Integrative Brain-Scale Computing

A Straw-Man Grand Challenge Proposal

Presented at the European Union's Future and Emerging Technologies FP7 Workshop on
"Grand challenges for basic research in emerging ICTs"

Doug DeGroot
LIACS, Leiden University
http://www.liacs.nl/~degroot • degroot@acm.org

31 March 2004

INTRODUCTION:

Advances in neuroscience continue at unprecedented rates. Further, the extraction of spinoff technologies from these advances into other fields has proven quite successful, and it promises to remain so for a very long time. In many ways though, the rapid pace of advancement has served to compartmentalize the activities within and across these fields, largely as a consequence of exploding information flows across the multiple domains and levels. Additionally, the fields of neuroscience, computational neuroscience, psychology, and computer science largely continue to proceed separately; cognitive and neural models remain largely disjoint and often even orthogonal; and the relations between brain structures, brain processes, and mind, especially as they relate to emergent, dynamic phenomena, remain largely unexplored. Thus while monumental successes have been achieved in many components of the neurosciences, little progress has been made toward developing a unified theory of the brain.

Many potential sources of this failure can be identified, including: the sheer complexity of the brain, the credit assignment problem, the scarcity of cross-field research scientists, the atomistic or reductionist approach to science, and the like. One could even point to the current, ubiquitous desktop computer paradigm, since current desktop software, modeling tools, visualization packages, and the like are quite slow and isolationist, by any reasonable measure. Hopes of computationalizing a theory or model and performing actual experiments often force researchers toward toy problems; to scale their neural structures and problem to the smallest sizes and levels of complexity, to extrapolate from the component to the system; to abstract away critical system-level details; to perform sub-critical numbers of experiments; and to simulate rather than emulate, build, and deploy. Further, the isolationist nature of the vast majority of tools does little to encourage and support hierarchical integration and the essential and complementary process of synthesis (or reconstruction).

These research practice limitations, coupled with the inherent complexity of the brain, make it likely that we will continue to suffer from the lack of a comprehensive theory of the brain for a long time – unless we change them, and change them significantly.

BRAIN-SCALE MACHINES:

Perhaps a possible impetus for doing so is on the horizon. The human brain's raw processing power is, by some estimates, equivalent to approximately 100 teraflops (100 trillion floating-point calculations per second), at least in terms of neuron firings. This is a daunting number; yet Moore's law predicts that by 2020 a single computer chip will be capable of this brain-level performance. Surprisingly, though, parallel supercomputers are already nearing and will soon exceed these numbers. For example, Japan's NEC Earth Simulator delivers performance at 35.8 Tflops, about one-third that of the human brain. The joint DOE/IBM Blue Gene/L computer, targeted for delivery in 2006, promises a full PetaFlop of performance, or 10 times the estimated raw performance of the human brain.

In addition to exponentially increasing performance, Moore's law also promises drastic reductions in the size and cost of these brain-class machines. Already, the world's third fastest supercomputer (and the fastest university computer --- built by Virginia Tech, et al) is a cluster-based parallel computer costing a mere $5M US and yet exhibiting peak performance of 17 TFlops. The situation is going to get only better, and soon.

Within the very near future, brain-class machines will become ever more powerful, highly affordable, and broadly available, making possible for the first time large-scale, integrated simulations of mammalian brains. This is a massively significant, near-term, technological discontinuity that promises break-through opportunities for the cognitive and neuro- sciences. The attraction of this discontinuity is not simply in being able to access more computing power in order to model increasing levels of detail of small neural structures (such as the impressive efforts in modeling single Purkinje cells on a Cray supercomputer). It is more about being able to develop and deploy common, integrated computing frameworks and platforms that support a "brain as system" approach, with highly structured computational models, high-resolution sensors, and within complex, real cognitive environments, all as part of an effort to provide a sort of “computational neuroscience infrastructure” for the research communities.

Overview of the Proposed Grand Challenge:

This *Integrative Brain-Scale Computing* grand challenge centers around the exploitation of brain-scale supercomputing to enable the formal development of holistic frameworks and computing platforms for integrating advances in brain, mind, computing, sensing, and behaving into an integrative whole. The frameworks and platform must allow the techniques, tools, models, and advances in one field to be readily fit into an evolving, integrated computational framework that incorporates the advances of the others.

The proposed challenge would target high-fidelity emulation of complete artificial nervous systems. This will require the development of isomorphic, hierarchical architectures and principles of operation of the brain, at multiple levels, from brain structures, to brain processes, to cognition, behavior, and consciousness. Such integrated, hierarchical frameworks and platforms are common in other fields. One familiar example is the computer networking framework based on the seven layers of the OSI model. Another (perhaps overly simplistic) example might be the isomorphic relationships that run from software applications and data sets, through high-level programming languages, compilers, operating systems, architecture, machine organization, machine instructions, logic cells, gates, and on down to the electrons.

Unfortunately, current research paradigms and methods in the Bio-IT-Cognition fields are ill prepared to exploit this imminent, encroaching discontinuity. Success will easily require truly integrative, interdisciplinary involvement at “grand challenge” levels, rethinking the problems, paradigms, approaches, and goals in many cases, and institutionalized innovation across the involved research communities.

The goals of the brain-scale platform should include being able to execute brain-scale simulations over millions and then billions of neurons - and at increasing levels of biological fidelity. New methods of experimental and theoretical neuroscience research—coupled with new computational methods—will need to be developed, ones involving reusable, plug-and-play functional-model libraries, massive computational experimentation, large-scale evolutionary computing and evolutionary cognitive models, direct methods of developing and evaluating emergent properties, "in silico" neurophysiological experimentation, and even integration of silicon sensory and motor systems to create complete, end-to-end, artificial nervous systems. Many of these methods involve computer science and engineering technologies that are already widely applied to other fields; now is the time to integrate them into computational neuroscience.

A final, important goal of the proposed effort would be the proactive involvement of the computational neuro-science researchers with the hardware and system designers of industry, in order to avoid the historical mismatch between applications and architectures that has plagued many other fields. Because the brain is hugely non-homogeneous in its structure and operation, mapping models of it onto today’s supercomputers will prove highly problematic, as these machines are highly homogeneous. The brain’s I/O devices also clearly differ from classical computer I/O devices, as does the real-time data distribution into the brain. Just by considering the straight-forward analogy of the brain’s white matter to the interconnection network of a supercomputer, it becomes readily apparent that the nature of brain-scale supercomputers and processors in 2020 might desirably be significantly different from the RISC-processor and homogeneous interconnection and computing models of today. These architectural changes are achieved slowly, if at all; thus the efforts and communications should begin now.

Comments:

It should be pointed out that the goal of this grand challenge proposal is aimed at a “brain as systems” approach. It is less about enabling advancements in any one area of neuroscience than about mutually enabling advance-ments in all areas by forcing them to relate to, integrate with, and validate each other as part of a unified whole. The direct goal of the proposed effort would not be to discover and implement specific theories, no matter how compelling, but to provide the computational frameworks and platforms within which increasingly sophisticated theories can be explored, analyzed, and validated; other FP7 grand challenges may address specific theories.

It should also be noted that there are at least two significantly different paths this proposed grand challenge could take: one targets bio-inspired frameworks and models, and the other targets biologically faithful frameworks and models. I am specifically assuming the biologically faithful approach in this proposal. It is my feeling that this approach will better allow the integration of neuro-prosthestics as peripherals, the integration of large-scale neuronal probes, the use of artificial lesioning for exploring brain diseases and their effects on cognition, and perhaps the eventual accurate, human-level spike interfaces for brain augmentation studies, among other benefits.