Roadmap of Neuro-IT Development

Edited by

Alois Knoll & Marc de Kamps

With contributions from

Igor Aleksander, Tobi Delbrück, John Hallam, Giacomo Indiveri, Marc de Kamps, Alois Knoll, Anders Lansner, Riccardo Manzotti, Rolf Müller, Guy Orban, Francisco Pelayo, Herbert Peremans, Eduardo Ros, Ricardo Sanz, Guilio Sandini, Erik de Schutter, John Taylor, Vincenzo Tagliasco, Richard Walker

Project funded by the Future and Emerging Technologies Arm of the IST Programme under Contract No. IST-2001-35498

Introduction

Overall aims of Neuro-IT

The aim of Neuro-IT.net is to create a new area of research at the interface between Neurosciences (NS) and Information Technology (IT) within the European Union. The term Neuro-IT has been used to express clearly the idea that the disciplines have been merged under the umbrella of nEUro-IT.net and have formed a new scientific working area, different to what was traditionally called Neuro-Informatics (NI).

The objective has been to move beyond well-established NI and AI (Artificial Intelligence) by fostering research to benefit both the NS and IT communities by helping to solve problems linked to the emergence and modelling of cognitive and awareness processes.

This is a groundbreaking move and the very first time that the fusion of these two scientific diciplines has been attempted. It is also a significant move for neuro reserach overall, clearly establishing Europe at the forefront of development, ahead of the US, and keeping Europe globally competitive in this field, with products and results fully expected in both the short and long terms.

The goal is for IT to profit from NS results to improve IT artefacts and for NS to validate models or hypotheses with a better use of IT. nEUro-IT.net is therefore particularly committed to:

- Publicising the potential of basic research conducted within the EU-funded Neuro-IT initiatives ALG (Artefacts that Live and Grow), LPS (Life-like Perception Systems), and NoS (Neuron on Silicon) to related scientific communities, to SMEs and global companies.

- Spearheading the emergence of new, visionary, long-term research objectives that could fully exploit the potential of collaboration between Neurosciences and Information Technology.

Why are we doing this?

Where our knowledge of the brain and brain function is expanding rapidly, our ability to make use of this information has somehow not increased at the same rate. Living creatures still outperform computers in a large range of skills, many of which are considered to be “simple.” Computer scientists can only dream of the object-recognition skills of humans, and roboticists would love to create service robots with the same degree of autonomy as an ant. The possibility for the development of artifacts that are able to learn over their lifetime and are able to adapt their behavior in the face of changing circumstances seems even more remote. In general, every complex artifact has to be programmed carefully, by hand, and for a new range of applications this has to be done again, from scratch.

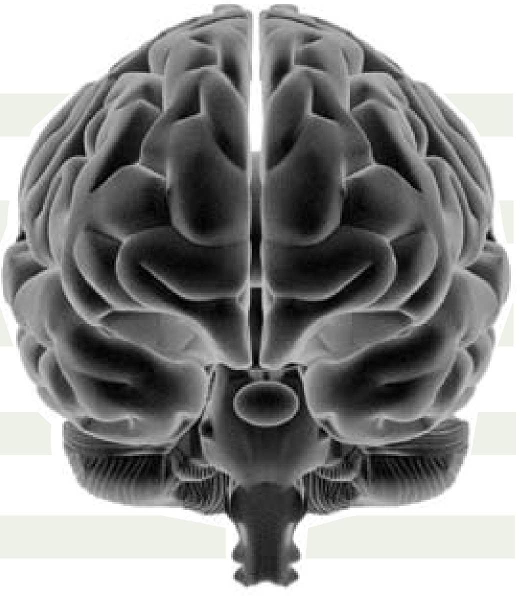
The relatively slow progress in the creation of bio-inspired artifacts and IT applications is a source of frustration for policy makers, scientists, and engineers alike. Scientists and engineers who try to emulate methods used by nature find that their bio-inspired approaches work very well on some problems, while failing on other, seemingly related ones. Or they find that approaches that are promising on toy problems do not scale well with the problem size. Behind these problems is a lack of systematic understanding of how nature accomplishes things, which, as we will see, is one of the central issues in the roadmap. This makes bio-inspired engineering difficult and solutions to many problems can only be found by trial-and-error. The haphazard development of bioinspired engineering is also undesirable from a political point of view: clearly there is great economic potential in some fields of research and in order to stay competitive it is important to know which research to fund. Moreover, the impact of new technology on society can be considerable, as we have seen recently with the Internet.

The Grand Challenges

With this in mind, nEUro-IT.net is not simply supporting incremental research, no matter how excellent it is, but is helping to discover new research domains that could lead to breakthrough in Neuro-IT in the long term. Neuro-IT really does have potential.

To sum up our objectives, we're looking for completely new research areas and breakthrough discoveries. Essentially: "What can neuroscience do for IT?".

In terms of this process to date, we have created a roadmap of our progress in the form of eight projects or '**Grand Challenges**'. These are:



The Brainship Project
The Bio-inspired hardware project
The Factor-10 Project
The Acting in the Physical World Project
The Conscious Machines Project
The Artificial Evolutionary Design Project
The Constructed Brain Project
The Tools for Neuroscience Project
The Brain Probe Project

This document is an abridged version of a more extensive report, which is available to view in full at the Neuro-IT.Net website, where you will also find other related scientific material.

www.neuro-IT.net

Summaries

The Brainship Project

Recent progress in research suggests a popular subject of science fiction may soon be technically possible: direct interfacing of the human brain with computers. Better electrodes and fast-signal processing techniques have spurred incredible breakthroughs in research using rodents and monkeys. But all current demonstrations are one-directional using signals from motor cortical areas to control virtual devices. For real-life applications, like the control of paralysed limbs, bi-directional interfacing (BBCI) will be necessary. The primary goal of this project is to develop an awake animal model where the brain interacts with the environment through BBCI techniques, and guidelines for ethical use of BBCI in humans.

The Bio-inspired Hardware Project

Technological progress has rapidly increased the number of transistors on a single chip. But most computing systems are limited by their dependency on clocked digital processing units with little parallelism. Silicon technology can create bio-inspired processing, outperforming conventional solutions and providing insights into the working of the human brain. Brain-like computing on silicon will be useful in a broad range of applications from real time control of robots, implantation of artificial cochleas and retinas, to large scale simulation (eg. of the brain). Existing technology may not be powerful enough for this application. A new generation of systems are needed which will emulate biological evolution. The potential impact on industry and human society is huge.

The Factor 10 Project

The emerging fields of epigenetic robots (embodied systems which develop through interaction with their physical and social environments) and smart materials offer a wealth of research opportunities. We envision the emergence of an artefact which evolves its cognition and motor control autonomously, based on multimodal/multisensory feedback in a fixed body – a humanoid. We will also evolve an artefact which develops new skills by coupling with its environment and using bodies/effectors that adapt their shapes to different tasks. Finally in this challenege, we will develop an artefact which co-evolves its brain and body in permanent interaction with the environment over an extended period of its lifetime. In summary, we will work towards a fully functional physical artefact which grows its body and IQ autonomously, by a factor of 10, over a period of, say, 10 months.

Acting in the physical world project

The objective of the successful ‘thinking-and-acting-in-the-physical-world’ challenge is to build complete systems making full use of sensors, actuators, body morphology and materials. Intelligence devices would be embedded in their peripheries which would enable artefacts to master tasks known to be performed by natural systems. These so-called smart peripherals will be able to be used by many different projects. Work should establish general rules which can be applied to problems for which no solution is known to exist in nature.

The Conscious Machines Project

There is considerable interest in Europe in the topic of ‘Machine Consciousness’. This topic, which even 10 years ago would have been dismissed as ‘crackpot’ science, is now taken seriously. Thanks to progress in brain science we now understand that the human brain sifts through a huge amount of information before selecting the elements that influence behaviour. Current IT applications must be programmed carefully and laboriously and struggle with tasks considered easy for human beings. The challenge is to implant a level of consciousness in these systems. This could benefit technical systems, such as power grids, which need to co-operate with other applications in order to monitor their own performance.

The Artificial Evolutionary Design Project

Traditional Artificial Intelligence (AI) has failed to scale up to real-world applications and similar problems face so-called new AI. This project aims to develop design techniques inspired by evolution. This project aims to develop new mathematical models of Complex Adaptive Systems (CAS), including techniques for the design of highly evolvable structures and behaviours, adaptable to different environments. This could lead to ‘generic’ autonomic (self-managing) robots whose body plans, sensor configuration and processing capabilities can be trained to do specific tasks. Hybrid chemical-computerised environments could be created for the evolution of complex ‘wetware’. This is close to artificially creating life.

The Constructed Brain Project

In the ‘constructed brain’ it is argued that for a systematic development of cognitive engineering principles in NeuroIT, what is needed is the simulation of an entire brain. Initially realised as software, it would later be linked to hardware, moving from a purely ‘virtual’ to an ‘embodied brain’. We need a good understanding of neural processes in order to find out why biological creatures are so good at ‘cognitive processing’. This will also tell us whether we can use biological information processing principles with existing hardware and find out how to transfer these concepts to artefacts. New ways to study the brain could emerge allowing ‘experiments’ which would bring leaps forward in treating psychological disorders.

The Tools for Neuroscience Project

Brain research is important for information systems and technology (IST) but the medical justifications are overwhelming. 35% of disease in Western Europe is due to brain disease. This project sets ambitious targets to record from a thousand electrodes in five different brain regions to obtain high-resolution images to better understand the integration between the molecules, and genes encoding them, and behaviour. The use of non-human primates is seen as critical for these developments of brain research. Non-invasive techniques still lack resolution. We suggest better public education is needed on this issue due to the ethical questions raised. Unless these are better understood, brain and other pharmaceutical research will continue to leave Europe.

The Brainship Project

The cyborg – a popular subject of science fiction movies, now seems possible thanks to recent advances in neurophysiological research which allow direct interfacing of the human brain with computers and machines. In 2004 the first implant of an electrode array in the brain of a quadriplegic patient allowed control of external devices including a robot arm – a new field called neuroprosthetics.

Advances have been made because of better electrodes and faster signal processing techniques which were tested by implanting recording electrodes in rodents and monkeys. Present hardware is too unwieldly and patients cannot yet have independent use of their neuroprosthetic implants. The first patient asked for his implant to be removed and, even then, its functional lifetime would have been too short for general use in humans.

This major breakthrough was made possible by the surprisingly high plasticity (ability to change to better cope with the environment) of neural coding in the mammalian cortex. Instead of researchers adapting the equipment to the way the brain works, the brain adapts itself to the task of controlling the equipment, even over a limited number of channels.

This led to an explosion of application of brain machine interfacing (BMI) from controlling cursors on computer screens to upper limbs in monkeys. Brain plasticity allows subconscious skill-learning and once learned, the patient does not have to attend to the task. More advances need to be made to move towards cyborg-like applications and move beyond motor control tasks towards cognitive tasks like memory functions. Much of the visionary research proposed here can be done using rodents.

The primary goal of the Brain Interface Project will be the development of awake animal models where the brain interacts with the environment only through BMI techniques based on implantation in several regions of the central and peripheral nervous system. More controversial applications lie in the direct control of remote robotic devices and enhancing humans with embedded machines as in the cyborg.

We aim to use animal models with brain implants or recording and/or stimulating electrodes to implement BMI in other contexts than limb control – such as artificial sensory perception, control of legged locomotory devices and control of navigation. In the final phase, fully bionic animals with bidirectional brain interfaces for control of high-dimensional systems will be created.

Examples of human application

There will, of course be many applications for this technology, not least in terms of the mental control of remote exploratory vehicles at both ends of the size scale – from microendoscopes to be used in surgery, through to vehicles to be used for deep-sea exploration.

Two other potential human applications will be the ability for mental interaction with information systems using direct perceptual input and the chance to help the severely disabled with neuroprosthetics – artificial limbs that can be controlled using the brain.

Finally, we will be looking to create bionic animals, which operate by brain computer interfacing with their environment.

Current state of technology

Recent studies have investigated the possibility of predicting limb movement from the activity of multiple single-neurons (nerve cells) in the motor cortex. This was applied to rats and then monkeys with the aim of controlling an external device such as a cursor on a computer screen, in real time using signals recorded from the brain. Such techniques could be the basis of neuronal motor prostheses, replacing the function of impaired nervous systems or sensory organs with artificial devices.

An important finding in both rat and monkey studies was that the animals continued to learn under closed loop conditions, implying an adaptation of the intrinsic neural coding signals recorded by the implanted electrodes to the properties of the external device, a process called brain plasticity. Recently this experience was used to put the first implants in humans but current electrode designs, with average life times of months to about a year, are not suitable for chronic implantation in humans. The state of technology is still very limited. Arm control, for example, is rudimentary because it employs visual feedback only. The absence of sensory feedback is a major problem.

Additionally BMI could be used in many other contexts like, for example, navigation or communication between individuals, but this has not been tested on animals yet. Finally, research needs to be done in order to understand the mechanisms, extent and limitations of brain plasticity in order to make even more breakthroughs in BMI. These areas of research include:

- Understanding the extent and limits of brain plasticity. How many overlapping codes can a brain area learn and when will plasticity interfere with original function?
- Developing strategies to optimally stimulate sensory regions by central or peripheral implants.
- Identifying the learning and coding strategy used during brain plasticity to develop better training methods.
- Understanding how the brain integrates multiple functions by using simultaneous recording from many different brain areas. This is necessary to develop BMI applications where some tasks are delegated to the robot side and to integrate multiple BMIs in more advanced bionic applications.
- Studying the usefulness of BMI as a new experimental model which enables closed-loop studies of the nervous system.

The IT implications

In order to implement this technology, we will be developing several IT solutions. Perhaps the most important of these will be to create real-time encoding/decoding software for brain input/output signals, algorithms robust to noise, changes in signal quality and brain plasticity.

As well is these, it will be important to develop training methods to maximize signal transfer over a limited number of channels of different qualities, as well as methods for shared control versus partial autonomy in real-time brain robot interaction. Finally, we will be devising effective strategies for the perception/decision/action chain in robotics which will be necessary for partial autonomous action.

Materials technology

As well as the IT solutions, we will, of course, need to develop our material technology to make these functions possible. We will be developing stimulation electrode arrays to allow direct input to the brain of spatiotemporal (space and time information) sensory data, as well as working on the longevity and durability of electrodes, which need to be suitable for long-term implantation in humans. As well as this work on electrodes, we will also be conducting research on alternatives to implanted electrodes, e.g. cortical EEG.

Meanwhile, it will be important to continue to miniaturise all the electrophysiological equipment (filters, amplifiers, spike detectors), combine it with the control software and put it into wireless, battery-operated configurations. The sensors and actuators must also have a performance as good or better than natural ones.

Ethical considerations

There are, of course, many ethical considerations to be made when working with the human brain. The invasive technology may cause brain damage, for example, while brain plasticity may interfere with the normal operation of the human brain. As well as these health considerations, we should also consider the moral one and ask the question, should we try to enhance humans with embedded BMI applications?

This challenge will develop new technology which can have great impact on human society, both at the personal and sociological level. Current technology allows only for highly invasive interface devices and therefore their use should be restricted to situations where they are deemed acceptable or necessary.

The Bio-inspired hardware Project

Brain-like computing on silicon

Technological progress has led to the rapid increase in the number of transistors that can be included on a single chip. But most computing is based on a single digital processing unit with a low degree of parallelism (simultaneous execution of multiple tasks). Waits for remote memory access increasingly dominate computation and large numbers of idle transistors consume a large proportion of the power. Designers try to enhance computing power by implementing multiple cores on a single chip but scaling of this approach will be very difficult. One alternative is to implement bio-inspired processing schemes providing the kind of dynamic adaptation found in the brain. Such schemes could be based on current silicon technology and could be the first step towards the development of novel high performance computing strategies, as well as offering new insights into the working of the brain.

Objectives

Bio-inspired processing systems can provide more efficient processing. Novel computational primitives can be embedded in robots and other artificial devices and be used to build Brain Computer Interfaces.

Examples

Implementation of vision systems on silicon. There have already been several single-chip implementations of motion processing but we are just starting to be able to build complex visual systems capable of emulating the early and middle stages of visual processing. In many cases the main problem is information transfer. A promising strategy is to separate sensing from computation and distribute the processing across several chips that communicate using pulse-frequency modulation to encode information, like spikes do in biological neural systems.

Embedded processing in robots. Bio-inspired, hardware implementations of circuits for vision, hearing, motor control, etc are the ideal embedded systems for use on robots which interact with the real world in real time.

Real-time sensory-motor integration. In simulated environments, the time dimension is very flexible but when working with real robots in dynamic environments time constraints are strong. Bio-inspired computing strategies that use time as a resource, could present interesting opportunities for designers.

Brain interfaces. Current technology allows real-time signal acquisition, processing and feedback but soon it will be possible to simulate brain areas and interface them to areas in real brains. This could help, not only to understand basic functions of the brain but also to build simulations of the brain and brain prostheses to support sensory or motor functions.

Simulation engines. Hardware implementations of bio-inspired processing architectures can provide a high performance computing platform on which to perform simulations of brain areas, at different resolutions. This kind of work prepares the way for silicon implementations of brain-like computing primitives.

Current state of technology

Progress in analog VLSI (Very-large-scale integration) and in digital, programmable circuitry has resulted in real-time visual processing devices for artificial perceptive systems, robots, implants and massively parallel processing architectures. Researchers are also working on ways to exploit the parallelism provided by FPGA (field-programmable gate arrays) and analog devices. Incorporating systems inspired by biological nervous systems, which can adapt to the peculiarities of specific tasks, might make it possible to devise scalable computing schemes using FPGA technology.

We call this “Building brain-like systems” and the real key here is adapting the technology to make it possible. The most promising solution in the long term is to develop dedicated VLSI solutions using both analog and digital circuits to exploit the physics of silicon to emulate biophysics. A parallel solution is to use the computational resources of modern digital FPGAs. However these types of devices are mainly designed for signal processing, image processing, etc. and may not provide the kind of primitives required by Neuro-IT.

In order to overcome the deficiencies in current design methodologies and associated technological constraints it may be necessary to develop strategies for simulating neural structures on FPGAs. This will lead us to towards brain-like processing.

Problem areas

There are many problem areas that currently exist in which hardware implementations of bio-inspired hardware could be useful. These include:

- Bio-inspired processing schemes requiring real parallel processing.
- Asynchronous systems (where each operation is started only after the preceding operation is completed).
- Systems using spike-based processing and communication.
- Real-time processing systems for vision, hearing, olfaction, complex robot control.
- Brain-machine interfaces.
- Brain-like processing primitives in silicon.

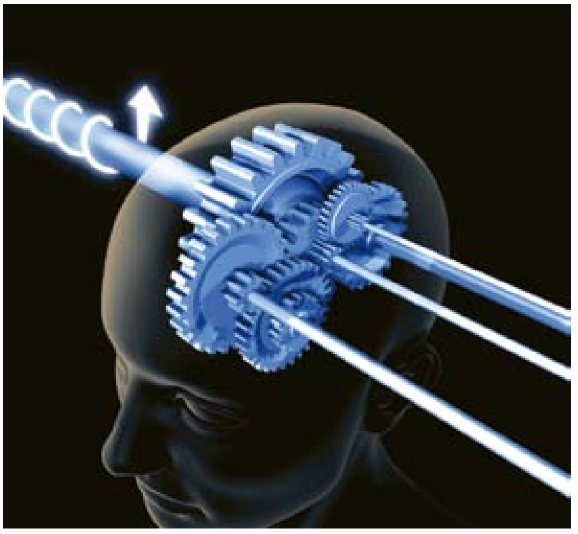
Where should the field go?

Circuit designers working in Neuro-IT should be more exposed to neurobiology and neuroscience, to learn about bio-inspired processing schemes. In the longer run, it will be necessary to adapt these schemes to simulate specific biological systems. There are many important Neuro-IT applications.

These include bio-inspired solutions to real world problems – artificial cochleas, artificial retinas, neuromorphic vision chips, neuro-cortical implants, and image processing.

There are also neuromorphic approaches. These are simply opportunistic design strategies which take new ideas from biological processing and combine them with techniques (such as time multiplexing) that take advantage of the implementation technology, while working round its constraints.

There are also several fields (such as spiking neural networks) that could benefit from ad hoc circuitry allowing efficient simulation of large scale systems and, as such, this will lead to the development of simulation platforms.



Designers will also be working on real-time processing systems, creating machines that close the perception-action loop, interacting with the real world.

These systems would need to evolve and become self-adapting in real world environments. Such systems would exploit information about task performance and resource use to generate more effective machines. This would need physical agents that evolve in physical environments.

In terms of robotics, on-board processing resources will make it possible to build autonomous machines with perceptive capabilities. Such machines could be used in studies of swarm intelligence or robot societies.

Immediate goals

In terms of the immediate work needed to be done to ensure progress towards those ambitions goals, we need to implement bio-inspired parallel processing schemes eg. smart vision systems, while pushing the technology to include primitives and design methodologies suitable for brain-like computing.

Hardware implementations for real-time processing, will provide robots with powerful sensorymotor capabilities and implementing neural-like processing schemes will also enable us to progress with our robot ambitions.

Ethical considerations

One of the arguments for bio-inspired processing schemes is that they allow us to 'understand by building'. When engineers reverse-engineer biological systems they meet the same problems nature resolved during evolution. We expect bio-inspired processing to outperform conventional computing, particularly in the areas where biological systems are most impressive: vision, reasoning, coordination of movement in complex bodies, model building for accurate movement control, etc. It might become possible to create new aids for handicapped people or to provide sensory augmentation for use in dangerous situations. This could have a huge impact on human society as a whole.

The Factor 10 Project

The emerging fields of epigenetic robotics and “smart” materials science offer innovative research opportunities and a large spectrum of new products. However, a completely new discipline may develop by combining key research in these fields to work towards three types of artefacts.

Type I
Artefacts evolving their cognition and motor control autonomously based on multimodal/multisensory feedback in a predefined and fixed body e.g. the “dancing robot” or the “classical” humanoid robots.

Type II
Artefacts that evolve new skills in structural coupling with the environment but with bodies/effectors that flexibly adapt their shapes to structurally different tasks, e.g. truly dextrous “hands”.

Type III
Artefacts that co-evolve their brains and their body in permanent interaction with the environment over an extended period of their lifetime (embodied artificial ontogenesis).

The third type may be seen as a new interpretation of smart materials with tailor-made functionalities for building up macro-structures with integrated sensing and cognitive abilities.

While artefacts of first and second type can be seen as classical allopoeitic machines, i.e. machines that are designed and built “from the outside in”, we hold that the third type of artefact needs a fresh approach as it can only be realised as a machine built from cells.

Following these lines of thought, we are defining a long-term research project called “Factor-10” or Factor-X, which aims at a physical artefact which autonomously grows its body and IQ by a factor of 10 over a period of 10 months.

This vision is largely inspired by the development of living organisms in embodied creatures through their permanent interaction with a real environment. One may argue that eventually the implementation of such artefacts would be based on biological substrates because nature has solved exactly the same problems of survival on earth through creating living organisms. We hold that it might not be desirable but necessary to begin Factor-10 related research by studying the challenges and promises of the concept of artificial growth using “dead matter” as a starting point. Meanwhile, mind development and bodily adaptation would have to be treated as two separate problems with the aim of overcoming this artificial separation as soon as possible.

Motivation and objective
For at least the last five decades the general public has been promised the advent of robots or human-like artefacts that would be of real help to us in our daily lives. However, as expectations rose, science consistently failed to deliver robots that could be compared to biological creatures.

Enormous progress has been made towards designing truly autonomous artefacts of types II and III outlined at the start of this chapter, such as brain and cognitive science, information technology and artificial intelligence, molecular biology and chemistry. The time is ripe to integrate them into new systems with autonomy and control intelligence distributed over their entire body, which in turn may adapt smoothly to a specific task. This could provide an economical market that cannot be underestimated.

While about 10 years ago the market for service robot systems (for both home and factory use) was projected to be larger than €1billion by the year 2000, less than a thousand have actually been deployed. The world market for standard, fixed production robots is about 100,000 units per year; it could also grow drastically if the perception and task-adaptation abilities of these robots increase and their programming efforts can be reduced.

Despite the current wave of euphoria for humanoid robots as impressive feats of engineering they hardly lend themselves to practical use outside of robot labs. The development of Type II artefacts could offer practical solutions to be appreciated by a broader public.

Looking at the preconditions for embarking on this research journey, we note that there is already a sizeable body of research in the diverse, necessary disciplines represented in Europe however with fragmentation across disciplines and countries. Apart from the scientific objective of developing the technologies, designing and building prototypes of type III artefacts – via type II as an intermediate goal – it is also the purpose of the project to establish a commonly accepted model for designing these artefacts.

Initially, recent resultshave been collected and translated into a language common to all the disciplines. More important, however, is the development of new theories, methods and paradigms by studying how methods from one field can guide research in another.

The goals of Factor-10 are demanding. Up to now, they have hardly been formulated as a common integrating challenge. We believe, however, that in view of the progress achieved in many disciplines, such as cognitive and neurosciences, Factor-10 comes at the right point in time. **If Europe does not take the lead now, it might miss yet another technology train.**

State of the art and projection from today's viewpoint
A small body of published work suggests some starting points for further research. We should examine the possibilities of building modular robots from a certain number of identical motor modules, which can be combined into different shapes and macro structures.

Another area of research that's needed is into the field of evolutionary and epigenetics robotics, both in the sense that robot shapes are optimised according to certain target functions and that the principles of autonomous learning is based on very basic instincts.

We also need to look carefully into microscale structures that can be assembled according to external conditions and that can serve as filters, modulators, etc. for chemical reactions. Alongside this, nanoscale self-assembling structures are also vital. These can grow to macroscopic size, “muscle tissue”, for example, and exhibit useful properties, such as joints without lubricants. This would rule out the use of processes relying on extremely high voltages or extremely high external pressures for the structures to form themselves.

There are a number of research areas that may directly contribute to this through elucidating principles of biological development in view of what is needed for type III artefacts:

1. **Developmental biology:** Compilation of the essential principles that enable living organisms to differentiate cells to form large bodies with specific organs.

2. **Genetics:** Contribute a set of rules that encode a set of “genes” which allow stable bodily development but control communication between the individual body cells so they can interact with the environment of the artefact

3. **Computational Neuroscience:** To develop basic processors (neurons) along with their interaction principles and communication networks/mechanisms that enable the parallel emergence of motor skills and cognitive skills.

It may be argued that there are good reasons to discuss carefully and review the size and functionality of the ideal basic block. Should it be the atom, the molecule, or perhaps assemblies of micromodules at the level of organs? Seen from today’s perspective, the basic block of type III artefacts will probably have to have most of the properties of stem cells in animals. From this standpoint, we see four essential threads of technology research that should form the basis for an integrated research plan. These are:

1. **Molecular Robotics:** exploration and design of useful materials lending themselves to build cells that can provide high mechanical stability (for “bones” and “joints”), energy transformation (for “muscles”), information exchange (“networks of nerves”), information processing (“neuronal assemblies”), etc.

2. **Distributed growable sensors:** for distributed areas of sensing “cells” that can sense force, light, odour, temperature. Of equal importance is the exploration of the role of preprocessing sensor data, (such as the preprocessing taking place on our retina).

3. **Growable distributed information processing:** this is a most demanding research area because the information processing system must control the growing artefact from the first moment of its “inception” on. It will need to grow in physical size as well as complexity and develop cognitive and motor skills in parallel with the sensors’ processing capacities. The challenge is not only to achieve a stable learning and growth behaviour for body control but also to develop new structural skills, e.g. “memory”.

4. **Growable motor entitites and spatially distributed actuators:** the actuators must be controllable as they develop both their actuation part (the mus-

cle portion) as well as the support structure (the skeleton/joint portion). Their evolution must be in sync with the size and mass of the artefact so that stability, locomotion effectiveness, energy efficiency and durability are achieved. Ideally, it will be possible to formulate principles which govern the growth processes in the artefact, such as the principles recently discovered for the development of the different types of neuronal cells.

From a technology point of view, we suggest a plan which centres about the basic building block (BBB) in view of the four aspects above:

Functional properties: what are the components the BBB consists of? What is the minimum amount of functions integrated into one BBB? Would it be possible to retain a certain amount of bodily plasticity/flexibility throughout the entire lifetime of the artefact?

Technological issues: how can the individual components be realised – and using what substrate material – including the ubiquitous question of a suitable source of power? Is it economical to use just one type of BBB that can differentiate into various uses or should there be more than one class of BBBs?

Interaction patterns: Studying the interaction patterns between individual parts over different communication channels is particularly important because, unlike with nanostructures whose interaction is completely static (i.e. binding forces), there can be a diverse range of patterns between the BBBs with different reach, with different time-scales, signal amplitudes, etc.

The development of convincing application scenarios should also be advocated in order to achieve useful deployment on the factory floor, in private homes, outdoor support etc., but it also involves the transfer of this technology to micoscale machinery for medical use.

Expected Results: What will it be good for?

In the following table (on page 16) you will see some of the possible applications of the spin-offs of potential research resulting from Factor-10 for adaptive and growing body structures. This table presupposes a development line from type II to type III artefacts with parallel basic research that in the first step is targeted at machines with relatively large BBBs using technology as available today, and then moves on to define the requirements for microscale BBBs, capitalising on nanotechnology.

Expected Result	Application of Result and Users
From Research targeted at type II artefacts	
Artefacts with early-cognitive properties such as context and attention-dependent visual scene analysis or with human-like pattern of intention-driven behaviour.	Applications that require only low-level adaptation to user needs, e.g. advanced human-machine interfaces.
Adaptive, cooperative prosthetics or physical support for senses, limbs or combinations.	Disabled and elderly people.
Artefacts with perception systems that share similar principles for human use and industrial automation and possess a high degree of robustness as typical of biological systems.	Medium and small scale production of goods not to be automated up to now. Revolution of the production of variants and a “batch size of one”.
Easily instructible “disappearing” robot systems for use in service (home and factory floor) that can adapt their body structure to become highly task-adaptive and that have some basic understanding of their own being there (self-awareness), react to and show emotions etc.	Small production shops and “home-workers”, new generations of handy “intelligent tools”, more demanding cleaning and housekeeping than just automatic vacuum cleaning, simple plumbing tasks, but also storage of all kinds of objects – even in small apartments.
From Research targeted at type III artefacts	
Artefacts that are capable of mind-body co-evolution and may adapt over a finite period of time to arbitrary environments (ultimate goal of the Factor-10 developments).	Unlimited range of applications. From microscale (inside blood-vessels) to creatures of animal-like shape up to free-form structures with intelligent behaviour and distributed sensing.
From Ongoing Basic Research	
In-depth understanding of the neural basis of human sensorimotor and cognitive processes and their development, the interaction of sensor/motor skills and the way mind and body interact during their respective development.	Researchers can simulate development (e.g. development of senses on fixed bodies and/or co-evolution of mind and body on growing structures) in a much more realistic way by using artefacts and test hypotheses on them; depending on the level of modelling-granularity as a supplement to animal experiments (in the long run possibly leading to a reduction of the need to carry out such experiments).
Basic Technologies in the field of: materials research, optoelectronics, sensors, actuators, information processing.	Industrial Automation Companies, Telecommunication Companies, new companies of still unknown profile.

Acting in the Physical World Project

Peripheral devices have been important areas of mainstream IT for decades. Strong trends towards embedded IT devices and pervasive computing (where almost any device can be fitted with chips connecting it to a network of other devices) are likely to increase the importance of periphery and system integration aspects even more in the future. NeuroIT can be seen as an attempt to close the sorely-felt gaps between natural neural systems and man-made artefacts by learning from nature. One hypothesis for this gap is that brains have computational capabilities for reasoning, planning, etc., vastly superior to man-made algorithms devised to reproduce these skills. When engineering a system intended to perform 'intelligently' in the physical world, designers have a wide variety of options. These can be grouped into four classes:

1. Choice of Computation and Control Strategies play a large role in determining the performance of a system. Depending on the task and circumstances, reasoning, planning, considered choices of action, or just reactive response to environmental stimuli may be necessary or sufficient to achieve the designer's goals.

2. Choice of Morphology. The right kind of body can be crucial. The body shapes of animals which live in the spaces between sediment particles are crucial to settling this environment. Likewise, the choice of wheels or legs for locomotion significantly influences accessibility and traversability of terrain.

3. Choice of Materials is crucial in high performance sensory and actuation systems. Sensing of mechanical waves, like displacement and acceleration detection or hearing of sound waves is accomplished in large part by the materials properties of the sensors.

4. The Environment itself can be engineered to facilitate the system's performance. Examples include marking the environment to store navigational information, such as pheromone trails or with street signs.

Natural neural systems are superior primarily because they are better integrated with respect to all these options. They are deployed in a system (for example, an animal) where coherent design choices are manifest across the whole space of options, rather than just at the computational/control level. Hence, the computational capabilities of the agent are distributed over the central nervous system, the peripheral system, the materials of the agent's body and the physical phenomena created by the interaction of the agent with its environment. Significant reductions in task complexity can be realized if each component in the control loop solves a simplified problem. Therefore, an intelligent, well integrated periphery may be the key to lowering task difficulty from 'impossible' to 'possible'.

One further goal of NeuroIT is to make IT artefacts sufficiently intelligent to interact with humans in a natural way ('Conscious machines' grand challenges) or interface successfully with the human brain as a useful replacement (prosthesis) or extension ('Brain interface' grand challenge). In either case, adequate periphery will be of prime importance.

Objectives

The objective of the 'Successful thinking and acting in the physical world' challenge is to build complete systems with distributed, embedded intelligence enabling artefacts to master tasks known to be performed by natural (neural) systems but currently elusive to technological reproduction. Research on this grand challenge will emphasise

- intelligent periphery
- system integration
- morphology and materials
- inspiration from the wide range of intelligent adaptations in non-human (neural) systems
- gathering and exploiting knowledge about the world and the tasks
- 'environment models' used to codify world/task knowledge

The focus of this grand challenge is on making the periphery smarter and integrating it better with central computations so the whole system gets more powerful and efficient. Knowledge about the tasks to be performed and the world they are to be performed in should be integrated at every stage.

Efficient ways to distribute the storage of this knowledge over different subsystems should be developed. This calls for advanced methodological achievements in gathering the relevant knowledge. Optimization processes in nature operate on large time-scales and vast numbers of prototypes for testing. For this research, shortcuts need to be found which narrow the search space so it can be managed within the scope of an engineering design process.

Research on this grand challenge will produce novel, smart peripheral devices for NeuroIT systems and thereby promote pervasive use of intelligent robotic systems. Work on the grand challenge should establish general design rules which may even be able to be applied to problems for which no solution is known to exist in nature. Research results should lead to the creation of universal standards for smart NeuroIT peripherals, which would enable closer cooperation between research projects. A pool of smart, readily available periphery should provide building blocks for powerful individual systems (robots) and establish new capabilities for robot interaction leading to the development of super-organisms.

Examples

Distributing intelligence over both central and peripheral stages should enable construction of 'minimalist' solutions thereby paving the way for cheap, low power yet capable robotic artefacts. Such systems should reproduce biological systems (in sensing, control, actuation and particular combinations of these) with the computing power of standard embedded systems.

Such artefacts could be made so ubiquitous we could have human-robot ecologies in which robots would enhance the quality of human life. Such symbiotic ecologies could be established in a variety of contexts, for example:

Smart home ecologies: Humans sharing their homes with unobtrusive creatures, which, for example, keep the house clean, establish and adapt wireless communication infrastructures and are rewarded with access to power.

Public spaces ecology: Perform cleaning of floors and windows, remove litter. Perform, for example, intelligent graffiti removal.

Office ecologies: Establish and continuously adapt communication/ teleconferencing setups, perform smart retrieval of tools and optimize the configuration of workspaces, office desks and storage.

Hospital/emergency room ecologies: Optimise sensors to monitor patients' health, fault detection, provide better comfort by reacting to symptoms of patients' discomfort with changes of environment (temperature, lightning, noise).

Communication ecologies: Optimise wireless communication channels to maximize transmission quality, minimise power consumption to increase battery life and minimise EMI related safety risks like interference with navigation or other vital systems. Future wireless communication devices carried by passengers on aeroplanes could intelligently adapt to the navigation/communication needs of the aeroplane, removing restrictions on their use.

Security ecologies: An alternative to fixed point surveillance cameras could be agile NeuroIT artefacts, which can change position and adjust to changing lighting conditions or respond to noise. If these agents are sufficiently agile, they could even escape attempts to disable them.

Playground ecologies: Improve recreational facilities for children and adults by making them more entertaining, more educational and safer by reducing the probability of accidents.

In order to achieve the those exciting goals, which will have such a dramatic impact on our day-to-day lives, we will be concentrating on several study areas. The study of simple 'organisms', both natural and man-made, that allow detailed analysis of their entire neural system, i.e. periphery and Central Nervous System, while performing natural tasks in challenging, natural environments or in a faithfully reproduced laboratory equivalent, is essential.

We will also be carrying out studies of non-human and possibly 'super-human' senses and actuation principles found in nature, to lay the foundation for artefacts which can not replace and surpass human labour by for instance being able to live in hostile environments. Finally, we will be studying 'environment models', i.e., finding the minimal amount of information necessary to get around intelligently in the environment, allow different organisms to integrate their peripheral and central processing into a control loop that efficiently guides them through their environment.

Current state of technology

The periphery and system integration of current NeuroIT artefacts is still lagging behind natural systems. For instance, in autonomous driving the limited dynamic range of cameras makes it impossible for these systems to cope with, for example, driving through a tunnel. Researchers in the field need ready access to much better embodiments or components. Projects in the life-like perception systems initiative, like Bioloch, Cicada, Circe, Cyberhand have been doing exploratory work on smart periphery and addressing the system integration challenge. This grand challenge should turn NeuroIT periphery into mainstream technology.

Projects like the DARPA-funded 'smart dust' project have been addressing the issue of deploying many cheap sensor modules. However, most research efforts seem to have been directed towards mass-manufacturing. The smart dust grains themselves could be a lot smarter. Mobile robotics is increasingly based on the use of GPS, restricting its application to environments where GPS-like information is readily available.

Problem areas

We have, of course, identified several areas that will need close attention as they present us with problems at present. These include the need to:

- Find ways to mass-manufacture and assemble the parts of advanced NeuroIT devices. For small structures, MEMS (microelectromechanical systems) technology may be a solution. Rapid prototyping technologies should be looked at in the context of using materials and creating shapes particularly well suited for NeuroIT artefacts.
- Find ways to analyze an organism's natural neural system while executing natural tasks in a natural, unstructured environment.
- Characterize and analyze the mostly unexplored physics describing the interaction between the organism's sensors/actuators, body and the environment during natural tasks.
- Improve understanding of design choices available, as a function of task, environment, cost and technology, and find ways to design a complete system.
- Develop novel sensor and actuator technology to support smart, biology-i nspired, peripheral systems.
- Application of non-linear dynamic systems theory for analysis of interaction organism and environment.

Future activities

A systematic effort should be undertaken to facilitate the development of next-generation

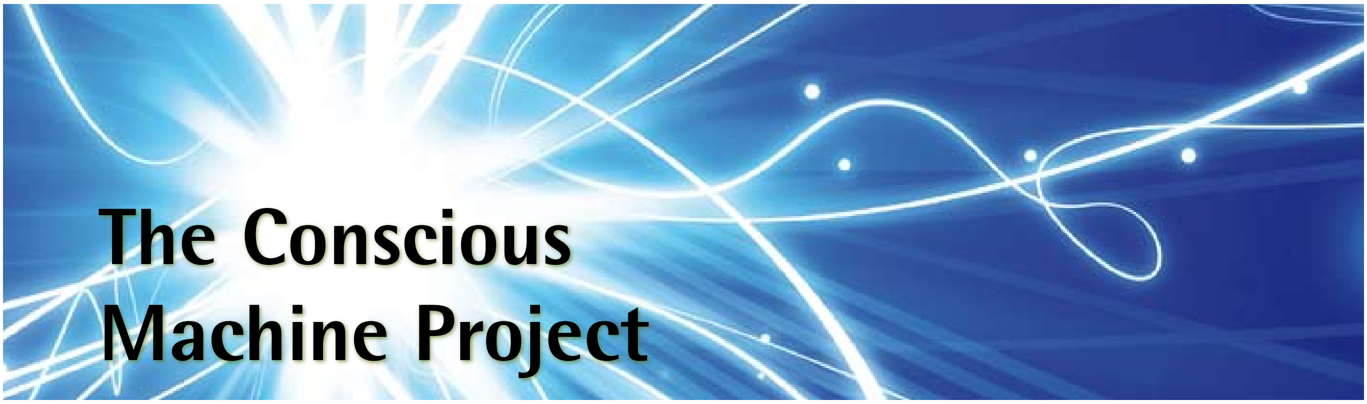
NeuroIT periphery. Today, embodiments are provided by small companies catering for the needs of experimental robots. These companies lack the resources for bold innovation and consequently their designs are conservative.

Alternatively, such embodiments are developed in research labs as one-of-a-kind systems which take a lot of man-power to develop and often enjoy very limited use. Remedies are needed to give researchers access to peripheral modules and system integration frameworks with capabilities and performance levels far beyond what is available today. Activities necessary to achieve these goals include:

- Benchmarking standards to stimulate and monitor improvements of NeuroIT systems.
- Standardize general, flexible protocols for interfacing NeuroIT periphery, to foster exchange modules between researchers and prepare the ground for industry standards.
- Organize an ‘organ donor data base’ for NeuroIT components (periphery as well as computation and control modules) to facilitate the exchange and reuse of existing periphery by researchers.
- Establish a repository (‘Noah’s ark’) of reference implementations, where information about periphery modules and system integration frameworks is stored along with a physical prototype.
- Establish shared manufacturing facilities (probably by way of cooperation with industrial partners) which make manufacturing technology specifically developed or adopted for building next generation NeuroIT periphery available to the entire research community in a cost effective manner.

Ethical considerations

Deploying capable, pervasive NeuroIT system within human society poses risks of failure and misuse. Ultimately strategies will have to be developed to make such systems failsafe. Suitable concepts for tackling these issues may again be inspired by nature.



The Conscious Machine Project

The last 10 years have seen rapidly growing interest in the scientific study of consciousness. Two of the most important research strategies are the neuroscience and the constructivist approaches (constructing meaning from current knowledge structures). It is this second approach which inspires this Grand Challenge. In brief, the scientific study of consciousness should be based on the investigation of physical systems and its final goal should be the construction of conscious machines.

Modern ‘service robots’ can perform a broad range of useful tasks. And yet it is very hard to adapt robots for tasks which deviate from the functions for which they were designed – or changes in the environment; we lack effective techniques allowing robots to communicate and cooperate with humans and other complex systems.

Because of these weaknesses, many people think robots are ’dumb’. The premise underlying this Grand Challenge is that robots need consciousness – a cognitive architecture that includes reflective control mechanisms akin to human introspection.

Consciousness can play many different roles. It can help the machine select sensory information relevant to its goals and motivations, reducing the amount of sensory-motor information it needs to store. It can help it to learn and thereby reduce the burden on programmers.

If these benefits of consciousness exist, they have been around for billions of years. Why, then, is it that only in the last decade or so people from other fields than philosophy of mind are starting to take in interest in it? The dramatic increase of computer power over last decades has not solved all our problems. Meanwhile advances in non-invasive imaging techniques, and more refined application of single-electrode measurements are starting to reveal the neural substrate of processes that were previously described in very abstract and generic terms. As a consequence, it is now possible to relate subjective notions such as ‘awareness’, to specific neuronal mechanisms. One of these neuronal mechanisms is ‘attention’.

Attention plays an important role in autonomous agents: it helps select those parts of sensory input which are behaviourally relevant, suppressing the rest.

Consciousness is a more complex phenomenon than just attention, although attention plays an important role. Already several attempts have been made to create a classification system for consciousness and the results are very controversial.

Advances in our understanding of the brain show consciousness is an essential ingredient for the performance of humanly simple but computationally complex tasks. This raises at least the question of how the property may be transferred into the machine domain in order to achieve similar competence. A study of this transfer may be called Machine Consciousness.

Such a study will require systematic analysis of signal-processing and conceptualisation in biological systems in order to effect action, enabling the transfer of these mechanisms to artificial systems. The key question is this. How can we design generic architectures that allow an autonomous system to develop its own cognitive structures? We have to bear in mind that human cognitive architectures are solutions to the problem of gene transmission and artificial systems should have very different user-determined goals.

Conscious machines could take many different forms. These could include self organising agents that develop in different environments and acquire skills and motivations which were not entirely predictable at the time of design; epigenetic conscious agents capable of entering into social relations with their human owners; intelligent 'situated artificial communicators', for situation-dependent human machine interfaces, for example; 'mentally adaptive' robot systems with new problem-solving abilities; complex technical systems that explain their 'state of mind' to a human user and understand what he/she feels. Early products are likely to focus on the enhancement of industrial robotics and driverless transport systems, service robotics (mobile service guides for customers in a supermarket), and edutainment.

Motivation and Objectives

Over the last 10 years, there has been rapidly growing interest in the scientific study of consciousness. The *Tucson Towards a Science of Consciousness Conferences* (1996-2002) helped create the right climate, playing a role similar to that of the *Macy Conferences on Cybernetics* (1946-1953), which prepared us for cybernetics and artificial intelligence.

Consciousness studies embrace a broad range of research strategies. Two of the most important are the neuroscience and the constructivist approaches. The first is summarised in a Nature Neuroscience editorial: "By combining psychophysics, neuro-imaging and electrophysiology, it will eventually be possible to understand the computations that occur between sensory input and motor output, and to pinpoint the differences between cases where a stimulus is consciously perceived and those where it is not." (Jennings, 2000).

In a recent book, Edelmann and Tononi, sketched out the alternative 'constructivist' approach: "To understand the mental we may have to invent further ways of looking at brains. We may even have to synthesise artefacts resembling brains connected to bodily functions in order fully to understand those processes. Although the day when we shall be able to create such a conscious artefact is far off, we may have to make them before we deeply understand the processes of thought itself." (Edelman & Tononi, 2000).

It is this second approach which inspires this Grand Challenge. Over the past few years, robots' computing capabilities and the quality of their mechanical components and sensors have increased to the point that 'service robots' can perform tasks such as transporting materials and tools in factories, providing delivery services in hospitals, performing household cleaning chores and undertaking underwater inspections. Yet despite these successes, achievements have fallen short of expectations. There are three main reasons"

1. Inadequate control strategies. Control strategies are inadequate for real-time tasks.

2. Poor adaptivity. It is very hard to adapt robots to tasks which deviate from the functions for which they were designed, or to changes in the environment.

3. Poor communications with humans.

Because of these failings, many people think robots are ”dumb” and do not believe they can provide useful services. The premise underlying ’Conscious Machines’ is that machines with real autonomy and adaptivity, which can communicate with human beings, will require consciousness – a cognitive architecture that includes reflective control mechanisms akin to human introspection.

Consciousness can help machines select the sensory information most relevant to their goal and motivations and help them to learn. Dancers train themselves consciously to move according to certain rules before acquiring the ability to perform the movement automatically, almost outside awareness. Conscious machines should possess this same self-teaching capability. Consciousness will reduce the burden on programmers.

Some of those who study machine consciousness lean towards the ’access’ side where action and behaviour is important, while others lean towards the ’phenomenal’ side, where the concept of self-awareness and internal representation of reality are important. The Grand Challenge is to develop systems that contain the balance between these two.

There are good technical reasons for trying to implement access consciousness:

- **Performance.** To maximise performance, robots need inner control loops which optimise the use of internal resources that can adapt the inner workings to optimally use resources. Self-awareness mechanisms are present even in the smallest and purest ICT systems.
- **Trust.** Before we trust a complex software system we require justification: the system has to be capable of explaining what it is doing. For ’justification’ the system needs introspection: access to its own ’thought processes’.
- **Robustness.** Robust ICT systems need to observe their interactions with the environment, understanding changes in their own state and that of the world in terms of the mission they are trying to fulfil. Systems need self-awareness to identify behavioural mismatches and diagnose and repair their own mistakes.
- **Cost.** IBM has made the cost argument for self-awareness very clear in its autonomic computing initiative: “Quite simply, it is about freeing IT professionals to focus on higher value tasks by making technology work smarter, with business rules guiding systems to be self-configuring, self-healing, self-optimising, and self-protecting“.

To achieve access consciousness, machines will need to be aware of many things. Some of these are external to the machine: aspects of the world that the robot has to perceive, if it is to interact with and/or adapt to them. Exteroception is an essential component in external control loops (mental processes that determine what the robot has to do to reach some desired world state).

But the robot also has to be conscious of its own inner mechanics – it needs propioception. Proprioception is the basis for the inner control loops which give the robot its autonomy. In brief, access consciousness is a sensor fusion mechanism, which creates integrated representations of the self and the outside world. The result is self-awareness.

One approach to machine consciousness would be to produce a mathematically rigorous and objective definition of consciousness, and then implement the theory in a model or a cognitive architecture. The alternative is to design and implement cognitive architectures that meet a set of performance requirements, specifying desirable features of conscious human behaviour. It is this option we propose here.

To achieve it, we will have to achieve a number of secondary objectives. In particular we will need to develop integrated software/hardware platforms (not simulations) where key ideas can be tested. Such platforms – which might include novel sensor and effector components – will require the ability to register the full complexity of the environment and to generate the very complex sensorimotor patterns, required to implement far-reaching actions or behavioural sequences

Examples of Applications

Conscious machines could take many different forms. For example:

- Self organising agents that develop in different environments and acquire skills and motivations which were not entirely predictable at the time of design;
- Epigenetic conscious agents capable of instantiating social relations with their human owners (consequently producing interesting opportunities for the consumer market);
- Intelligent ‘situated artificial communicators’ – for situation-dependent human machine interfaces;
- ‘Mentally adaptive’ robot systems with qualitatively new problem-solving abilities;
- Complex technical systems that can explain their ‘state of mind’ to a human user and understand what he/she feels (communication between different ’kinds of consciousness’).

The advent of these conscious machines will inevitably give rise to new products and services, unthinkable today. It is not difficult, however, to imagine some of the possible applications. The implementation of self-awareness, for example, will lead to major improvements in industrial robotics (programming interfaces through the integration of images and language; learning of complex action sequences, for example). The main marketing problem will be the creation of attractive prototypes, acceptable to potential customers.

There are many possible applications for adaptive service robots, too – many of which lend themselves to practical demonstration, with a good chance of attracting public attention and a large market. A typical example might be a navigation-capable artificial porter at a large airport which could carry a passenger’s luggage to the desired check-in desk, observing and adapting to the behaviour of the customer.

Conscious machines could also be used in animated displays in science centres and theme parks to enhance both education and entertainment.

Mechanisms of self-awareness can also be used to control complex technical systems, such as chemical plants, or electrical power grids. The systems could monitor their own performance.

In general terms, there is one class of machine that could leverage 'machine consciousness' more than any other – autonomous systems. The obvious examples are planes that fly themselves or cars that park automatically. Machine consciousness could also be useful for machines whose environment is very different from those where human consciousness evolved: avatars inhabiting virtual worlds, game playing engines, telecom infrastructures, nuclear reactors, hard drives, are good examples/

A business case for Consciousness Machines

Is there any real business case for conscious machines? Indeed there is – and not one but many business cases. Let's mention just two in quite different niches and then do some analysis that may serve as a general business drive for this technology.

The first business case is the case of the software systems we use to support human activity processes on our laptops, PDAs and mobile phones. There is no longer a central deity that decides when to release new updates of this software so we have to take on the complicated task of keeping our working environment in line with evolving realities. This new Flash 8 media file which can't be properly executed on my Linux Firefox browser; this just-released sequencer plug-in that my OS X music software rejects to incorporate. All are changing in a world without a coherent configuration management. There is no single authority that can do that.

The second business case is the case of electrical system internetworking between countries. National power production plants, transport and distribution grids are operated by companies or governments that have quite different objectives and strategies. Cross-border interconnection seems necessary from many points of view but from a purely technical point of view, the task of controlling such a system is hopeless. While the technical issues may be solved relatively easily (standardisation bodies do help in this) the main problem remains: integrated, unified decision-making. These processes are not only political or commercial decision processes but also include technical, even automatic, decision processes that happen ubiquitously in the network and that can produce electrical ripples that may manifest catastrophically in a remote place. The butterfly effect is becoming a daunting fact of our infrastructures.

The question is for a technical system to be able to reason about i) how it is able to think and act ii) how others do the same and iii) how can I communicate with them to achieve my objectives. Some of these topics have been addressed in the agents community but agent technology still lacks the level of self-awareness that is needed to properly drive these processes.

The business case is clear. Software intensive systems – real-time or not – are getting so complex we're no longer in the position of fully controlling them and their environments to make them robust enough. The classic zero-defect engineering or replicative fault-tolerance approaches do not scale well to systems of such a size and such a rate of uncoordinated change.

The possibility we envision is also clear – make systems responsible for providing their function. Instead of having a single production engineer, producing either software or electricity- in charge of change let the systems take care of themselves. Make the systems self-aware. This is somewhat happening in the field of software (IBM's autonomic computing or Sun's conscientous software). We need it to also happen with physically embedded systems.

Current Technology and Knowledge

Drawing heavily on Aleksander & Dunmall, (2003), in the near future, even hardened sceptics expect to hear claims that non-biological machines have attained consciousness. Finally, it is becoming possible to have a serious discussion about the object of consciousness. Why, for instance, should a bat be conscious? How does the bat's consciousness differ from my own? How might a computational machine benefit from being conscious?

In the last few years, computer scientists, neurophysiologists, psychologists, philosophers and engineers, with very different views of consciousness, have been thrashing out their differences, developing a shared understanding of what it means to be conscious. There have been a growing number of conferences and projects dedicated to 'conscious machines'.

In 2002-2003, two calls for projects (FET, Future Emergent Technology), 'Beyond Robotics' and 'Presence' from the European Union explicitly encouraged projects to investigate 'machine consciousness', 'phenomenal experience in machines and robots', and 'machine awareness'.

Nobel Laureate, Gerald Edelman – who is very much a sceptic with regard to traditional Artificial Intelligence – has a goal to build an intentional robot capable of mimicking the neural structure of the human cortex, drawing his ideas from the theory of evolution. Edelman argues the survival of biological organisms depends on their 'values': for instance 'light is better than dark' or 'fear is to be avoided'. They are encoded in neural circuits in the organism's brain, biasing its behaviour to favour survival. To test these ideas, Edelman and his colleagues have developed the Darwin series of robots (Edelman & Tononi, 2000). Darwin III, for example, has a mobile eye and a moving, jointed, arm and learns that small bright objects are important and should be touched with the tip of the arm (finger). Darwin is not conscious but it shows how we can use machines to understand the role of concepts like 'value' in consciousness.

Another important US researcher in this area is Bernie Baars, a psychologist associated with Edelman's Neurosciences Institute in La Jolla. Baars has proposed that consciousness should be seen as a 'Global Workspace' within which incoming sensory information activates dormant memories – making the organism conscious of the input. Memory fragments activated in this way stimulate the organism to look for more input, activating additional memories and so on (Baars, 1997).

But if we are really going to claim we have built a conscious machine, it will be necessary to go further. If consciousness is something material, what we need to demonstrate is not just that our machine is 'functionally conscious' in the sense described earlier but that it shares substantial material properties with organisms we generally recognise as conscious.

Other examples of work which might be considered 'functionalist' are the work of Aaron Sloman and Ron Chrisley (Sloman & Chrisley, 2003), Haikonen's Cognitive Neural Architectures (Haikonen, 2003).

There is also research whose emphasis is largely with the material nature of mechanisms and what it is about these that can be said to capture the conscious state. This inevitably examines living neurological machinery for appropriate design clues (Aleksander, 2005). Examples of research that is more to the physicalist end of the spectrum are, for example the work of Taylor (2003) and (Aleksander & Dunmall, 2003). Taylor's model (CODAM: Corollary Discharge of Attention Movement) is based on the principle that without attention to an input there can be no awareness of it. The corollary discharge is a copy of the attention movement signal generated by an inverse model control system. The attention signal itself will have influence of the input which it will selectively amplify or suppress. The corollary discharge is a precursor to the contentful signal from lower sensory cortices. Taylor claims that it is this precursor to content which in CODAM gives the experience of 'I' (Taylor, 2003).

Aleksander (2005) has sought to identify mechanisms which through the action of neurons (real or simulated), are capable of representing the world with the 'depictive' accuracy that is felt introspectively in reporting a sensation. The model stems from five features of consciousness which appear important through introspection;

1. **Perception of oneself in an 'out-there' world**
2. **Imagination of past events and fiction**
3. **Inner and outer attention**
4. **Volition and planning**
5. **Emotion**

An implementation of this structure has been used in a variety of applications ranging from the assessment of distortions of visual consciousness in Parkinson's sufferers (Aleksander & Morton, 2003) to identifying the possibility of a brain-wide spread of the neural correlates of 'self', models of visual awareness that explain inattention and change blindness (Aleksander, 2005).

The research just described is only a small sample of current work. As Owen Holland of the University of Essex has written in a recent editorial on the Journal of Consciousness Studies (Vol 10, no. 4-5), "We cannot yet know how fast and how far the enterprise will progress, and how much light it will be able to shed on the nature of consciousness itself, but it seems beyond doubt that machine consciousness can now take its place as a valid subject area within the broad sweep of consciousness studies."

Future Research

Achieving the goals of 'Conscious Machines' requires analysis of biological systems at the signal-processing and conceptualisation level, enabling the transfer of these mechanisms to artificial systems. Although an understanding of attention by itself is not sufficient for understanding consciousness, most researchers agree its role in selecting relevant and suppressing irrelevant information is an absolute prerequisite for humanly simple, but computationally complex behaviour.

The co-ordinated development of sensory systems, cognitive activities, and effector capabilities during the lifetime of the artificial system and over multiple generations (the equivalent to biological evolution) is also vital, while the control and/or exploitation of the growth dynamics of the sensorial system and external structures (morphology) needs to be developed.

Meanwhile, the creation of cognitive mechanisms such as attention control, complexity reduction, category formation, concept learning, object naming, and transformation of knowledge into forms intelligible to humans is also be necessary for robots to function consciously.

There are some key questions that need to be answered before these functions can be realistically realised

- To what extent is physical structure responsible for the sequence of cognitive processes or for their development?
- How are representations created and how do they interact with the machine structure? What knowledge should be hardwired into the machine?
- What does the machine need to know about itself?
- What representations will be required for self-awareness?
- How can these representations be used to support model-based behaviour? What strategies should it use to explore the environment ?
- How does a machine optimally attend to specific elements of a complex sensory world?
- What information can it gain by fusing internal and external knowledge?
- How do meanings arise, and how are they connected with sensorial and behavioural patterns?
- How can machines exchange experience, when they may have very different bodies?
- How can we control such machines – adapting their 'minds' to perform new tasks in new environments.

The answers will help us answer the broader question underlying the Grand Challenge, namely: how can we design generic architectures that allow an autonomous system to develop its own cognitive structures, creating a model of the world, a model of itself and mechanisms to exploit these models? In proposing answers to these questions, we have to bear in mind that artificial systems should have very different user-determined goals than human systems. To quote McCarthy, "The useful forms of computer agent self-awareness will not be identical with the human forms. Indeed many aspects of human self-awareness are bugs and will not be wanted in computer systems."

Relation to other chapters

Clearly, there are strong interdependencies between this chapter and other chapters. Before 'machine consciousness' will become a regular engineering discipline, we will have to make considerable progress in the understanding of cognition. In a certain sense this involves almost all other chapters.

An idealised approach would be: first we understand how the brain works, then we try to find abstract general principles, which underlie 'natural computation'. Here we expect consciousness to pop up somehow, among other high-level cognitive concepts. In order to improve our understanding of these principles, we implement them in technical systems, which should function precisely in those situations where adaptive, flexible, intelligent behaviour is required.

Important research will certainly proceed along these lines, but there are two reasons to consider possible alternatives: in the first place, it may not be simple to find abstract general principles that hold universally. Human behaviour in any given cognitive task involves a large complex of neuronal structures, perception, action and physical interaction with the outside world. Who ever said that it would be simple to produce technical systems that will function like humans or animals?

Will the problems in constructing such artefacts ever justify creating them for technical applications? It may well be possible that a rough idea on how some aspects work may be very beneficial in the creation of technical applications. Franklin's work already is an example of that. The investigations of attentional structures and cognitive architectures, such as, for example, CODAM, and EU project ALAVLSI, or various vision projects which take inspiration from visual cortex, suggest a rough understanding of their function may be a guide in creating novel technology, even if the fundamental neuroscience details have not been figured out completely.

So, while an overall understanding of the basic principles of the brain is desireable, there is no reason at all to sit around and wait until this has come about.

Research that would take a 'basic understanding first' approach clearly would push for advances in fundamental and cognitive neuroscience and psychology.

Research that would take an 'inspired approach' may roughly be divided into two categories. We can learn much from non-invasive imaging techniques, because they can be used to study large-scale functional neuronal networks. This will provide new insight in the architectural organisation of the brain and the dynamics between brain areas. Again the relation to 'constructed brain' seems quite prominent. Research in the area of Brain-Machine Interfacing (*see the Brainship Project*) may also be very important here, because it may help to show how higher level cognitive processes interact with the outside world and how consciousness plays a role in this.

This is some sense a 'top-down' approach. A second approach, 'bottom-up' would in fact start with technical systems and perhaps only very high level abstract ideas on how consciousness could be used in technical systems. Franklin (2003). Such technical systems could be simple at first, but become more complex, designed specifically to function in those situations were we expect 'machine consciousness' to be useful (and which therefore would need to display, adaptive, flexible, intelligent behaviour). This may lead us to discover machine equivalents of awareness.

The Artificial Evolutionary Design Project

Artificial Intelligence (AI) has not lived up to its initial promise as systems that performed impressively in “toy worlds” failed to scale up to real-world applications. Similar problems face so-called new AI. While biologically inspired approaches such as neural networks, evolutionary computing, and evolutionary robotics have produced interesting results, applications have been highly specialised and had only limited impact on the engineering community. AI applications such as expert systems, machine translation, data mining, scene analysis, robotics, and intelligent software agents have been less successful than once hoped. In many domains artificial systems have yet to come close to the routine performance of humans and even relatively simple animals. Molecular and neuroscience investigations of real-life systems suggest that the mechanisms required cannot be compressed into a compact piece of code. If we want to implement artificial systems that emulate human and animal cognitive competencies we will have to implement very complex algorithms. Current techniques in software engineering are inadequate to this task. Paleontology has shown that nature can, when required, evolve complex structures and behaviours in what, in geological terms, are very short periods of time. Examples include the rapid evolution of human intelligence from that of simpler primates (in less than 10 million years). It would thus seem logical to look to nature for guidance on how to design complex artificial systems. Current attempts at biologically-inspired design have only scratched the surface of what might be possible.

Objectives

This Grand Challenge’s long-term goal is the production of artificial cognitive systems capable of performing useful tasks effectively performed by human beings and other animals. The project is based on the premise that current techniques of software engineering are inadequate for the design of highly complex artificial cognitive systems and that to build such systems we have to draw inspiration from the way in which nature designs complex systems and behaviours. This project, therefore, aims to:

1. Develop mathematical theories and models of the evolution and dynamics of natural Complex Adaptive Systems (CAS);
2. Validate these models through simulations of known biological systems and processes;
3. Investigate the computational complexity of the models developed under point (1).

4. Design and implement tools for the automated design of Artificial Cognitive Systems based on (possibly simplified) versions of these models;
5. Demonstrate the engineering effectiveness of these tools with respect to benchmark problems (perhaps generated by the other Grand Challenges) which are hard or impossible to resolve with existing techniques of manual design or artificial evolution.

Examples

The ultimate goal of the project is to develop models and tools allowing the automated design of Artificial Cognitive Systems in low-order polynomial time. Examples of possible realisations include:

1. 'Elastic' designs for autonomic (automatic) robots, whose body plans, sensor configurations, actuator configurations and processing capabilities can be easily evolved (or trained) to meet the requirements of specific industrial and domestic applications;
2. Highly flexible software for pattern recognition and categorization, which can be 'evolved' or 'trained' to solve currently insoluble tasks in data mining (extraction of useful information from large data sets), scene analysis, speech recognition etc.;
3. Development environments for the automated design of complex electronic circuitry and hardware to be incorporated in Artificial Cognitive Systems.
4. Generic models for the development of domain and language-specific machine-translation tools.
5. Self-adapting systems for the protection of autonomous systems against threats, which have not been specifically identified at the time of design.
6. Hybrid chemical-computerised development environments, for the artificial evolution of complex 'wetware' e.g. bio-electronic sensor, actuator and cognitive systems.

Current state of technology

The concept of biologically-inspired automated design is not new. Artificial Neural Networks, first conceived in the 1940s, attempted to model animal and human learning as an alternative to explicit design; Evolutionary Computing, originating with Rechenberg's Evolutionary Strategies of the 1960s adopted an alternative approach, drawing inspiration from Darwinian models of evolution in populations of organisms. Work in the 1990s and the 2000s moved beyond the symbolic world of computers towards the automated design of physical, chemical and biological artefacts: electronic circuitry, autonomic robots, and biologically active molecules for use in the pharmaceuticals and materials industries.

Artificial Neural Networks (ANNs) are based on highly abstract models of biological neurons and synapses. The defining feature of modern ANNs is that they can ‘learn’ and ‘generalize’ from what they have learnt. In other words you don’t have to explicitly program an ANN to perform desirable tasks: ’learning’ algorithms allow human operators to ‘train’ networks to recognise patterns. Common applications include Optical Character Recognition (OCR), analysis of Mass Spectography Data for explosives detection and medical diagnosis, as well as data mining for commercial and medical applications.

Evolutionary computing adopts an alternative approach to automatic design. In this approach, randomly generated ‘artificial organisms’ are tested for their ability to perform a task defined by the operator. The organisms that perform the task most effectively produce ‘offspring’ which inherit the characteristics of the parent organisms and can also effectively perform the task assigned by the operator.

In Evolutionary Strategies individuals are represented by a vector of real values (e.g. the parameters describing the shape of an aircraft wing). The key genetic operator is a Gaussian mutation (an evolutionary algorithm), in which a (small) random value is added to one of the elements in the vector. Applications of evolutionary strategies have focussed on problems such as the design of aerodynamic surfaces, road networks and the solution of the vehicle routing problem.

Genetic Algorithms, introduced in the 1970s, represent competing problem solutions as a rich and flexible coding scheme which allows a natural implementation of ‘genetic operators’. There is an extremely broad range of applications ranging from the evolution of control systems for autonomic robots to job scheduling, mechanical component design, and the design of VLSI (very large-scale integration) circuits.

Genetic Programming adopts an alternative approach, in which the evolutionary process constructs a computer program, represented as a ‘tree’ of functions and values. Like biological evolution, Genetic Programming is open-ended. Applications have included the automated development of control mechanisms for prosthetic limbs, and the discovery of classification rules for medical diagnosis. Recent research has moved into the area of Evolvable Hardware, Evolutionary Robotics, and Combinatorial Chemistry

Evolvable hardware (EHW) combines computer-based techniques with physical tests designed to measure the performance of the resulting circuitry and to select the ‘fittest’ circuits for further rounds of evolution. EHW’s use of physical hardware is paralleled in Evolutionary Robotics where researchers combine computer-based methods and physical testing (or detailed simulation of robot physics) to evolve not only robot control systems but also robot morphology (body plan, motor and sensor placing etc). Much work in Evolutionary Robotics is designed to exploit the physics of the physical machinery in which designs are implemented.

In a completely different domain, molecular biologists have recently made great breakthroughs in the field of combinatorial chemistry, which involves the rapid synthesis or computer simulation of a large number of different but structurally related molecules. Combinatorial Chemistry, which represents a first step towards the ”artificial evolution” of useful molecules, is now standard practice in the pharmaceutical industry (combinatorial drugs for pain, cancer, HIV, lupus, and asthma are currently in clinical trials) and is being watched with keen industry by scientists seeking to develop innovative materials for use in other fields such as communications and electronics.

Problem areas

Each of the technologies, described in the previous section has produced impressive results in the laboratory but in the majority of cases industrial applications have been highly specialised and of limited economic significance. One intrinsic weakness in current technology is its lack of resemblance to natural processes. Other weaknesses are summarized below.

1. Compared to the systems we would like to emulate, or develop, the majority of systems studied by current research into ANNs, evolutionary computing, EHW, Evolutionary Robotics or Combinatorial Chemistry are extremely small. The robots developed by evolutionary robotics do not go beyond “insect intelligence”; the circuits developed by EWH are elementary in the extreme, combinatorial chemistry is limited to the synthesis of single ligands (atom, ion or molecule) for individual molecular receptors.
2. Experimental experience shows the time required to ‘train’ ANNs or ‘evolve’ other classes of evolutionary system, increases rapidly with the size of the problem to be resolved. This suggests that even rapid increases in computing power will have a limited impact on what can be attacked now.
3. In the majority of models there is no differentiation among the sub-units. Such approaches fail to model the complex interactions among diverse agents (molecules, genes, neurons, organs etc.) which are an essential characteristic of biological systems. As a result, there are no internal constraints on the way in which the CAS (complex adaptive system) can develop during training or evolution.
4. The majority of current techniques in ANNs Artificial Evolution pre-define the basic size, architecture and morphology of the ‘organism’ they are training or evolving leaving no room for the open-ended evolution which characterises the development of life on earth.
5. Virtually all current ANN or evolutionary algorithms emulate a single period of ‘training’ or ‘evolution’. In nature, organisms experience long sequences of learning and adaptation constantly building on the results of earlier phases. The failure to model this limits the ability to design complex systems.
6. In the majority of evolutionary models, there is no distinction between genotype and phenotype (the difference between an organism’s heredity – genotype – and what that heredity produces – phenotype) and thus no room for artificial organisms to ‘develop’. This makes for inefficient coding schemes that contrast with the coding efficiency of animal genomes. In biological elements genetic networks can code for growth processes which can involve large numbers of differentiated cells. No equivalent process occurs in AE.
7. The majority of evolutionary models also make no attempt to model the “regulatory” elements which govern gene expression in biological organisms. This absence, together with the failure to model development, leaves little room for the developmental plasticity that characterizes biological organisms.

8. In reality, current models of Artificial Evolution fail, not only to model the regulation of gene expression but to include any biologically realistic mechanism for macro-evolution. Mainstream models make no attempt to model duplication of genes, chromosomes or regulatory code – a common event in biological evolution – which many theories believe is a pre-condition for the development of novel function.
9. With the obvious exceptions of EVH, Evolutionary Robotics and Combinatorial Chemistry, the majority of computer-based studies of Artificial Evolution pay little attention to the physics, chemistry and mechanics of evolving complex systems. This means they are unable to take advantage of natural self-organisation.
10. Most evolutionary models make no provision for the transmission of information between individuals. This theoretically precludes the emergence of 'major evolutionary transitions' which depend on the evolution of novel mechanisms of communication among organisms.
11. The majority of ANN and evolutionary models are single level, making no attempt to model the hierarchical organisation of real life biological systems into genes, gene networks, cells, organs, organisms, demes, or species. Evolutionary theory predicts that in the absence of this kind of hierarchical organisation relationships between genetically unrelated systems will be purely competitive.

The key factors hindering progress are partly cultural but primarily technical. From a cultural viewpoint, studies of Artificial Evolution have suffered from inadequate communications between different disciplines thanks to separate conferences and the failure to develop interesting lines of research into sustained programmes of research. At a higher level, approaches whose founders drew inspiration from biology have tended to cut ties with these disciplines. But the real reasons underlying the lack of progress are much deeper than this. The biological processes Artificial Evolution needs to model and emulate are immensely complicated and difficult to model. Traditional evolutionary theory lent itself naturally to mathematical modelling. What is missing today, and what Artificial Evolution requires, is a 21st century theory of the evolution of biological complexity. Such a theory would make it possible to identify the intrinsic limitations of evolution, identifying and characterising problems that no evolutionary process – natural and artificial – will ever resolve.

Future research

Achieving effective techniques for the automated design of artificial cognitive systems is a long-term goal. In fact, it should itself be seen as a problem in evolutionary design. There is no hope of achieving practically useful systems in a single step. Inevitably, the development of such techniques will be an incremental process; problems will have to be resolved one at a time; much will depend on communications and symbiosis between different disciplines and approaches. To achieve our strategic goals, future research will have to develop a broad range of innovative techniques, each of which represents a scientific challenge of its own.

Challenges to overcome

1. Models incorporating mechanisms for change above the level of species, for example, modifications of regulatory mechanisms with visible consequences, symbiotic relationships and transfer of genetic material to another cell which is not its offspring; gene, chromosome and genome duplication.
2. Techniques for the automatic design of highly evolvable structures and behaviours with the ability to rapidly adapt to the requirements of a broad range of different environments.
3. Open-ended models in which large-scale adaptive change emerges as a sequence of adaptations to a changing environment.
4. Techniques for modelling "multi-level selection", allowing competing 'Darwinian individuals' (e.g. genes, organisms) to cooperate in higher level systems (organisms, demes).
5. Models which combine evolution and development, in which "artificial genomes" include both regulatory and structural elements, coding a development process rather than directly representing the phenotype.
6. Models of evolutionary and developmental change in which the 'search space' is constrained by "grammatical rules" (as in current genetic programming).
7. Techniques for exploiting the intrinsic chemistry and physics of artefacts produced via Artificial Evolution.

However, these techniques on their own are not enough. To make links between different teams using different approaches the project must identify benchmark problems which are sufficiently complex to be insoluble with current techniques yet which are sufficiently simple to give developers real hope of success.

Ethical considerations

The attempt to create design processes, with the capability to produce Artificial Cognitive Systems comparable to natural systems, is, like most of the challenges in this booklet, very close to an attempt to artificially create life: Inevitably, it contributes to the "disenchantment of the world" (Weber, 1919), to what Bill McKibben has called "The End of Nature" (McKibben, 1990) – the illusion of a human-generated, human-managed world, in which civilisation creates itself, independently of nature. As such there can be little doubt that it is ethically ambiguous, even dangerous. However, a theory of the evolution of biological complexity implies not only the ability to perform tasks which we are currently incapable of performing, but also new knowledge of the intrinsic limitations of human design; the theoretical characterisation of problems which are by their very nature intractable. Such knowledge is perhaps not without ethical value.

The Constructed Brain Project

During the last decade the amount of experimental data coming from the brain sciences has exploded as a result of new experimental techniques but our understanding of the fundamental principles on which the brain operates has not progressed at the same rate. We still do not understand the principles of cognition well enough to apply them to artefacts. But we are also not using the data available to us now as it is scattered over thousands of books and journal papers and thus hardly accessible for modellers. Much modelling in brain science is done on the basis of home-grown software and there is no software repository for cognitive modelling, which leads to enormous duplication. We only have to look at CERN (The European Organisation for Nuclear Research), with its well-maintained software infrastructure, to see there is nothing comparable in brain science. These problems are aggravated by poor communication across traditional scientific boundaries, leading to replications in research between the areas of Cognitive neuroscience and psychology, for example.

There is an enormous potential for Europe to make better use of experimental data simply by improving coordination of theoretical efforts to understand the brain. We suggest one way to do this is a project which involves a simulation of brain processes. Such a project may be the core of a much needed theoretical infrastructure for brain science.

Within one or two decades computers will be sufficiently powerful to implement brain-like computing of a scale and complexity matching the brain of large mammals, primates and eventually humans. These truly intelligent machines will have important applications in many sectors of society and will be of great commercial interest. In parallel, our understanding of the healthy and diseased brain will increase dramatically. This development is now in its infancy but it is already of strategic significance for Europe. It is crucial to strengthen the research basis within FP7. The design and use of these artefacts may to some extent also need to be regulated in order to avoid unwanted effect on human life and society.

In this chapter we sketch two ways to approach this Challenge: a ‘bottom-up’ development of ever-more realistic simulation of the brain’s neuronal networks, and a ‘top-down’ approach, focusing on the cognitive abilities as they exist in humans and animals and investigating how these abilities can be realized in neuronal networks.

Motivation and Objectives

Today we are close to having enough computer power to simulate a complete brain in considerable detail. We could call our simulation a ‘virtual brain’. If we integrate hardware devices into the computational architecture, we would have an ‘incorporated brain’. Or it might be possible to give the brain a body, creating an ‘embodied brain’. In more general terms, we can talk about a ’constructed brain’ and this is the name we have given to this Grand Challenge. Biological neuronal networks are huge. Artificial neuronal networks with comparable performance will need comparable numbers of units and connections. Most neuronal network simulations today are very limited in size, typically because researchers do not have the computing power for larger-scale work. But computers are becoming faster and faster and if current trends continue, it is likely that some time between 2015 and 2020 performance will reach levels capable of simulating networks the size of the human brain.

Today the key obstacle to the development of ‘intelligent’ applications is our lack of insight into why the brain performs cognitive functions so well and so fast. Is it because the brain is massively parallel, on a scale unmatched by current hardware or software? Or is it because of the computational architecture of the brain – which we are still far from understanding? Or could the crucial factor be spiking neurons and the way they code information? We need a comprehensive view of the processes that take place in the brain when it is performing cognitive tasks and when it interacts with the outside world. In a 10 to 20-year perspective the objectives are to:

- Bring the brain and information sciences closer together taking advantage of inherent synergies between the two fields;
- Strengthen theoretical and computational brain science;
- Investigate the fundamental principles of brain function;
- Design large-scale architectures and algorithms inspired by the brain;
- Develop dedicated hardware optimised for brain-size computation;
- Design artificial nervous systems and brains capable of perception, attention, decision-making, action and movement control, and learning.

Current technology and knowledge

Massively parallel computing is developing fast. In a few years we will have affordable compute boxes providing clusters of thousands of processors on the desktop. This will allow us to implement and investigate very large-scale, brain-inspired architectures and algorithms providing the logic needed to exploit the inevitably random nature of molecular scale computing.

At the same time, it is possible to study cognition at a higher level. Language, for example, has been notoriously difficult to implement in artefacts. Even basic aspects of language provide constraints at the neuronal level. Another approach could be to model behaviour first and move downwards to the neuronal level.

Large-scale biophysically based brain simulation

Measurable computational modelling is today accepted as an important tool in brain science and research. Models have been developed for many different levels of brain processing, ranging from the molecular processes underlying cellular and synaptic properties to brain-scale neuronal networks. A good review of simulations at the neuronal level can be found at http://www.neuroinf.org. These packages make it possible to set up sophisticated simulations. But with the exception of CATACOMB, they are very much ‘stand-alone’ tools, with few links to other programmes and databases. A number of publications have described sophisticated algorithms for the simulation of large groups of neurons but mostly the programmes used to obtain key theoretical results are not publicly available. On a higher level, the most common simulation tools are oriented towards artificial intelligence and machine learning, rather than high-level cognitive modelling. An exception is NSL, the Neuron Simulation Language which supports both ANNs, and neuronal simulations.

Taking a broader view of current simulation techniques, we see that researchers use two complementary modelling strategies. 'Bottom-up' simulations start from biophysically realistic models that mimic the system under study. 'Top-down' approaches use abstract models or pure mathematics to cast the general principles of the system under study into a minimal model. The central role of modelling is to bring together experimental data from different sources and demonstrate how seemingly unexplained phenomena are a consequence of what is already known. Exploration of the model can also produce truly unexpected findings, which inspire new experiments. Increased use of modelling will speed the build-up of knowledge and reduce the need for animal experiments.

It is sometimes claimed there is no point in building a biophysically detailed model of a brain scale neuronal network since the model would be as complex as the system it represents and equally hard to understand. This is clearly untrue. An exact quantitative model of a human brain would provide researchers with full access to every nitty-gritty detail in the simulated brain, dramatically speeding up progress in understanding. Much brain research demands computational modelling at the large-scale neuronal network level. We cannot represent the dynamics of a global brain network by modelling local networks, or by letting one model neuron represent an entire cortical column or area. In such models one significant problem is how to provide enough synaptic input current to activate model neurons. Researchers are forced to exaggerate connection probabilities or synaptic conductances, and most of the time both. This creates a network with a few strong signals whereas in real cortical networks many weak signals interact.

Another unavoidable but disturbing fact is that the shift from a single cell model to a large network with many different cell-types involves a rapid increase in the number of 'free' parameters. Any brain-scale network model would contain many billions of parameters. Fortunately typical neuronal networks comprise a limited number of cell types, each with roughly the same properties. Thus the parameters used for one neuron of a certain type are likely to do for the others as well. This is also true for synaptic interactions. Thus, the good news is that the number of truly free parameters is more or less independent of the actual number of neurons and synapses in a network. A huge network model can use the same averages, distributions and learning rules as a tiny one. The realism of the model increases as the number of neurons and synapses approaches those of the system being modelled. In really large networks there is little extra cost associated with complex cell models.

From brain simulation to brain theory and brain-like computing

Thirty years ago, there existed a plethora of theories about how the brain worked. Today, these have been reduced to a handful of qualitatively different hypotheses. This pruning is likely to go on until we can say that, at least in principle, we know how the brain works. Numerous papers have been published on the behaviour of individual neurons, beginning with the seminal work by Hodgkin and Huxley (1952). Today the literature contains literally thousands of papers. This is an important and very active line of research, involving large numbers of participants. On a somewhat higher level, some of the most important topics in neural modelling and theory are the working of the cortical code (rate coding, precise inter-spike times, the mechanisms underlying Long-Term Potentiation and Depression (LTP and LTD), the role of these mechanisms in learning, and the way in which cortical and subcortical structures use them to produce behaviour. Somehow we have to find ways of incorporating this information in higher-level descriptions of the brain. Even though it is becoming possible to simulate billions of neurons, this in itself is not very helpful. One would still have to extract information from the spike trains of these billions of neurons. Just as physics used statistical mechanisms to describe the macroscopic behaviour of large numbers of molecules in gasses, so neuroscience is beginning to use techniques from statistical physics to model the behaviour of large groups of neurons. Though this behaviour in response to stimuli is very complex, recent work has discovered, for example, that large groups of neurons can be described by powerful sets of equations, and that such descriptions can even incorporate a degree of neuronal detail. Although solving such equations is important, it is computationally much more efficient than direct simulation of a large group of neurons.

The same techniques have been applied to the cortical circuits, believed to underlie working memory, attention, and the formation of orientation columns in visual cortex etc. This work could possibly be the first important step towards the description of the large-scale cortical networks, identified by modern developments in fMRI. If we can find good descriptions of neural activity for higher-level cognitive processes, it may be possible to simulate fMRI and EEG signals.

In cognitive science, researchers try to simplify the computational models they use, so as to achieve better understanding of the system under study. To this end, they often use so-called connectionist models. For instance coordinate transformations between various frames of reference (head-centered, eye-centered) and long-term memory formation in the hippocampus complex, have been modelled using Perceptrons (binary classifiers) and back-propagation. In the future it may be possible to use findings from neuroscience to constrain these models, identifying the cellular and synaptic mechanisms which play such a critical role in global, network level phenomena. This would allow connectionist models to be used in a biological (and cognitive) context. Biologically constrained connectionist models represent brain structure at different levels of detail. At the most abstract level, it is still common to use so-called connectionist style models in which small local groups of neurons, for example, cortical minicolumns, are represented as one computational unit and connections between units may represent a bundle of synaptic connections. Given the number of minicolumns in real brains the this kind of model constitutes the key interface between brain modelling and brain-like computation.

It is likely that competitive learning in local modules, modelling cortical hypercolumns, or in different cortical layers, will be very important. Today, we are unable to describe these learning and adaptation rules in exact terms but we are making very rapid progress. Some of the most important challenges will come at the systems level, when we seek to compose known dynamic and learning principles into a functioning whole, with a complexity approaching that of biological brains. At the moment we have a long way to go even to match a rat brain.

The key issues these models have to face

- Sparsely and globally connected architectures, composed of modules and layers;
- Multiple modalities and cross-modality interactions;
- Perception-action interactions;
- Closely interacting short-, intermediate-, and long-term memories;
- Temporally fine-tuned motor control and learning;
- Goal-directedness and emotional/motivational aspects of attention, learning, decision-making and behaviour selection.

Theory

Finally, theory is bound to play a critical role in the determination of the computational architecture of the brain. The human cortex is remarkably uniform which suggests the brain's computational capabilities may rely on a relatively small set of cortical configurations. The idea is that the entire variety of complex computational tasks performed by the cortex might be based on a small set of basic 'cortical circuits'. There is relatively strong evidence that the visual cortex uses a so-called 'blackboard architecture' in which different high-level features of visual stimuli, such as colour, form and motion are processed by high-level visual areas. Feedback information from higher to lower visual areas can lead to a re-evaluation of information in lower visual areas. Several researchers have suggested similar principles could also be involved in language processing and production. The investigation of 'computational architectures' like these is important, because it relates to future hardware implementations. If the cortex uses a small number of 'computational architectures' and if we can understand how these architectures function, we should also be able to understand how a massively parallel structure of relatively slow elements can perform complex computations. The road would be open for the construction of brain-like, genuinely intelligent artefacts. Obviously, this would be a technological breakthrough of great significance.

Databases

Theoretical modelling is constrained by data so it is essential that researchers should have access to the data they need. The number of databases, and the variety of data available on the web is astounding. For an overview see http://www.neuroinf.org. Some of these databases are designed very professional. One example is COCOMAC (Stephan et al., 2001), which provides extensive information on macaque brain connectivity. Another website, created by van Essen and co-workers, provides extensive information on surface-based atlases of human and macaque cortex. fMRIDC, an initiative announced in the *Journal of Cognitive Neuroscience,* is an interesting attempt to create a database that can be used for reanalysis of fMRI data. fMRIDC invites authors to submit datasets, supporting their publications. Other databases cover topics as varied as hippocampus neuron data, ion channels, cortical connections in cat areas and so forth. But in general, the quality of publicly available databases is poor with many broken links. Few conform to the high standards set by COCOMAC for example. Recently, however, there have been a number of attempts to standardise data formats. For example NeuroML and BrainML (http://brainml.org (its American counterpart) look beyond single problem domains - something very few other initiatives even seek to achieve.

Dedicated hardware

IBM's Blue Gene computer family is currently the most brilliant example of a trend for more processors in parallel computers. As designers incorporate increasingly powerful compute nodes, performance will continue to rise. Meanwhile, the extension of parallel computing towards the consumer market, currently in its initial stages, can be expected to produce a drop in prices. With a hundred compute nodes on one of today's cluster computers, it is already possible to simulate networks a third the size of the mouse cortex. Thus, in the near future, we will have sufficient computer power to build large, and even full-scale models of global brain networks.

Traditional computers designed for high-precision deterministic computing are not ideal for running brain-sized computation. In the future, we should design compact, low-power dedicated hardware, perhaps using molecular-scale substrates with computational characteristics similar to those of real neuronal networks.

Future Research

To move as fast as possible towards the goals of the Grand Challenge, it will be necessary to concentrate research in a number of priority areas. These include:

- Theoretical and hypothesis-driven approaches to brain science, taking advantage of computational modelling and associated hypotheses and theories;
- Hardware for massively parallel and scalable computing;
- Development of advanced scalable simulators for computational neuroscience;
- Robust low-precision molecular scale stochastic computing;
- Learning and dynamics in scalable spiking neuronal networks;
- Recursive network-of-network architectures, merging attractor networks, competitive learning networks and other components into a common framework;
- Temporal aspects of brain function, sequence learning, predictive control, role of cerebellum;
- Memory systems with short-term, intermediate term and long-term memory;
- Motivational and emotional mechanisms underlying decision making, attention, gating of learning etc.

In the meantime it would be useful if we could start work immediately beginning with small-scale projects. The goal of these projects would be to model relatively limited aspects of human cognition, or to emulate parts of human cognition or motor behaviour in hardware (artificial retinas, cochlear implants, robot arms/hands). The most important deliverables from this kind of project, should be software libraries which conform to high-quality standards for code, interfaces, documentation, and maintenance and are verified by independent experts. This would be a significant step towards re-usable models which could be used as building blocks for more complicated models or re-implemented in hardware. For instance, a retinal implant could be interfaced with a model of early visual processing; sensors in an artificial hand, could be interfaced to a model of sensory-motor cortex, a model of visual processing could be interfaced with a model of auditory processing. The emphasis on software libraries, would distinguish this 'start-up' project from other projects in the same area of research.

Aggregation of databases with data for different levels of brain organisation
This work should bring together databases providing neuronal data as well as data on small scale brain structures (e.g, details of the structure of a cortical column), structural and functional connectivity, and large-scale cortical structures. These databases already exist or are in the process of being created. The goal would be to allow modellers to access data for all levels of brain organization via a single tool.

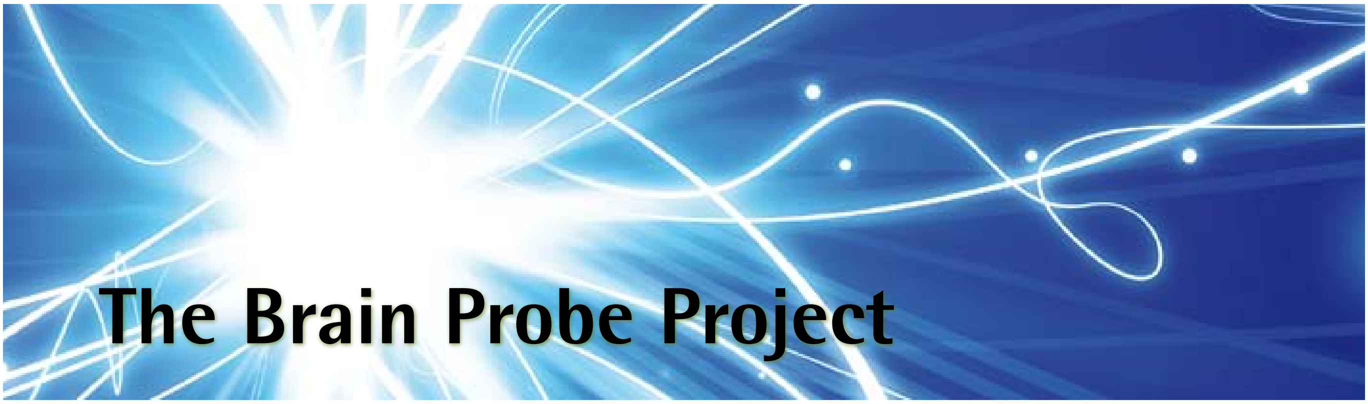
Creation of an external environment for the brain to interact with
Starting with simple, 'abstract' simulations of the 'real-world' and moving on to 'real' sensory input and motor output.

The development of theoretical methods
These would be capable of bridging the gap between phenomena at different levels in the organisation of the brain. One example might be statistical mechanical methods describing the collective behaviour of large groups of neurons.

Investigation of computational architectures
These would be done from a theoretical and an experimental point of view. If, as we have argued, there exists a 'cortical principle' of parallel computation, or at least a few relatively simple principles, we could implement these principles in hardware and select the most suitable hardware for the emulation of specific cortical functions.

Creation of visualisation tools
These would provide an overview of the 'constructed brain', at every possible level.

Finally, if this research is to be effective, it will be necessary to break down the cultural and academic barriers between brain science and information technology. The best way of achieving this is probably through the promotion of interdisciplinary activities in computational neuroscience – an area which brings together experimental neuroscientists with a strong interest in synthesis and theoreticians and engineers with good knowledge of biological nervous systems and brains. It is essential that FP7 should contribute to facilitating this communication. There is much we can learn from other sciences which have established large multi-disciplinary collaborations. This impressive computing infrastructure has been developed by many people, from different disciplines, working together in a single highly ambitious project to a good end. This is exactly what we are seeking to achieve in this Grand Challenge.



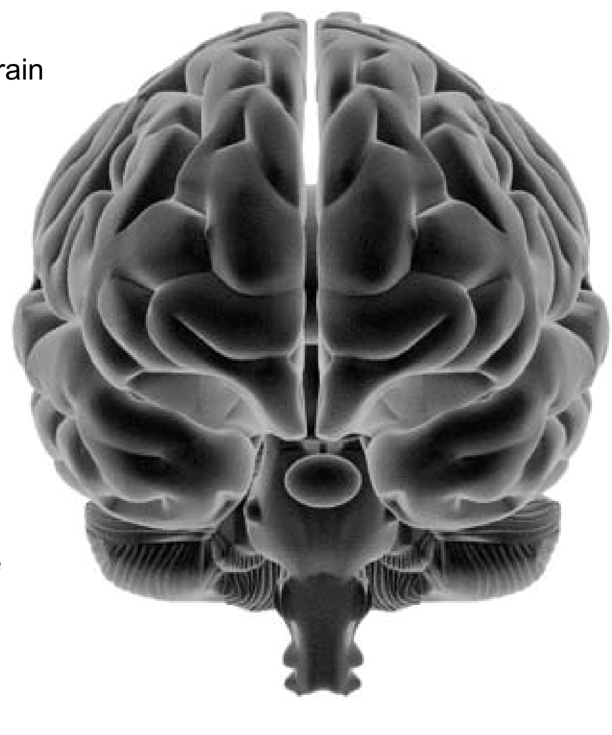
The Brain Probe Project

It is now accepted that Information Technology falls far short of our requirements. The robotics dream of cognitive robots which can interact with humans is more possible than ever thanks to recent European advances in functional imaging in primates which establishes the bridge between the human work and the knowledge from invasive techniques, accumulated in the last 40 years. The fact that these advances happened in European labs opens an extraordinary opportunity for the EU to lead the world in linking neuroscience and IT, in particular, computer science and robotics.

Despite the difficulties it is facing, European neuroscience, or at least its most performing laboratories, has been very responsive, not just because FET (Future of Emerging Technologies, an IST programme nursery which has launched a proactive initiative called Beyond Robotics) has provided them with much needed support. The classic example of why it would be good for neuroscientists to collaborate with engineers is the problem of segmentation in vision (the process of partitioning a digital image into multiple regions). What is present on the retina is a spatio-temporal distribution of light, not the image of an object. The brain is so complex that even models are insufficient to understand this complex reality. In order to strengthen neuroscience, it is crucial to understand the complexity of the brain.

What makes the brain so special?

Each of the 10 billion or so neurons in the human brain is connected to 1,000 or more other neurons. Brain function arises from the concerted action of anatomically organised groups of neurons. These determine the supra-neuronal levels of integration typical to the brain: the local network level (e.g. cortical columns), the functional map level (e.g. primary visual cortex) and the system level (e.g. the visual system). In addition, there are the neuronal and sub-neuronal levels. To understand the brain we need to integrate information across these levels by modelling (see chapter on the constructed brain). While we have powerful techniques to address the neuronal level it's the supraneuronal levels which are most relevant to neuro-IT, because they embody the computational principles we want



to endow artefacts with. We propose that the combination of dense multiple recording with functional imaging can do this but the techniques have to be developed further at maximum strength. Advances in this area will have the capacity to improve quality of life by offering breakthroughs in medical science.

Objectives

1. Strengthen the knowledge base of European neuroscience, to enhance the cooperation between information technology and neuroscience
2. To be able to record simultaneously and chronically from 1000 neurons in 5 or more brain structures and to able to relate these measurements to the different non-invasive, high resolution brain imaging modalities: fMRI, EEG, MEG, PET.
3. To be able to use these measurements to understand the operations performed by the different brain structures, not just simple input-output relationships but representations emerging in complex networks.
4. To obtain these measurements under a wide range of conditions including in realistic sensori-motor and sophisticated cognitive tasks.
5. To combine these measurements with physical (electric stimulation, cooling) or chemical (pharmacological local injection) manipulation of neural activity or transmitter systems

Examples of realisations

1. Understand how primates and humans head through the environment, grasp, catch or manipulate objects.
2. Understand how primates and humans classify objects and actions in a scene and perform other cognitive tasks.
3. Understand how learning and training change the representations in the brain and enhance performance.
4. Provide the underpinning of systematic use of brain imaging for clinical and pharmaceutical investigations.
5. Decrease the need for invasive experiments

Current state of technology

Positron emission tomography (PET) is a brain imaging technology that uses radioactive tracers to visualise brain function. The amount of tracer injected is minimal but subjects can only participate in a single session per year. Depending on the tracer used the PET scanning will measure either regional cerebral blood flow (using radioactive water) or label receptors or other molecules related to synaptic transmission or cell to cell communication. During the 1985-1995 period this was the main avenue for functional study of the human brain but it has deficiencies and has now been taken over by functional Magnetic resonance imaging (fMRI) which allows comparison between different activation regions in a single subject, and allows repeated testing of that subject. Studies of language are particularly effective with this system. PET remains unsurpassed for its other main application – studies of neuronal transmission, but the technique depends heavily on the development of tracers and on radioactive chemistry laboratories to produce them locally.

Functional magnetic resonance imaging (fMRI)

fMRI is based on the BOLD (brain oxygen level dependent) which is dependent on three hemodynamic variables: blood flow, blood volume and oxygen extraction. fMRI measures neuronal activity indirectly and needs to compare MR signals in different conditions. In the simplest design, MR activity in two epochs is compared. By adding a low-level control we can disentangle small differences in activation of in deactivation between the main conditions.

This subtraction design has been criticised in the sense that it is difficult to isolate a single cognitive factor, since the factor interacts with other cognitive factors, already present in the lower level condition. This is far less a problem in simpler sensory experiments in which the subtraction design has proved very useful. To isolate the effect of cognitive factors in more complex experiments, other designs such as factorial and parametric designs have been used. Factorial designs have the additional advantage that interactions between factors can be studied.

Although fMRI provides signal strong enough to study single subjects, one needs to record from several subjects to ensure the generality of the finding. On one extreme are group studies in which all subjects are averaged together which will ensure that a finding is representative. But to ensure general conclusions, one needs to use the random effect model.

In between we find the ROI analysis. The differences between magnitudes of the MR activity averaged over the ROI in different conditions can be tested statistically across subjects.

The time course of the BOLD effect is slow, yet fast enough to be convolved with brief trials or with different sub-periods of long trials, in what is referred to as event-related fMRI. In the brief trial version, activity is measured only when the subject is engaged in the trial. This technique allows the comparison between different types of trials, for example; correct and incorrect trials, trials with targets and without distracters, trials with stimuli in different parts of the visual field. The cost of these more specific activation patterns is the loss of statistical power. This lack of power can be offset by increasing the number of subjects.

An increasingly used application of event-related fMRI is the repetition paradigm. In this paradigm, trials with two identical or different stimuli are compared with trials in which it is unknown whether or not the brain treats the two stimuli as different. The MR activity will be lower for identical stimuli than different ones. Depending on whether the MR activity is low or high in the trials with unknown stimuli, the experimenter can separate processes that operate at different instants of the trial, such as visual processing, maintenance and response selection in working memory trials.

fMRI only indicates that signals related to average neural activity differ between conditions. It is badly in need of validation and even more so the adaptation paradigm. In humans fMRI can be compared to neuropsychological data – if a region, active in a task is critical, its destruction should impair the task. In practice this rationale is difficult to apply since lesions generally are vascular in origin and affect large, stereotyped regions of cortex, for example, the territory of the middle cerebral artery. Therefore, fMRI has relied very heavily on comparison with single cell data obtained in the awake monkey – the only adequate animal model for higher-order brain functions. But there are severe comparison problems that can only be solved by resorting to a new technique, fMRI in the awake monkey.

It is worth pointing out that Europe has a leading position in this new technique, monkey fMRI, which has not at all been exploited at the European Community level.

Functional connectivity

Activation studies performed with fMRI are still inadequate and what is really needed is a functional description of the cerebral network active in a task, i.e. not just a description of the nodes but also of the links between them. Depending on the task, the anatomical connections will be used differently and functional connectivity refers to these adjustable strengths of existing connections. In order to investigate the functional connectivity between active brain regions, structural equation modeling (SEM) technique is commonly considered for computing the connection weights in a predefined network. Alternatives to SEM have also been introduced. Recently, Dynamic Causal Modelling (DCM) and SPM2 beta release have been introduced.

Tracing anatomical connections with MRI

In vivo tract tracing refers to local injections in to a brain region of a tracer that can be visualised in the MR. So far only one study has been performed in the monkey using Magnesium and investigating connections of basal ganglia. The interpretation of such studies is compounded by the influence of magnesium on the neuronal function.

An alternative for in vivo tract tracing that can be used in humans as well as animal models, is Diffusion Tensor Imaging (DTI). DTI exploits the asymmetry of motion of water molecules in nerve axons, but is in its infancy. Major problems are absence of signals within the cortex and disentangling the multiple crossing axons.

Increasing the temporal resolution: EEG and MEG

The main shortcoming of fMRI is its relatively low temporal resolution, even in event-related mode, especially in comparison with the time course of single cell activity. Since a few years it has been repeatedly suggested that this can be remedied by integrating fMRI with EEG or MEG, which suffer from the opposite limitation. Although several attempts have been made, this problem is not completely solved in humans and neither has this fusion of imaging techniques been tested on animals.

Other imaging technologies with limited use

Other imaging technologies with limited use include Optical recording, which has restricted use on old world monkeys, 2-deoxyglucose technique, which has excellent sensitivity and spatial resolution but can only study one or two conditions and is a very invasive technique.

Multiple single cell recordings

Obviously more information can be obtained from recording multiple single neurons rather than a single neuron. Initial techniques allowed recording of small numbers of neurons, typically 2-5. The aim was to study synaptic connectivity or to increase the number of neurons tested. More recently attempts have been made to record from large numbers of neurons, as initially done in the rat by Nicolelis. The transfer of this type of experiments to the monkey has been difficult but has now been achieved. Arrays of 100 electrodes have been used even in different parts of cortex. One drawback of multiple recordings is that all neurons are tested with a uniform set of stimuli or conditions and stimuli cannot be tailored to the requirements of each neuron. The technique, however, opens much wider perspectives as many problems can be addressed, for example, functional architecture, synchronisation of signals between areas and control of a robot arm by the brain signals obtained.

Manipulation of brain structures

Lesion studies in which part of the brain is permanently damaged, either by surgical excision or by local injection of neurotoxic substances, are usually combined with behavioural testing. Note that lesions are more specific than surgical excision, as fibres of passage are spared. This was an important step forward to disentangle the role of hippocampus and overlying perirhinal cortex in delayed match to sample tasks.

Pharmacological agents can also be injected locally to manipulate the local neuronal activity. So-called inactivation studies rely on transitory silencing of neurons in a given region, typically with drug injections such as lidocaine (local anaesthetic). This has been combined with behavioural measures or single cell recordings. The problem is to inactivate large enough regions to obtain reliable effects. An alternative is local cooling, which generally can affect large enough regions and can be more rapidly reversed, but is difficult to restrict to a given anatomical region.

Finally it is worth mentioning that in humans systemic injection of pharmacological agents is used in pharmacological challenge studies in which task/stimulus and drug interactions are imaged. Extension of these studies to animal models should enhance considerably their use for the clinical and pharmacological purposes.

Visual and sensory systems

Monkey visual system

It is now more than ten years since Felleman and Van Essen (1991) compiled the visual cortical areas in the monkey. Beyond primary visual cortex, the monkey cortex contains about 30 different extrastriate (areas of the cortex next to the visual cortex) visual cortical areas. Each of these areas is on average connected to 10 other regions. The primate visual system is extremely complex and adapts its configuration to the visual task at hand. In comparison rodents have only a few extrastriate areas so the exploration of the rat visual system has no interest for understanding the human visual system.

The nice maps of monkey extrastriate cortex should not hide the fact that our knowledge of the best known sensory system is still very fragmentary. In a number of instances the boundaries of a number of areas are not firmly established. Even those regions for which the boundaries are established have not all been explored in detail: only one study has been devoted to area DP to give an example. Often these studies are performed by young PhD students and the supervisor will choose a well-known area in which the stimuli will work. Hence, most of the progress is achieved by young independent researchers, such as assistant professors, who can afford to take risks because they have proven themselves as PhD and post doc. In Europe the ultra-conservative policy for academic recruitments, related to job security, hampers the recruitment of exactly this sort of innovative researcher.

Many of the main functions of the primate visual system, the knowledge of which is needed by those building artificial systems, are still little explored.

Human visual system

Functional imaging has shown that in general terms the visual systems of all primates are similar. The visual system in both species is divided in dorsal and ventral streams and these streams process to some degree different attributes for different behavioural purposes. As imaging in both species progresses differences start to appear. V3A has similar retinotopic organization in both species, yet is both stereo and motion sensitive in humans but only stereo sensitive in monkeys. The IPS of humans processes motion information, and in particular extracts 3D from motion, much more than its monkey counterpart. For years there have been heated discussions about the colour-processing region, largely based on an absence of relevant data. Now that both brains can be imaged exactly in parallel, these problems can be rigorously addressed.

Other sensory systems

There is a general lack, also in Europe, of primates studies on other senses. This is particularly true for the tactile sense. Here also a number of cortical areas have been mapped and it has been proposed that the tactile system, also includes a dorsal and ventral stream reaching the parietal cortex and the insula respectively. Even more so we have little clues about the role of these different regions.

Motor systems

The frontal lobe of primates is formed by two main sectors: a rostral one (prefrontal cortex) that has essentially cognitive functions and a caudal one that is related to the control of movements. Five areas lie on the lateral cortical surface, two on its mesial surface. The subdivision of the motor cortex into 7 areas was originally described in monkeys. A similar subdivision starts to become clear also in humans although some aspects of it as not yet clear such as the border between the dorsal and ventral motor areas and within the ventral premotor cortex.

Why there are so many motor areas? Such a multiplicity is surprising, especially if one accepts the classical view that motor areas had as their only functional role the control of body part movements. But recent neurophysiological data has shown that motor areas play a broader role in behaviour. First of all, motor areas are involved in a series of sensory-motor transformations. Among them, particularly complex are those that transform visual information on objects and object location into the appropriate goal-directed actions. Second, motor areas are endowed with a mechanism that matches observed actions on the internal motor representations of those actions (mirror mechanism). This mechanism may contribute not only to action recognition and preparation but also to learning of actions. Third, motor areas are involved in decisional processes that lead to action initiation. Finally, some premotor areas are involved in the control of sequences of movements.

This subdivision of motor areas is in accord with their connections with other motor areas (”intrinsic connections”). Another anatomical finding that strongly supports the validity of this subdivision is the organization of cortico-spinal projections. Certain motor areas send direct projections to the spinal cord, while others do not. From these anatomical data, it appears inescapable to conclude that the two sets of areas play different roles in motor control. It appears logical to posit that these areas have a control function.

Cognitive systems

It is well established that prefrontal neurons display delay activity in the interval between two stimuli or between a stimulus and a response in delayed match to sample or response tasks. In addition to delay activity, the task dependency of prefrontal activity has been recently documented physiologically as well as its role in categorization. While the lateral aspect of prefrontal cortex is heavily engaged in cognitive processing, the medial and basal prefrontal cortex is engaged in motivational and reward processing. Selectivity of medial prefrontal neurons for type or value of reward has been demonstrated. In addition to prefrontal cortex, parietal cortex has been shown to contribute to cognitive functions.

Problem Areas

Long-term recordings with multiple electrodes

The two main problems are the damage to the cortex and the recording of the same neurons over long time. It has become amply clear that the monkey (and perhaps human) cortex is much more vulnerable than say rodent cerebral cortex. Thus methods to evaluate damage and to restrict damage are urgently needed. The stability of the recordings is probably the most important problem since it would extend the use of the technique tremendously, e.g. many training experiments would become possible.

Scanning moving subjects

The present day scanning situation is dramatically restricted. The head of the subject has to be precisely fixed, the subjects lie in a confined space. Auditory stimulation is difficult because of noise of the scanner, access to the body is restricted, visual stimulation is generally restricted to a part of the visual field.

Going for lower field strength in which wide bore magnets can be used and in which some subject movement is tolerable might be a better way forward. The development of new sequences providing new type of information about brain function remains important, as is the development of new coils.

MEG for monkeys

All brain imaging modalities suffer from the same limitation: lack of validation in animal models. Do they really measure what they are claimed to measure? This can only be tested if other sources of information (a ground truth) is available, as it is the case for monkeys in which many invasive experiments have been performed. Thus the new brain imaging techniques and their fusion should be tested in monkeys. EEG and now fMRI are readily performed in the monkey, but MEG would require adaptation of the present equipment, perhaps equipment designed for children could be used.

Few mathematical tools
Just as mathematics were developed for physics then for economics, we need mathematics for biology and in particular for neuroscience. Of course statistics is used, but what we need are new mathematical tools to deal with the multiple electrode signals and/or the MRI signals.

Education of the public
In its majority the public is supportive of medical research even that involving animals, particularly when it has clinical applications. We need to educate the public about the distance between basic and clinical science: that a clinical success builds on years of basic research. This is even more true for neuroscience, because of the complexity of brain function.

Young Investigators
The dramatic trend of loosing brilliant post-doc' to the US must be reversed. The main reason is often the lack of support (including laboratory space) for independent research of these young investigators.

Future Research

Improve recording and local manipulation techniques

The electrode arrays can be further improved to record from more sites, increase the likelihood of recording single neurons over long periods of time, and without damaging cortex.

Study the possibility to inject electric signals back into the electrodes for stimulation, perturbation of brain regions or other use. Methods to assess damage and to visualize in vivo electrode location are important. To miniaturize the connections and introduce wireless methods is imperative so that the animal could move its head.

To improve ways of delivery of local chemicals to influence neuronal activity (and control the size of the effect), as well as to increase the range of such chemicals is useful.

Improve and diversify brain imaging techniques

To improve the Signal to Noise (S/N) ratio and consequently spatial resolution either by increasing field strength, better coil design or MR sequences, or by improving on contrast agents are important topics.

To make the contrast agents available and acceptable for human use, even for restricted clinical applications would be valuable.

Meanwhile, it will be critical for the better interpretation of fMRI signals to get a better understanding of the vascular phenomena and neural activity phenomena underlying the different MR signals.

While it will take some time before we can scan a human subject who walks in a field, we should try to lift many of the restrictions on the motor and sensory side imposed on the subjects during scanning.

All brain imaging techniques used in humans and even in clinical settings, have yet to be properly validated. For higher cognitive functions, which are the essence of human functional imaging, so validation in the monkey is essential.

Monkey fMRI, especially in the awake animal, also opens an almost unlimited avenue of pharmacological research. Pharmacological companies suffer from a large gap between assessment of new potential drugs in small animals and in humans. Many drugs fail in that interval which could be bridged by pharmacological monkey fMRI studies.

New mathematics for Neuroscience

We badly need more incisive techniques to treat multi single cell recordings. We should go beyond correlation techniques, which are now the main tool used. Techniques to provide general interpretation and integration schemes, such as new coordinate systems, brain atlases, and warping algorithms to compare brains, are important. Development of new signal processing tools to extract relevant signals from fMRI measurements are key. Finally, one needs to develop mathematical tools to relate the multiple single cell recording, or their local field potentials equivalent, to global signals such as fMRI or EEG signals that will have been recorded simultaneously. Again we should go beyond correlation.

Visual system

In vision the main issues of segmentation, extraction of 3D shape and motion, of building shape, material, action and scene representations for recognition, categorisation and visuomotor control as well as cross modal integration should be addressed. While we can link at a coarse level the different visual cortical areas with these different functions (dorsal and ventral streams), the detailed functions of the different (over 30) areas are largely unknown.

In the same vein, coding of a number of image features has been documented, but we largely miss the dynamics of the visual system, which adapts itself to the task at hand. While top-down modulations of all sorts are very important, their study cannot replace the investigation of the visual functions as such, which are largely neglected.

Motor system

In the motor system the multiplicity of cortical areas also calls for further investigation

1. The role of the fronto-dependent motor areas (F6 and F7) is only hypothetical. Understanding this control may be of enormous advantage for constructing robots or other artefacts that, on one side, code external visual stimuli in a format ready for action, on the other emit responses only when particular contingencies are met.

2. The transformation of the intrinsic properties of objects into the selection of appropriate hand action, that takes place in the parieto-premotor circuit AIP-F5 needs further study. For example, how does the AIP-F5 premotor circuit know the quality of objects? This knowledge will be of enormous value for the construction of artefacts able to interact with objects in an intelligent way.

3. The discovery of mirror neurons provided the basis for understanding the mechanisms of imitation. Yet, there are virtually no data on how mirror neurons are re-organised when an individual learns a new action. There are enormous economical possibilities for construction of robot that could learn by imitation.

4. The link between the motor areas coding action and the primary motor cortex M1 (F1) coding mostly movements are little understood. Understanding how motor knowledge is formed could be a fundamental step in the construction of really intelligent artefacts.

As mentioned above for the visual system, many of the fundamental problems outlined can be solved by chronic recordings of action potentials and field potentials in the behaving primates using multiple electrodes and by exploiting the now available possibility to employ fMRI techniques in monkeys. This latter technical development represents a fundamental step that allows one to link classical neurophysiological monkey studies with human brain imaging studies (PET, fMRI, MEG, quantitative EEG). The combined used of these techniques and combined use of monkey and human experiments will solve many problems not discussed such as the understanding of why an action has been performed (its distant purpose).

Other systems

Given its importance in cognitive functions prefrontal cortex should be explored more vigorously on this side of the Atlantic. Also we should ensure a minimum coverage of other important regions of the primate brain such as the tactile cortex, medial temporal cortex and deep brain structures and auditory regions.

Theories of cognition and design of cognitive tasks

However, powerful the tools at the disposable of the neuroscientist, the quality of the experiment depends in the end on the paradigms used. As well as achieving advances in such paradigms, we should aim for them to contain less abstraction compared to real life situations (for example, monkeys using tools, taking elevators (in virtual reality)).

Theory of brain function at neuronal, network, functional region and system level

Modelling is generally accepted as necessary in order to understand the huge complexity of the brain. However in addition to modelling theories about the brain and its functions are important. Beyond that we have to understand the data. We will have to run lots of simulations on the model to understand what each neuron, circuit and functional region is contributing to the behaviour we observe.

First impetus

A concrete plan for the immediate future could be concentrated on four directions.

1. Strengthen the research basis. Most performing European laboratories easily match the competition of the American and Japanese labs but we need more of them, and to encourage young researchers, who have often been lured to the US, to set up their own independent group. We may have to devise special grants or awards, and include provision for laboratory space.

These proposals should primarily use primates, unless a specific question can be addressed in lower animals.

2. To foster by all possible ways the introduction and widespread use of multi-electrode recordings. Since many of the problems are technical, we should favour proposals linking neurophysiological teams with SMEs e.g. Thomas recording in Germany or with engineering groups. Here the EU could play the role of catalyst in bringing these groups together.

Also proposals that favour the understanding of the blood supply and other physiological requirements of the brain, should be welcome. As multi-electrodes are introduced we should support the development of software to record, display and analyse this wealth of data.

3. To foster the integration of different non-invasive imaging techniques in the primate, notably fMRI, EEG and MEG. In non-human primates the verification is easy: compare the generators postulated from on invasive imaging with the direct recordings in the corresponding brain area. This aspect is extremely attractive from the EU perspective, as the validation of this imaging integration is the condition for widespread clinical use and can also lead to industrial products. After all two of the four major companies producing brain scanners are European. Progress in these non-invasive techniques allows reducing the number of animals in research.

The particular problem will be to find money for European groups performing monkey fMRI to acquire the expensive equipment required for MEG. The creation of EU sponsored centres of excellence linked with companies producing the equipment and functions as a test site, is a possible mechanism. Again we are envisioning projects linking academia with industry.

4. Increase the awareness in mathematical circles for the need of neuro-mathematics. The traditional way would be to call workshops and symposia.

Ethical problems

Although funding for the neuroscience experiments is justified here from the point of view of information technology, it should be clear that the rationale for the experiments themselves is the understanding of the brain critical for human health.

The need for primate research is not always well understood by the general public so it is crucial to inform the public and the authorities of the following three points;

1. The need and particularities (slow, painstaking, tortuous nature) of basic research.

2. The distance between basic biological research and clinically relevant medical research is long in general but especially so in brain research.

3. The need for using the adequate animal model. At this point non-invasive techniques of brain imaging not only lack the resolution in space and time compared to single cell studies, but also have not been validated, hence the need of using animal models. On the other hand when invasive techniques are being used the choice of animal model depends on the function investigated. For higher order functions and most cognitive function primates are the only option.

In counterpart it should be clear that the neuroscientists are using all possible ways (including the development of imaging) to reduce the need of invasive investigation and animal models in general, but also that they take great care of the physical and psychological well being of the subjects in their care.

The way forward

Ambitious challenges

It is clear that the challenges laid out in this roadmap are very ambitious and will occupy the minds of researchers for many years to come. Indeed, some of them can only be started in the future, when our knowledge and skills have developed sufficiently.

But the ambition of these challenges and the scope of the research being attempted take Europe to the very forefront of this important work, which will not only have a dramatic impact on the way we live our lives, but on life itself. We are talking here of developing conscious, learning, thinking, adapting machines; we are close to artificially creating the essence of life itself.

The potential is enormous. The more we understand brain function and construction, for example, the better able we are to treat brain disorders and illness and prolong life. The development of robotics, cyborgs and humanoids will lead to any number of commercial applications and subsequent economic benefits will surely follow.

By creating Neuro-IT.net and forming a completely new area of research at the interface between neurosciences and information technology, we have given Europe the lead globally in a field that will have direct benefits for IT generally, as well as helping to discover completely new research domains.

Part of this is our commitment to seeking partnerships with industry, the public sector and other consortia and institutions and, by creating an effective network, we are also fostering closer cooperation between existing research groups.

All this is vital as it is quite clear that many aspects of the work set out in the Grand Challenges have strong interrelations. Conscious Machines, for example, will need to utilise much of the research being carried out in the Constructed Brain and Brainprobe Projects.

We have also recognised the need for collaborations to extend beyond small research groups with, for example, the Acting Challenge calling for standardisation of peripherals. Meanwhile, those involved in the Constructed Brain Project have recognised that the models created by individual research groups cannot possibly capture the complexity of the brain and so is calling for a framework to be created to connect such models.

All this cooperation clearly identifies the need for the NeuroIT community – as well as for more permanent funding initiatives to create communities and institutions to create the standards and other collaborations in a more systematic way than at present.