the first time, or searching for a document posted on the Internet by specific keywords. The critical difference, however, is that much of the world we perceive is highly structured, and our brains excel at learning structure from sensory data, within their context, with remarkable efficacy and efficiency in generalization and dimensionality reduction (4). Thus, we may expect neuromorphic models of brain-line cognitive computing to get the upper hand over conventional von Neumann digital computing for problems dealing with naturally structured data in very high dimensions, such as high-level visual cognition, where the linear scaling between task and machine complexity implied by the conventional digital-state machine implementation becomes prohibitive.

Thus, the exciting and motivating premise for neuromorphic engineering is that by taking the necessary bold step of reaching into the vastly unexplored regime of high task complexity, not only may it become feasible to unravel the mysteries of the cognitive brain, but do so with great impact at full scale, and with efficiency unreachable by conventional computing approaches. What does it take to get us there? For one, it is unlikely that a neuromorphic avatar can accomplish tasks of complexity at the human brain level without expending at least as many resources: the equivalent of roughly 10^{12} neurons and 10^{15} synapses, which if not strictly needed would have been pruned through evolution for greater fitness in size and metabolic use. Fortunately, the extreme density and energy efficiency of the human brain, at a mere 0.002 m^3 volume and 20 W power consumption, is within reach of current advances in neuromorphic silicon nanotechnology—at 10^{11} synapses per square centimeter (13) and 1 fJ energy per synaptic operation (10)—for applications of embodied cognition on active mobile or implanted platforms.

Clearly, merely scaling up neuromorphic hardware to numbers of neurons and synapses on par with the human brain is not sufficient to warrant its cognitive intelligence. In their article, Neftci et al. (1) make a critical contribution in demonstrating that elements of context-sensitive cognitive behavior arise in neuromorphic systems that are modularly structured with reliable soft winner-take-all (WTA) dynamics. These neuromorphic "soft-state machines" are fundamentally different from digital finite-state machines with corresponding linear scaling in task complexity (Fig. 1B). These soft state machines represent the equivalent of cognitive "habits" that are effortlessly recalled through learned soft transitions between distributed states of working memory,

bypassing the excessive mental effort required to dynamically instantiate heaps of nested working memory in cognitive "rule"-based symbolic reasoning (12). In the implementation of Neftci et al. (1), selective amplification and signal restoration inherent in soft WTA networks (14) provide stable modules of working memory for neural state-encoding of context, whereas cognitive habits of input-dependent transitions between these contextual states are induced by sparse synaptic connections across the soft WTA modules. Although the proof-of-concept in Neftci et al. (1) by design targets the low-end of the task complexity scale, the concepts readily extend across the multiscale hierarchy of neural representations (Fig. 1A). Heteroclinic transitions between metastable states of working memory (15) may thus recur at various levels in the hierarchy, modulated by lower-level cognitive processes. Such a hierarchical modular structure is necessary for a cognitive agent to learn habitual responses that combine to complex nested cognitive behaviors (12). Hierarchical, deep learning (16) is equally essential for such cognitive behavior to be efficiently embedded in compact neural code. Indeed, efficient learning across multiple scales is key to dimensionality reduction in minimizing machine complexity for a given task complexity (Fig. 1B). By virtue of Occam's razor, such reduced representations also tend to offer superior generalization, further benefiting cognitive performance.

Our brain offers an existence proof that assemblies of imprecise and unreliable analog circuit components are capable of producing highly reliable and resilient, if not reproducible, cognitive behavior. The open question is to what extent such behavior could also emanate, perhaps more efficiently, from carefully crafted functional abstractions in more traditional computer architecture

(2). Recent trends in neuromorphic design toward large-scale cortical models of cognitive computing favor a hybrid approach that combines highly efficient analog continuous-time emulation of dendritic synaptic integration and neuronal excitability (gray matter), with highly flexible digital routing of action potentials along axon fiber bundles (white matter), offering reconfigurable long-range synaptic connectivity dynamically instantiated in memory tables (9, 17, 18). Purely digital alternatives are also surfacing (19, 20), offering comparable energy efficiencies in the picoJoule per spike range, but abstracting the continuous-time analog nature of neuronal and synaptic function by digitally accumulating and transforming spikes. Time will eventually tell the most favorable mix of physical realism and digital abstraction in the cognitive race along the task-complexity axis (Fig. 1B), but recent developments in deep learning (16) bolster the fundamental perspective of neuromorphic engineering that Boltzmann statistics are an essential physical foundation to computational intelligence. What better substrate for emulation does silicon offer than the innate Boltzmann statistics of electrons and holes moving across channels of field effect transistors, akin to sodium, potassium, and other ions moving through ion channels across the cell membrane? As Carver Mead told us all along: "Listen to the technology and find out what it's telling you" (5). Before long, and with a sufficient dose of providence and persistence, we may all immerse in the beautiful sound of a silicon symphony emerging through a concerted effort of scientists and engineers, all tuned to the dual goals of holistic understanding of what the brain is telling us, and leveraging this understanding to newly achievable levels of health and intelligence.

1 Neftci E, et al. (2013) Synthesizing cognition in neuromorphic electronic systems. *Proc Natl Acad Sci USA* 110(37):E3468–E3476.
2 von Neumann J (1957). *The Computer and the Brain*. (Yale Univ Press, New Haven, CT).
3 Alivisatos AP, et al. (2013) Nanotools for neuroscience and brain activity mapping. *ACS Nano* 7(3):1850–1866.
4 Churchland PS, Sejnowski TJ (1992) *The Computational Brain* (MIT Press, Cambridge, MA).
5 Mead CA (1990) Neuromorphic electronic systems. *Proc IEEE* 78(10):1629–1636.
6 Mead CA (1989) *Analog VLSI and Neural Systems* (Addison-Wesley, Reading, MA).
7 Indiveri G, et al. (2011) Neuromorphic silicon neuron circuits. *Front Neurosci* 5:73.
8 Liu S-C, Delbruck T (2010) Neuromorphic sensory systems. *Curr Opin Neurobiol* 20(3):288–295.
9 Schemmel J, et al. (2010) A wafer-scale neuromorphic hardware system for large-scale neural modeling. *International Symposium on Circuits and Systems, ISCAS 2010* (Institute of Electrical and Electronic Engineers), pp 1947–1950.
10 Karakiewicz R, et al. (2012) 1.1 TMACS/mW fine-grained stochastic resonant charge-recycling array processor. *IEEE Sens J* 12(4):785–792.
11 Bellman RE (1957) *Dynamic Programming* (Princeton Univ Press, Princeton, NJ).
12 Dayan P (2008) Simple substrates for complex cognition. *Front Neurosci* 2(2):255–263.

13 Kuzum D, Jeyasingh RG, Lee B, Wong HS (2012) Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. *Nano Lett* 12(5):2179–2186.
14 Indiveri G, Chicca E, Douglas RJ (2009) Artificial cognitive systems: From VLSI networks of spiking neurons to neuromorphic cognition. *Cogn. Comput.* 1:119–127.
15 Rabinovich MI, Huerta R, Varona P, Afraimovich VS (2008) Transient cognitive dynamics, metastability, and decision making. *PLOS Comput Biol* 4(5):e1000072.
16 Salakhutdinov R, Hinton G (2012) An efficient learning procedure for deep Boltzmann machines. *Neural Comput* 24(8): 1967–2006.
17 Serrano-Gotarredona R, et al. (2009) CAVIAR: A 45k neuron, 5M synapse, 12G connects/s AER hardware sensory-processing- learning-actuating system for high-speed visual object recognition and tracking. *IEEE Trans Neural Netw* 20(9):1417–1438.
18 Yu T, et al. (2012) 65k-neuron integrate-and-fire array transceiver with address-event reconfigurable synaptic routing. *Biomedical Circuits and Systems, BioCAS 2012* (Institute of Electrical and Electronic Engineers), pp 21–24.
19 Jin X, et al. (2010) Modeling spiking neural networks on SpiNNaker. *Comput Sci Eng* 12(5):91–97.
20 Merolla P, et al. (2011) A digital neurosynaptic core using embedded crossbar memory with 45pJ per spike in 45nm. *Custom Integrated Circuits Conference, CICC 2011* (Institute of Electrical and Electronic Engineers) pp 1–4.