The Architecture Of Brain And Mind

Integrating Low-Level Neuronal Brain Processes with High-Level Cognitive Behaviours

A Proposed Grand Challenge For Computer Science[1]

Mike Denham[2] (Moderator)

The last twenty years has seen an explosion in the application of molecular biology, genetics, and cell biological techniques to the problems of neurobiology, and a growth in neurobiological experimental research which has dramatically increased our understanding of the nervous system and its function. However the primary function of the nervous system is to gather, represent, interpret, use, store and transmit information, and neuroscience is inherently a computational discipline. Despite the insight neurobiology provides, a mature science of the brain thus ultimately requires a computationally based understanding of how information is represented, organised and used within the structures of the nervous system, and how such brain processes create the high-level cognitive capabilities which are manifest in the human mind.

In addition, in a world that day-by-day becomes increasingly dependent on technology to maintain its functional stability, there is a need for machines to incorporate correspondingly higher and higher levels of cognitive ability in their interactions with humans and the world. Understanding the principles of brain organisation and function which subserve human cognitive abilities, and expressing this in the form of a computational architecture of the brain and mind, will provide the foundations for a radical new generation of machines which act more and more like humans. Such machines would become potentially much simpler to interact with and to use, more powerful and less error-prone, making them more valuable life-companions, whether for learning, information finding, physical support or entertainment.

The Challenge.

To create a computational architecture of the brain and mind which is inspired both by the neuronal architecture of the brain and high level cognitive functioning in humans; captures the information processing principles present in the brain; describes how low level neuronal processes are linked and integrated with high level cognitive capabilities, such as adaptability, self awareness and creativity; provides a major input into the worldwide scientific endeavour to control or eliminate a range of human mental disorders; and will allow the creation of intelligent artefacts which incorporate a significant subset of human cognitive functional capabilities.

[1] **Acknowledgements**: the proposal draws together a number of submissions to the UKCRC Grand Challenges Workshop, Edinburgh, November 2002, that are explicitly acknowledged in the text. Thanks are also due to fellow members of Panel D at the workshop: Aaron Sloman, Leslie Smith; Simon Colton, Mark Steedman and John Sutherland for their substantial contributions to and considerable assistance during the production of this proposal.

[2] School of Computing and Centre for Neural and Adaptive Systems, University of Plymouth

Background.

This Grand Challenge, as summarised above, aims at drawing together a number of Grand Challenges submitted to the UKCRC Workshop on Grand Challenges for Computing Research, held at the National E-science Centre, Edinburgh, 24-26 November 2002. In particular it attempts to incorporate some key aspects of the challenges described in the following submissions:

Mechanisation of Thought Processes	David Al-Dabass
A Theory of the Brain	Roman Borisyuk
Towards a Testable Theory of Meaning	Yaw Busia
A Mathematical Theory of Creativity	Simon Colton
The Neocortical Microcircuit	Mike Denham
What is the Functional Specification of the Brain?	W.H.Edmondson
A Real-Time Computer Simulation Of The Human Brain	Steve Furber
Creating Machine Consciousness	Owen Holland
Self-Reflective Machine Learning	Dimitar Kazakov
e-Brain - A Large Scale Brain Modelling Experiment	Mark Lee
How can we build human scale complex systems?	Julian Miller and Catriona Kennedy
The Challenge for 21st century Computer Science is Biology	Ray Paton
Architecture for a mind: requirements and designs	Aaron Sloman
Building a synthetic sensory-motor system	Leslie Smith
Unifying High- and Low- Level Cognitive Systems	Mark Steedman
Computers as Part of Humanity	John Sutherland

Why is it a Grand Challenge?

Significance.

Is it driven by curiosity about the foundations, applications or limits of basic Science?

Understanding the brain is arguably the most fundamental and important scientific challenge facing mankind in the 21st century. A mature science of the brain ultimately requires a mathematically based understanding of the information representation, interpretation, use, storage, and communication that occurs within the structures of the nervous system, and the way in which the neural processes which constitute these structures support the cognitive capabilities observed in human and animal behaviour.

Likewise understanding aspects of human cognitive behaviour has been a topic of concern to diverse sciences, such as all branches of psychology, linguistics, anthropology, economics, education and AI. By linking information processing in the brain to the cognitive capabilities it engenders means that the proposed challenges addresses both kinds of scientific curiosity in a single, coherent manner.

The challenge is to understand and model brain function at different levels of abstraction, including
– physiological properties of brain mechanisms, e.g. cortical microcircuits;
– neural information coding, storage, processing and communication functions;
– higher level cognitive and affective functions of many sorts;
and in particular to understand how the levels are linked to form an integrated functioning system. Central to this objective is to develop an understanding of the fundamental principles involved and the mechanisms of self-organisation and adaption inherent in the brain.

The aim will be to abstract and formalise principles of operation rather than attempting to directly mimic the intricate chemical and physical mechanisms of biology. However in order that the development of the brain/mind architecture engages the interest of neurobiologists and neuropsychologists, a strong biological and psychological realism will be important from the start, and as the project progresses and successive models are created at all levels with more and more detail based on what is known about brain and mind, this should increase.

Is there a clear criterion for the success or failure of the project after fifteen years?

The challenge will be to use the developing understanding and models of brain function to design and build a succession of increasingly sophisticated architectures, demonstrable as working models of behaviour in either a physical or simulated/virtual environment, meeting carefully selected combinations of requirements. The intention would be that each step will be both achievable and, in itself, a major challenge, capable of pushing forward the frontiers of knowledge.

The criterion for success would be the achievement of each of the steps, over a period of fifteen years, leading at the end of that period to an architecture and working model which is able to demonstrate, in either a physical or virtual environment, or both, the capability of performing many cognitive tasks that reveal different competences, such as the tasks of communicating with others, asking for and giving help, and the purposeful manipulation of physical objects.

Does it promise a revolutionary shift in the accepted paradigm of thinking or practice?

Research into the brain has been very successful in the sense of providing knowledge and understanding of the workings of the nervous system at the neuronal level, but there has been far more limited success understanding how the brain operates as a holistic entity and in particular how it gives rise to a mind.

Several revolutionary shifts are likely to result:

- the need to understand how neural processes provide support for higher level cognitive behaviours, such as perceiving affordances and reasoning about them; self-monitoring, self evaluation, self control; empathetic and simulative reasoning; and creativity, can be expected to lead to both radically new computing paradigms, and radically new questions for neuroscience to pursue;

- the need to create working models of the brain/mind architecture can be expected to lead to radically new brain-inspired hardware architectures;

- the requirement to create a single integrated brain/mind architecture will require major shifts in the areas of AI and cognitive science, since at present little of this research has faced the issues of integrating symbolic and neural-level theories;

- the ability to demonstrate aspects of cognitive function and malfunction within a working brain/mind model may lead to revolutionary shifts in the scientific study of many aspects of human and animal mental functioning, including replacement of many purely verbal forms of explanatory theorising that accompany descriptions of psychological and neurobiological experiments;

- the task of formalising and mathematically analysing a hierarchy of architectures and working models provides a formidable challenge to mathematics and theoretical computer science and is likely to require new mathematical theories and analysis methods.

Does it avoid duplicating evolutionary development of commercial products?

It is highly unlikely that industry will proceed along this direction through any normal evolutionary commercial development process. Humanoid robots have been developed by a small number of companies, but these, whilst mechanically sophisticated in some cases, possess no computational architecture/ model which is capable of providing cognitive capabilities. Indeed, without diminishing the difficulty of the task or extent of the achievement, they are currently built solely as physical prosthetic models of humanoid movement. In future it is likely that they will provide a computational architecture which is limited to supporting programmability for various particular tasks, but not the kind of multi-functional, biologically grounded, brain/mind architecture proposed here.

Impact.

Will its promotion as a Grand Challenge contribute to the progress of Science?

The primary function of the nervous system is to gather, represent, interpret, use, store and transmit information; thus, neuroscience is inherently a computational discipline. A mature science of the brain and mind ultimately requires a theoretical, computational based understanding of the information coding, storage and processing that occurs within the structures of the nervous system. It also requires a new deep understanding of features of the environment on which the success of biological information processing systems depends, for instance an analysis of the types of affordances available in various environments and the ways in which they can be detected, represented and used. Through this Grand Challenge computer scientists have the opportunity to make a primary and key contribution in this international, multidisciplinary scientific endeavour.

In addition the complexity of the brain is at least as complex as any of the most complex artifacts which exist in the world today. Insights into, and paradigms developed in this Grand Challenge, for understanding the architecture of the brain and how it leads to high-level cognitive behaviours, will provide tools for coping with the complexity of understanding the behaviour of the most complex artefacts.

Does it have the enthusiastic support of the general scientific community?

Achieving this may require that the work proceeds, as proposed, by means of a carefully selected sequence of sub-challenges, meeting carefully selected combinations of requirements which are perceived as achievable, each sub-challenge building on the achievements of the previous one.

Does it appeal to the imagination of other scientists and the general public?

Understanding the brain, and in particular the basis of such human capabilities as self-awareness, consciousness, creativity, has been a topic of fascination for both scientists and the general public for many years. Melvyn Bragg has remarked that his biggest post from listeners comes when his radio series ("Start The Week", "In Our Time") deal with issues in brain science.

The idea of building something which possesses the cognitive capabilities of a person has fascinated people for centuries, e.g. the old "golem" idea, stories and films about man-made monsters such as Frankenstein; likeable robots in science fiction films such as Star Wars, Forbidden Planet, AI; fearsome intelligent machines, such as COLOSSUS: The Forbin Project; and most recently AI-based synthetic agents in computer games. The growth of functionality in increasingly visually realistic and rich video games has created virtual systems of immense complexity. These video games particularly capture the imagination of the young. The richness and 'fun' of the environments in which the working models of the brain/mind architecture will be tested should have strong public appeal.

What kind of benefits to science, industry, or society may be expected from the project, even if it is only partially successful?

Benefits to science will include:

- a new understanding of human capabilities, through a detailed specification of the requirements of the new brain/mind architecture. We understand very little about what the requirements for a human-like system are. Although we talk confidently about humans as seeing, thinking, learning, communicating, acting, being creative, having desires, intentions, feelings, emotions, and being conscious, we have no clear idea of what we mean, and what is required for us to attribute these qualities to a machine. A major intermediate benefit of the challenge will be a detailed analysis of requirements for satisfying these everyday descriptions of humans;

- new insights into how neuronal processes can support simple high-level cognitive functions, such as reactive decision making; more complex functions such as creativity; and affective states and processes, including desires, preferences, emotions, etc., providing the basis for novel intelligent systems and human-computer interfaces;

- novel paradigms and processes for learning and memory based on principles derived from how the brain carries out these functions;

- novel tools for modelling neurobiological and cognitive processes;

Benefits for industry will include:

- new insights into the principles underlying information processing in neuronal circuits and systems, leading to novel artificial sensory systems (vision, audition), and speech production and recognition systems which approach human levels of performance and display similar properties of graceful degradation;

- novel simulated virtual environments for the demonstration and emulation of intelligent behaviour, applicable to the computer entertainment industry;

- the potential for putting cognitive capabilities into industry developed humanoid robots, leading to progressively more sophisticated "robot assistants" ;

- feedback to the world's largest entertainment industry - video games - on new ways of using their technologies to work with customers, and possibly increase market penetration.

Benefits for society will include:

- a major contribution to the international endeavour to control or eliminate a range of human mental disorders.

- the increased possibility of "domestic robots" with sophisticated capabilities, to support, for example, mobility and care of the aged and the physically challenged;

- novel educational methods and tools based on an increased understanding of learning and cognitive capabilities;

- novel systems for intelligent access to information sources;

- a better understanding of what, if anything, makes us uniquely human and differentiates us from machines, eg machine versus human consciousness.

Scale.

Does it have international scope?

The challenge can easily accommodate a world-wide collaborative effort, drawing on expertise from many disciples and many scientific and industrial centres in many countries. Indeed international collaboration is already mandatory in this field.

There are a number of international projects which aim at providing specific challenges to the cognitive capabilities of machines, e.g. the international RoboCup Rescue project: http://www.r.cs.kobe-u.ac.jp/robocup-rescue/ and DARPA's new Cognitive Systems project:
http://www.darpa.mil/body/NewsItems/pdf/iptorelease.pdf

How does the project split into sub-tasks or sub-phases, with identifiable goals and criteria, say at five-year intervals?

The overall challenge can be identified with a sequence of sub-challenges, each building on the goals of the previous sub-challenge, each with achievable aims and scope. Each of the sub-challenges will itself require a collection of sub-goals to be achieved by researchers working on different aspects of the architecture, working models and demonstration physical and virtual environments, from within each contributing discipline.

What calls does it make for collaboration of research teams with diverse skills?

Understanding the brain, creating an architecture, and building a working model and a virtual environment in which its cognitive capabilities can be demonstrated, requires the collaboration of scientists and engineers from a wide range of disciplines which will substantially exceed that which currently exists.

How can it be promoted by competition between teams with diverse approaches?

The establishment of competitions for robots, such as the international RoboCup and RoboCupRescue competitions, commensurate in their cognitive demands with the goals of the identified sub-challenges will allow teams to test their approaches in a range of intellectually stimulating real and virtual environments.

Timeliness.

When was it first proposed as a challenge? Why has it been so difficult so far?

This challenge has been in the minds of scientists for as long as people have considered what is the nature of mind and brain (the word "brain" appears on the ancient papyrus called the Edwin Smith Surgical Papyrus. This document was written around the year 1700 BC, but is based on texts that go back to about 3000 BC. This document is considered to be the first medical document in the history of mankind).

Up until very recently, our knowledge of the mechanisms of the brain has been very sparse and limited in depth. It was only fifty years ago, in 1952, that A.L. Hodgkin and A.F. Huxley first described the voltage clamp method for measuring neuronal response which has formed the basis for much of the neurobiological experimental investigations since that time. In the last ten years methods of imaging of living human brains (PET, fMRI) have provided a wealth of new knowledge about the relationships between brain activity and cognitive function. One of the most recent techniques to be developed is an MRI-detectable, neuronal tract-tracing method in living animals, which recently demonstrated MRI visualization of transport across at least one synapse in the primate brain. Transsynaptic tract tracing in living primates will allow chronic studies of development and plasticity and provide new valuable information about brain anatomy.

The challenge of devising symbolic and algorithmic representations of cognitive abilities such as reasoning and understanding has an almost equally ancient origin in the Logic and mathematics of the Ancient Greek and Arabic cultures, and was recognized in its modern form in Boole's "Investigation of the Laws of Thought" and Ida Lovelace's account of Babbage's Analytical Engine, and such direct descendants as Frege and Turing. A major advance was achieved when Chomsky, Minsky, and others succeeded in applying formal and computational methods to the analysis of particular human cognitive abilities. More recently the exponential growth in performance and decrease in cost of computing machinery has made it possible to apply these methods much more extensively.

One of the most significant reasons why this challenge has proved so difficult is that until recently the two research activities outlined above, the neurally- based and the symbolically mediated, have largely proceeded separately. It is now clear that many of the obstacles that both approaches have encountered arise from this separation: symbolic representations are often intractable and fail to scale because they fail to make any contact with natural categories that might arise from the physical interaction of beings with the world. Neurally embedded mechanisms, on the other hand, and the related machine learning techniques, typically fail to capture the higher levels of representation and situation-independence that the analysis of symbolic cognition indicates must be involved.

Why is it now expected to be feasible in a ten to fifteen year timescale?

Advances in neuroscience, including neurobiology, computational and theoretical neuroscience, and neuropsychology have accelerated in the last ten years at an unprecedented rate, using increasingly sophisticated technology, eg two-photon laser scanning microscopy, fMRI, direct imaging of neuronal activity and connectivity *in vivo*.

Moreover computer speeds and memories are now of a size and power that most researchers could not hope for even ten years ago, and this trend is likely to continue. This makes possible exploratory research of the type considered here at far greater speeds than ever before.

Developments in materials science and in miniaturisation make feasible the design of robots with sensors, motors, and animal-like limbs with size/weight/strength/power/cost ratios that offer new opportunities for building working models of the brain/mind architecture in which to test the ideas. Also, it is only in the past two years that virtual reality development and run-time environments have reached the stage where complex, virtual worlds can be built for testing and widely disseminating the working models of the brain/mind architecture.

This rate of increase in knowledge and technology is expected at the very least to be sustained over the next fifteen years of the Grand Challenge. This will lead to previously inconceivable experimental procedures and insights into brain processes. Computational tools will correspondingly increase in sophistication and power, coupled to the development of new theories to elucidate the principles of information processing in the brain.

However, whilst this Grand Challenge is guided by the long term scientific goal of understanding how the human brain functions in supporting the full range of human mental processes, it is not claimed here that this goal can be achieved in fifteen years: on the contrary a far longer time will be required. Nevertheless, within ten to fifteen years major progress is possible that will provide a solid foundation for further research in the decades that follow.

What are the first steps?

The first steps will involve creating a set of functional requirements for a computational architecture of the brain and mind. From this it will be necessary to firstly identify important core subsets of the requirements which could be put together in complete working architectures/models able to demonstrate their competence, and secondly, to specify an initial sequence of such architectures/models to aim at, that are expected to be achievable (albeit with great difficulty), demonstrable, in a succession of working (physical or simulated) architectures, and provide a launch pad for achieving the next architecture/model in the sequence. It will also be necessary to define, as part of the requirements analysis, a specification of a set of metrics, eg precisely defined competences, which will allow a principled approach to determining progress.

What are the most likely reasons for failure?

The most likely reason is a reluctance of the funding agencies to maintain adequate levels of support over such a long period of time.

Initially an obstacle will be the scarcity of researchers with sufficient breadth in their knowledge and experience to be able to contribute to such a project. One of the most important initial tasks and immediate benefits of embarking on this challenge will be the rapid creation of a new generation of researchers equipped to carry the work to completion.

Other difficulties which could arise are a possible reluctance on the part of neuroscientists and other scientific communities to get sufficiently engaged with computer scientists in the challenges posed. Waning public interest may also pose a problem, and there will need to be some early significant breakthroughs, leading to featured articles in popular science magazines and presentations in television programmes such as Horizon and Equinox, in order to maintain strong public interest in the Grand Challenge.