Roadmap of Neuro-IT Development

Edited by Alois Knoll & Marc de Kamps

Version 2.0: July 12, 2006

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

Contents

Chapter 1

Introduction

1.1 What is Neuro-IT.net ?

Neuro-it.net is a Thematic Network (TN) dedicated to NeuroIT. [1] NeuroIT.net has various tasks: the organization of reviews and workshops where members from various disciplines can meet and get to know each other, to establish contacts between academic institutes and small and medium enterprises (SMEs), to stimulate schooling and training in the field of NeuroIT, and last but not least, to draft a Roadmap. At the start, 54 institutes and SMEs comprised NeuroIT.net. At the time of writing over 200 researchers from more than 100 institutes and SMEs, divided over 16 nations make up NeuroIT.net. NeuroIT.net has its own website: `http://www.neuro-it.net` and a mailing list, which can be found via the website.

1.2 What is a Roadmap ?

A Roadmap is a document, which describes the current state-of-the-art of the field, as well as a vision of which research will interesting, challenging and beneficial for the field over a relatively long period of time (a decade or more). It serves as a reference for funding agencies, but also for scientists. It has played a considerable role in earlier formulating calls during the sixth framework programme, in particular the BIOI3 call was influenced considerably by the contents of the Roadmap at that time. At the moment of writing, the seventh framework programme (FP7) is in the planning stage. In particular the first two years will be defined over the next months. It is important for the Neuro-It community to present input to these programmes, and this Roadmap is intended to provide such input.

Currently, one other Roadmap already exists: the Roadmap for Nanoelectronics (NE), which can be found on the web at:
`http://www.cordis.lu/ist/fetnidqf.htm`

---

[1] A Thematic Network is one of the Instruments of the employed by the 5th Framework Programme(FP) of the European Union. Framework Programmes are the EU's main instrument for research funding in Europe. FP6 has been in operation since January 1, 2003 (European Commission, 1998), which coincides with the start of NeuroIT.net, although officially this Neuro-IT.net an FP5 initiative.

1.3 Introduction to the purpose of this roadmap

The aim of Neuro-IT.net, the EU Neuro-IT Network of Excellence, is to build a critical mass of new interdisciplinary research excellence at the interface between NS (Neurosciences) and IT (Information Technologies) within the European Union and its Associated States. The term Neuro-IT was coined to express clearly that the disciplines merged under the umbrella of Neuro-IT.net form a new scientific working area, which is different from what is traditionally called Neuro-Informatics (NI).

The objective is to complement and move beyond the well established NI (NeuroInformatics) or AI (Artificial Intelligence) domains by fostering research that would benefit both the NS and IT communities by helping solve the fundamental problems linked to the emergence and the modelling of cognitive and awareness processes as well issues related to physical growth, phylogenesis and ontogenesis. 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

1. making known the potential of the basic research conducted within the EU funded initiatives and
2. spearheading the emergence of completely new visionary long term research objectives that could fully exploit the potential of collaboration between Neurosciences and Information Technology.

In this context, the role of Neuro-IT.net is not to support incremental research, no matter how excellent, but to help to discover new unexplored research domains that could lead to breakthrough in Neuro-IT in the long term. A central guiding question in this respect is: ’What can neuroscience do for IT’ ?

To make this highly abstract goal concrete, the members of Neuro-IT.net have agreed to develop a roadmap, which is in the form of “Grand Challenges”. Taken together, these challenges cover a broad scope of what we image to result from Neuro-IT research. However, each of these challenges is in itself a demanding research programme, laid out to lead to tangible results, both in terms of very basic scientific research and in the development of technology leading to prototypes, which can show the potential for new products.

The challenges are summarized in the executive summaries, below.

The Roadmap for NeuroIT will be the second Roadmap that will be created, the Nano-Electronics (NE) Roadmap was the first. It will be very different from the NE Roadmap, which was described from the perspective of a mature and powerful industry.

1.4 What is the status of this document ?

This document, version 2.0, is a draft. It contains a new Challenge (Bio-inspired hardware) and most of the other Challenges have been thoroughly updated. **At this moment, we will are still open for input, new ideas or comments.** Due to the fact that FP7 is still in flux, we will keep the map open, and forward your comments directly to the FET project officers. Version 2.1 will stabilize at the time where the FP7 programme itself has stabilized.

1.5 Where can this document be found ?

The latest version of this document will be maintained on NeuroIT.net's web site: `http://www.neuro-it.net`.

1.6 Executive summaries

1.6.1 Executive summary of the Brainship Project

Recent progress in fundamental neurophysiological research suggests that a popular subject of science fiction movies may soon be technically possible: direct interfacing of the human brain with computers, either in embodied external devices or incorporated into the human body. Development of better electrodes and of fast signal processing techniques have allowed chronic implantation of large arrays of recording electrodes in rodents and monkeys. The major breakthrough was the demonstration of a high level of plasticity in the mammalian brain, allowing it to adapt its signals to communication over a limited number of channels. Nevertheless all present demonstrations are one-directional, usually involving the use of signals from motor cortical areas to control virtual devices. For real-life applications, like the control of paralyzed limbs or complex prosthetic devices, bi-directional interfacing will be necessary so that the brain can use its sophisticated feedback control strategies.

Bi-directional brain computer interfacing (BBCI) holds therefore great promise in the treatment of neurological and trauma patients. More controversial applications of BBIC lie in the direct control of remote robotic devices and information systems. Before this highly invasive technique can be applied to humans further development is needed on multiple fronts. Particular areas of concern are the lack of direct sensory input, necessary for feedback motor control in locomotion, poor understanding of neural coding outside of primary motor regions and longevity of implanted electrodes. These goals problems will be covered under the primary goal of the Brain Interface project: the development of an awake animal model where the brain interacts with the environment only through BBCI techniques, in other words both sensory input and motor activity will be channeled through computer interfaces. Additionally we will promote discussion and development of guidelines for the ethical use of BBCI in humans.

To augment human interaction with its environment by enabling direct interfacing to sophisticated perception, robotic, prosthetic and information systems. Present technology requires invasive methods which will be enhanced to create bidirectional brain interfaces for control of high-dimensional systems. Both neurophysiological and IT technologies will need to be greatly advanced to allow interfacing at a natural cognitive level or to embed the brain at an appropriate level in an artificial sensori-motor loop.

1.6.2 Executive Summary of the Bio-inspired hardware project

Technological progress has rapidly increased the number of transistors that can be included on a single chip. Nevertheless, most current computing architectures are based on clocked, digital processing units with a low degree of parallelism. This scheme has established limits, and designers are struggling to find alternative design and computing paradigms to overcome these limits (e.g. see INTEL's recent multi-core CPUs, or IBM's latest CELL processor). One attractive solution is to use silicon technology (e.g. analog VLSI, FPGAs) to implement bio-inspired processing schemes, with massively parallel, scalable architectures. In the short term, it is unlikely that such architectures will outperform conventional solutions, except perhaps on the sensory periphery. But they could be the first step towards new high performance computing strategies and provide new insights into the working of the brain.

Biological nervous systems are composed of large ensembles of noisy, slow, and low-power neural structures which can be composed into fast, robust, and fault-tolerant systems. The performance of these systems on specific tasks improves during development, allowing them to adapt to the peculiarities of specific tasks, bodies and input/output systems. Designers would like to incorporate these mechanisms in their circuits. But the adaptation of bio-inspired processing schemes for implementation on silicon is non-trivial. The computational resources (multipliers, embedded processors, embedded memory blocks, etc) provided by modern FPGAs were designed for signal processing, image processing, etc. and may not provide the kind of primitives required. New analog and digital design principles are needed - starting with the simulation of neural structures on FPGAs and the development of adaptive analog and mixed signal circuits.

Brain-like computing on silicon will be useful in a broad range of applications, from real time control devices for robots to implantable devices such as artificial cochleas and artificial retinas, to large scale simulation (e.g. of the brain). A new generation of systems will emulate biological evolution, exploiting information about task performance and resource use to generate a new generation of more effective machines that emulate specific biological systems. The potential impact on industry and on human society is huge. It is this vision which inspires our GrandChallenge.

1.6.3 Executive Summary of the Factor-10 Project

Both the emerging fields of epigentic robotics and "smart" materials science offer a wealth of innovative research opportunities and promise a large spectrum of new products to become feasible in the mid to very long term. However, a completely new discipline may develop by combining key research in these fields to work towards new types of artefacts. We envision three types of such artefacts to emerge from this combination: (i) artefacts evolving their cognition and motor control autonomously based on multimodal/multisensory feedback in a predefined and fixed body, (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 and (iii) artefacts that **co-evolve** their brains and their body in permanent interaction with the enviroment over an extended period of their lifetime (embodied artificial ontogenesis).

While the implementation of the first type of artefacts largely depends on progress in control architectures of cognitive systems, the latter two will also draw heavily on methods and technology developed in materials science for self-assembling materials and structures. Even more so, type (iii) above may be 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, the third type of artefact needs a completely fresh approach in that it can only be realised as an *autopoietic* machine built from cells, i.e. "from the inside out". To make these extreme challenges an easy-to-communicate research project that everybody can relate to, we propose to define a long-term venture called "Factor-10", which aims at a *fully functional physical artefact* (i.e. not a computer simulation), which, during an extended but limited period of time (e.g. 10 months) autonomously grows

- the volume of its body by at least a factor of ten, thereby differentiating out "organs" and "effectors" as well as
- its cognitive abilities (its "IQ" and its repertoire of sensorimotor behaviours), also by at least a factor of ten.

Issues central to the development of living creatures that would have to be followed to a higher or lesser degree for such an artefact, i.e. synchronous evolution of morphology and mind, have hardly been formulated, let alone been tackled. Fortunately, due to the need for qualititative breakthroughs at diverse research frontiers, there would be a window of opportunity for Europeans to compete with Japanese research at the next stage of development – the Japanese advantage in humanoids (type (i) above) research will hardly be caught up to. Looking at the preconditions for embarking on this research journey, we note that a there is already a sizeable body of research in the diverse necessary disciplines represented in Europe (see the non-exhaustive list in the appendix), however with fragmentation across disciplines and countries.

Although the goals of Factor-10 are reaching far out into the future and well beyond what is currently the set of obejctives in FP6, there are a number of items that can be related to the "priroty thematic areas of research in FP6", i.e. research will address *autonomous self adaptation* [1.1.2.ii] of physical systems (artefacts) capable of *responding intelligently to speech, gesture or other senses* [1.1.2.iv]. To obtain such artefacts, dedicated combined research both in the areas of cognitive sciences and in the areas of *self assembling materials and structures* [1.1.3.i] is both necessary and highly innovative. Moreover, some applications (e.g. intelligent prosthetics) resulting from work undertaken in this NoE require substantial advances in the integration of biological and non-biological entities. Finally, the growing artefacts can be seen as a completely new interpretation of "smart" materials with tailor-made functionalities and for building up macro-structures [1.1.3.ii].

Apart from the scientific objective of developing the basic technologies and actually designing as well as building prototypes of type (iii) artefacts, it is also the purpose of the project to establish a commonly accepted paradigm for designing these artefacts. The goals of Factor-10 are indeed very demanding. Up to now, they have hardly been formulated as a common integrating challenge because of the deterring technological impediments in every single area involved. We believe, however, that in view of the progress achieved in many of the disciplines, particularly 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.

1.6.4 Executive summary of the Acting in the physical world Project

The objective of the ‘Successful thinking and acting in the physical world’ challenge is to build complete systems which make optimum use of distributed intelligence embedded in the periphery (sensors, actuators, body morphology and materials) and at a system integration level. Research on this grand challenge will emphasize

- 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

Distributed, embedded intelligence should enable the artifacts to master tasks known to be performed by natural (neural) systems but currently elusive to technological reproduction. It should have a significant effect on central neural computations taking them to until now unattained levels of efficiency. Despite neural processing remaining an integral part, the focus of this grand challenge is on making the periphery smarter and integrating it better with central computations, so that the whole system gets more powerful and efficient. In addition, knowledge about the tasks to be performed and the world that they are to be performed in should be integrated at every stage. Efficient ways to distribute the storage of this knowledge, i.e. ‘environment models’, over the different subsystems should be developed. Ultimately, it should be possible for designers to have confidence that each part of such a system exploits all reasonable available prior knowledge. The same should be true for the system integration level. This calls for advanced methodological achievements in gathering the relevant knowledge. Optimization processes in nature operates on large time-scales and vast numbers of prototypes for testing. In order to apply such optimization to every part, every integration level and every task, shortcuts need to be found which narrow the search space so that it can be managed within the scope of a engineering design process.

In following the outlined approach and addressing the involved issues, research on this grand challenge will produce novel, smart peripheral devices for NeuroIT systems and thereby promote the pervasive use of intelligent robotic systems. While individual projects will probably have to include many case studies, work on the grand challenge as a whole should establish general rules, how these objectives can be achieved for any tractable problem. Obtaining a set of design rules will enable tailoring solutions to IT problems without further need to study specific natural solutions. Consequently, the design rules may be applied even to problems for which no solution is known to exist in nature.

Research results should lead to the creation of universal standards (e.g., ‘bus standards’) for smart NeuroIT peripherals, which would enable closer cooperation between research projects (parts developed in one project can be reused by another) and also facilitate the inclusion of novel technology into product design. A pool of smart, readily available periphery should not only provide the building blocks for powerful individual systems (robots) but also establish new capabilities for robot interaction and collaborative behaviors, either between self-contained individuals or parts forming ‘states’ or super-organisms.

1.6.5 Executive summary of the Conscious Machines project

There is considerable interest in Europe for the topic of ’Machine Consciousness’. This topic, which even ten years ago would have been dismissed as ’crackpot’ science, draws an increasing number of established researchers from well-respected universities. Why? What has changed? And why is the topic considered to be relevant to present day engineering problems by a large group of people?

In the first place, considerable progress was made during the last ten years in brain science. In particular the role of attention is now understood in much greater detail than before. An enormous amount of information enters our senses and although this information can processed in parallel to some extent, prior knowledge and experience are essential to process only that part of the data which is most likely to be behaviourally important. Moreover, similar processes seem to be at work in preparing actions, but also in planning and remembering. A picture of the brain as a sophisticated control structure emerges, with internal controllers which predict sensory input, which monitor the results of actions currently carried out and which predict the result of actions yet to be carried out. Although the picture is by no means complete yet, it becomes possible to define more precisely, in terms of neural substrate, what consciousness is and what it is not.

From an engineering point of view, many are unhappy with current state of IT applications. These applications must be programmed carefully and laborously, for well-defined situations, and their performance beyond their original specifications is usually poor. Moreover, there are many tasks which are considered to be easy for human beings, which we can not emulate in technology. Increasingly, people believe that we should look more closer to how the brain has dealt with these problems and increasingly they feel that consciousness plays an essential role in the flexible adaptive behaviour of human beings.

Another motivation is provided not by single applications, but by technical system which are comprised of a large number of components, but which are not under control of a single manufacturer anymore. Power grids which extend over various nations are one examples, technical applications, which have to co-operate with another technical applications (such as in communication networks) are another. Classical methods for certifying safety and corectness do not apply here anymore. The idea is to delegate responsibility for corectness and saety to the individual components that make up the system. In order to be able to that, these systems must be able to monitor their own performance, and to negotiate solutions to problems with other parts of the system to solve problems. This is line of research is heading towards a control structure not unlike that found in the human brain.

1.6.6 Executive summary of the Artificial Evolutionary Design project

Traditional Artificial Intelligence (AI) has failed to scale up to real-world applications. Similar problems face so-called new AI. Current AI techniques have failed to produce systems that match the routine performance of relatively simple animals. The Artificial Evolutionary Design Project is based on the premise that such systems requires novel techniques of design, inspired by biological learning and evolution. Such techniques, it is suggested, could build on existing work in Artificial Neural Networks, Evolutionary Computing and related disciplines, while enriching the research programs of these disciplines with new insights from modern research into the evolution and development of biological organisms.

The concept of biologically-inspired automated design is not new. **Artificial Neural Networks** have modelled animal and human learning as an alternative to explicit design; **Evolutionary Computing**, has drawn inspiration from Darwinian models of evolution. More recent work, has involved the automated design of physical, chemical and biological artefacts. In the majority of cases, however, industrial applications have been rare. The intrinsic weaknesses of current approaches include: the small size of artificial compared to natural systems, the use of undifferentiated sub-units, as opposed to the highly differentiated cells and organs found in nature, the pre-definition of architectures and morphologies, the failure to model development and related regulatory mechanisms, the lack of any realistic attempt to emulate macro- (as opposed to micro-) evolution, the failure to take advantage of natural self-organization through chemical and physical mechanisms, lack of attention to transfer of information between interacting subunits, and the tendency to model systems at a single level ignoring the hierarchical organization of real life biological systems.

In this setting, the Artificial Evolutionary Design Project aims to develop new techniques of artificial evolution and learning. To this end it will be necessary to develop mathematical models of the evolution and dynamics of Complex Adaptive Systems (CAS). These are likely to include models of macro-evolutionary change, techniques for the automatic design of highly evolvable structures and behaviors with the ability to rapidly adapt to a broad range of different environments, open-ended models in which large-scale adaptive change emerges as a sequence of adaptations to a changing environment, models of 'multi-level selection', allowing competing 'Darwinian individuals' (e.g. genes, organisms) to cooperate in higher level systems (e.g. organisms, demes), models which combining evolution and development, models of evolutionary and developmental change in which the 'search space' is constrained by 'grammatical rules' (as in current genetic programming), techniques for exploiting the intrinsic chemistry and physics of artefacts produced via Artificial Evolution based on (highly constrained), multi-level interactions among large numbers of heterogeneous agents.

Key project goals include the validation of these models through simulations of known biological systems and processes; investigation of their computational complexity, the design and implementation tools for the automated design of artificial CAS based on simplified versions of the models; the demonstration of the engineering effectiveness of these tools with respect to benchmark problems

Examples of the kind of products to which such models could lead include: 'generic' autonomic 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; flexible software for pattern recognition and categorization, which can be 'taught' to solve currently insoluble tasks, development environments for the semi-automated design of complex electronic circuitry and hardware; generic models for the development of domain and language-specific machine-translation tools, self-adapting systems for the protection of autonomous systems against threats, which have not been specifically identified at the time of design; and hybrid chemical-computerized development environments, for the artificial evolution of complex ’wetware’.

The attempt to design artificial systems with competencies comparable to their natural counterparts, is close to an attempt to artificially create life and as such is ethically ambiguous, even dangerous. It is likely however, that the endeavor will lead not only into an expansion of our current design capabilities but also to a better understanding of the intrinsic limitations of human design. The proposers suggest that such knowledge is not without ethical value.

1.6.7 Executive summary of the Constructed brain project

In the ’constructed brain’ it is argued that for a systematic development of cognitive engineering principles in NeuroIT a comprehensive framework is necessary that allows for the simulation of an entire brain. Initially realized in software, lateron its protocols will allow interfacing with hardware, thereby moving from a purely ’virtual brain’ to an ’embodied brain’, which may lead to artefacts that have a substantial degree of autonomy and adaptability. As such it may be considered a ’top-down’ approach for NeuroIT, whereas other challenges take a ’bottom-up’ approach.

The benefits of such a ’constructed brain’ are manifold. First of all, we have not been very succesful in the creation of autonomous, flexible, and adapable artefacts, whereas even simple biological creatures have amazing capabilities in this respect. A good understanding of neural processing is clearly essential to understand why biological creatures are so good at ’cognitive processing’. With such an understanding, it will be easier to judge, whether we can cast biological information processing principles into existing hardware and the systematic design of cognitive engineering principles for NeuroIT will be possible. It may even lead to formulations of ’awareness’ and ’consciousness’, which are defined in terms of large-scale neuronal systems and this may be instrumental in studying the transfer of these concepts from biological systems to artefacts.

Secondly, such a framework would lead to new and better ways to study the brain, for instance, because it would allow for ’experiments’ which would be difficult or unethical in human beings. This, in turn, should have a profound impact on the treatment of psychological disorders, which are a source of distress to many and a cause of substantial economical damage.

In the ’constructed brain’, the methodological and sociological issues are explored, which are currently hampering the integration of the vast knowledge that we have on the brain already. It is argued that, as long as there is no strong incentive for the various disciplines in brain research to cooperate on common projects, this situation is not likely to improve. A ’constructed brain’ could provide such an incentive, since its creation would require the collaboration of scientists from many disciplines.

A review is presented of techniques, which are potentially useful in integrating the vast, but fragmented, knowledge on the brain, which is distributed over many disciplines, into a framework like the ’constructed brain’. Finally, suggestions are made to start up such a project.

It is interesting to look at other sciences which have established multi-disciplinary collaborations, such as bioinformatics. It is clear that the Humane Genome Project has provided an enormous drive for the coordination of many activities in this field. Another field which is centered around large projects is high energy physics. The existence of only a few large accelerators in the world has also created natural 'condensation points' for this branch of science. In high energy physics knowledge of electronics, heavy engineering (accelerators and detectors are huge), detector physics and the underlying theoretical concepts of particle physics come together. High energy physics has created `WWW`, and has developped software suites for detector simulation, data analysis and visualisation, which are used by virtually every high energy physics laborotory in the world. Moreover, its database techniques and projects for distributed computing (the `GRID` project) draw much attention from other branches of science. This impressive computing infrastructure of high energy physics was developed by many people, from various disciplines, who were working together to bring a highly ambitious single project to a good end.

1.6.8 Executive summary of the Tools for Neuroscience project

This chapter reviews the importance of brain research for information systems and technology (IST). Notice that the importance of brain research does not just derive from these technological needs. The medical justification for brain research is overwhelming as 35 % of the disease burden in Western Europe is due to brain diseases. Understanding the brain is also a human objective, as we love and learn with our brains. Central to the strategy of brain research is the uniqueness of the brain amongst the organs of the human body and the importance of the 10^{12} connections in the human brain. These connections define at least 5 levels of integration between the molecules, and the genes encoding them, and behavior. Neglecting these intermediate levels, as has been done recently in some programs, dooms brain research. While we have potent techniques to address the lower levels of integration, we largely miss those addressing the supra-neuronal levels of integration critical to understanding brain function. This chapter sets ambitious goals to overcome these shortcomings: to record from thousand electrodes in 5 different brain regions while simultaneously obtaining high resolution multi-modal brain images and to develop new mathematical tools to organize and understand this wealth of information. This chapter stresses the need for education of the public in view of the ethical questions raised by the use of non human primates, which is seen as critical for these developments of brain research. Indeed non invasive techniques still lack in resolution. The alternative, in vivo studies, have to use adequate animal models and for higher cognitive functions monkeys are the only valid model. Unless these issues are better understood, brain and other pharmaceutical research will continue to leave Europe, further undermining the position of Europe in the world. In conclusion a strong investment into Brain Research will boost the collaboration between Neuroscience and robotics which is fast developing and provide a major source of inspiration for the whole of IST.

Chapter 2

The 'brainship' project

2.1 Introduction

Recent progress in fundamental neurophysiological research has made a popular subject of science fiction movies seem possible: direct interfacing of the human brain with computers and machines to create the cyborg. Indeed in 2004 the first implant of an electrode array was performed in the brain of a quadriplegic patient, allowing control of external devices including a robot arm (Fig. 1),(Hochberg, Mukand, Polykoff, Friehs, & Donoghue, 2005). This rapid development of a new field called neuroprosthetics has been made possible by the development of better electrodes and of fast signal processing techniques. This allowed chronic implantation of large arrays of recording electrodes in rodents and monkeys, leading to preliminary human application.

Nevertheless, present hardware is not sufficiently integrated and miniaturized to allow patients independent use of their neuroprosthetic implant. This lack of mobility led the first patient to request for removal of the implant. Even if this had not been the case, the expected functional lifetime of the implant was too short for general use in humans.

This major breakthrough was made possible by the surprisingly high plasticity of neural coding in the mammalian cortex. This plasticity allowed researchers to bypass the problem of figuring out normal neural codes and adapting the equipment to them, instead the brain adapts itself to the task even with communication over a limited number of channels. This discovery led to an explosion of application of brain machine interfacing (BMI) to control of cursors on a computer screen or artificial upper limbs in monkeys. In the first human application the brain plasticity was experienced as subconscious skill learning and once learned the patient did not have to consciously attend to the task (e.g. moving a cursor).

While these rapid advances are quite spectacular many more ambitious projects are now possible and needed if we want to progress towards cyborg-like applications. In particular we should consider other brain areas and other tasks than motor control tasks such as using an artificial arm. This will require interfacing of sensory input to the brain and interfacing to cognitive tasks like memory functions. To achieve such goals the BMI needs to be developed further and we will need a much better understanding of the underlying brain plasticity and brain coding mechanisms. While research oriented towards neuroprosthetic application needs to be done in monkeys to prepare for human implementation, much of the visionary research proposed here can first be done in cheaper rodents, as was the case for the original BMI studies (Chapin, Moxon, Markowitz, & Nicolelis, 1999). 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 in central and peripheral nervous system. Both sensory input and motor activity will be channeled through computer interfaces and in the ultimate experiment also social interaction among animals.

Obviously advanced BMI holds therefore great promise in expanding the neuroprosthetic treatment of neurological and trauma patients. More controversial applications lie in the direct control of remote robotic devices and information systems or enhancing humans with embedded machines as in the cyborg. The latter necessitates discussion and the development of guidelines for the ethical use of BMI in humans.

2.2 Objectives

To augment human interaction with its environment by enabling direct interfacing to sophisticated perception, robotic, prosthetic and information systems. In a first step animal models using brain implants of recording and/or stimulating electrodes to implement BMI in other contexts than limb control will be made. Necessary BMI implementations include artificial sensory perception, sensorimotor integration, control of legged locomotory devices, control of navigation and brain-computer-brain interfacing. In all these applications several functions normally performed by the brain may be implemented in the connected machines or robots, e.g. the pattern generation in locomotion may be done in silico, but it is important that the interaction with the brain is of high dimensionality. In a final phase fully bionic animals with bidirectional brain interfaces for control of high-dimensional systems will be created. Both neurophysiological knowledge and IT technologies will need to be greatly advanced to allow interfacing at a natural cognitive level or to embed the brain at an appropriate level in an artificial sensorimotor loop.

2.3 Examples of future human applications

- Full-immersion teleoperation by mental control of remote exploratory vehicles equipped with non-human sensors, ranging from microendoscopes to deep-sea vehicles with acoustic sensing to total teleaction over long distances.

- Interaction with information systems using direct perceptual input.

- Neuroprosthetics: repairing paralyzed or amputated humans by interfacing the brain with external sensorimotor devices to control limbs or prosthetic devices, reintegrating a severely disabled person into society.

- Initial test beds of the technology will be bionic animals that are completely dependent on brain computer interfacing to interact with their environment.

2.4 Current state of technology

Several recent studies investigated the possibility of predicting limb movements from the activity of multiple single-neurons in the motor cortex (A., 2003; Schwartz, 2004). After initial studies in rats (Chapin et al., 1999), this was applied successfully to monkeys (Wessberg et al., 2000; Schwartz, 2004; Taylor, Helms-Tillery, & Schwartz, 2002). The main objective of these studies was to find ways to control an external device in real-time (e.g. a cursor on a computer screen or a robot arm) using signals recorded from the brain (Fig 2). Such techniques could potentially be the basis of neuronal motor prostheses. A major conceptual breakthrough in this direction was achieved by experiments demonstrating the feasibility of real-time brain-control of a computer cursor in two (Serruya, Hatsopoulos, Paninski, Fellows, & Donoghue, 2002) or three (Taylor et al., 2002)dimensions under closed-loop conditions (i.e. the brain signals directly control the cursor). An intended target reach could be decoded from the activity of only 7- 30 neurons, when the subject had visual feedback of their brain-driven cursor movement. 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 (Lebedev et al., 2005), a process called brain plasticity. Recently this technology was also brought to the clinic with first implants in humans, demonstrating that the experience gained in monkeys translates to clinical application (Hochberg et al., 2005). Nevertheless, it should be emphasized that current electrode designs, with average life times of months to about a year, are not suitable for chronic implantation in humans. Despite these apparent successes the present state of the technology is still very limited: the arm control is rudimentary because it employs visual feedback only. The absence of sensory feedback and of sensorimotor integration is a major problem and requires extensive fundamental research and technology development. Additionally BMI could be used in many other context like, for example, navigation (Yoo et al. 2004) or inter individual communication, which could have direct applications but have not been implemented yet in animal models. Finally, the mechanisms of the brain plasticity underlying the current success of BMI applications (Lebedev et al., 2005)are poorly understood. It is important to understand the extent and limitations of brain plasticity, e.g. can BMI interfere with original function of the brain area used and what are the practical limits on new control tasks that can be trained, before one can propose human application of BMI technology in non-clinical contexts.

2.5 Problem areas

Neuroscience:

- Understanding the extent and limits of the brain plasticity: is this area specific, what is the long-term stability of plasticity, what is the capacity (how many overlapping codes can a brain area learn), when will plasticity interfere with its original function, how does it compare to normal learning, etc. This requires specific experimental designs with sufficient controls.

- Develop strategies to optimally stimulate sensory regions by central or peripheral implants.

- Identify the learning and coding strategy used during brain plasticity to develop better training methods and improve the corresponding software representations for both input and output to the brain.
- Understand how the brain integrates multiple functions both at a hierarchical level and across modalities and different brain areas to come to unified behavior. This requires simultaneous recording from many different brain areas, including subcortical areas. This is necessary step to develop BMI applications where some subroutines are delegated to the robot side and to understand how to integrate multiple BMIs (sensory, motor, memory, ...) in more advanced bionic applications.
- Study the usefulness of BMI as a new experimental paradigm which enables closed-loop studies of the nervous system. This is expected to provide for deeper understanding than the classis open-loop stimulus-response designs.

IT:

- Real-time encoding/decoding software for brain input/output signals, algorithms robust to noise, changes in signal quality and brain plasticity.
- Calibration/training methods to maximize signal transfer over limited number of channels of different qualities (using all signal properties like local field potential, multi-unit spike data, etc.). Make use of ability of natural brains to switch fast between different coding schemes.
- Methods for shared control versus partial autonomy in real-time brain robot interaction.
- Effective strategies for perception/decision/action chain in robotics necessary for partial autonomous action.

Materials technology:

- Stimulation electrode arrays to allow direct input to the brain of spatiotemporal sensory data.
- Longevity and durability of electrodes which need to be suitable for chronic implantation, also in humans. Study long-term effects of implantation.
- Research on alternatives to implanted electrodes, e.g. cortical EEG.
- Miniaturize all the electrophysiological equipment (filters, amplifiers, spike detectors), combine it with the control software and put it into wireless, battery operated configurations.
- Sensors and actuators which must have a performance as good or better than natural ones.

Ethical:

- Invasive technology may cause brain damage. When is this acceptable in patients? Is it at all acceptable in normal humans?
- Brain plasticity may interfere with the normal operation of the human brain.
- May we enhance humans with embedded BMI applications?

2.6 Future research

The IST program will focus on fully establishing BMI technology in animal models and on developing ethical guidelines for future use in humans. Technology and materials development geared toward human neuroprosthetic application should be covered by the life sciences programs. The goal of this challenge is to develop bionic animals, defined as animals which use BMI both for sensory input and to interact with their environment and with other animals. Interaction can be by controlling a robot, a prosthetic device which allows the animal to move itself or another animal which performs a task. In these models the brain closes the loop between computer controlled sensory input and computer driven action or motor activity. Some of this technology will be developed in rodents and will be subsequently ported to application in monkeys.

2.7 Immediate goals

The following are goals which fall within the goals of the FET program:

- Study the extent and limits of brain plasticity in BMI.
- Implant in multiple cortical and subcortical areas and study there in interaction in a closed-loop BMI context.
- Develop 'simple' new BMI animal models: perception, sensorimotor integration, locomotion, navigation, inter-animal communication. In a later step these can be integrated to make the fully bionic animal.
- Study methods for shared control versus partial autonomy in real-time brain robot interaction.
- Development of real-time encoding/decoding software for brain input/output signals which is robust to noise, changes in signal quality and brain plasticity.
- Create public fora to define ethical standards for use of BMI in human patients.

The following are goals which should be covered by the nanoscience program:

- Development of stimulation electrode arrays to allow direct input to the brain of spatiotemporal sensory data.
- Miniaturization of wireless neurophysiological equipment.
- Development of better sensors and actuators.

2.8 Ethical considerations

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.

References Chapin JK, Moxon KA, Markowitz RS, Nicolelis MAL (1999) Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. Nat Neurosci 2:664-670 Craelius W (2002) The bionic man: restoring mobility. Science 295:1018-1021 Fetz EE (1999) Real-time control of a robotic arm by neuronal ensembles. Nat Neurosci 2:583-584 Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. Science 233:1416-1419 L.R. Hochberg, J.A. Mukand, G.I. Polykoff, G.M. Friehs, J.P. Donoghue (2005) BRAINGATE NEUROMOTOR PROSTHESIS: NATURE AND USE OF NEURAL CONTROL SIGNALS Program No. 520.17. 2005 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2005. Online. Johansson RS (1996) Sensory control of dexterous manipulation in humans. In: Hand and Brain: The Neurophysiology and Psychology of Hand Movements (Wing AM, Haggard P, Flanagan JP, eds), pp 381-414. San Diego: Academic Press. König P, Verschure PF (2002) Neurons in action. Science 296:1829-1832 Laubach M, Wessberg J, Nicolelis MA (2000) Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. Nature 405:567-571 Lebedev MA, Carmena JM, O'Doherty JE, Zacksenhouse M, Henriquez CS, Principe JC, Nicolelis MA (2005) Cortical ensemble adaptation to represent velocity of an artificial actuator controlled by a brain-machine interface. J Neurosci. 25:4681-4693. Nicolelis MA (2001) Actions from thoughts. Nature 409:403-407 Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, Donoghue JP (2002) Instant neural control of a movement signal. Nature 416:141-142 Taylor DM, Helms Tillery SI, Schwartz AB (2002) Direct cortical control of 3D neuroprosthetic devices. Science 296:1829-1832 Todorov E (2000) Direct cortical control of muscle activation in voluntary arm movements: a model. Nat Neurosci 3:391-398 Todorov E, Jordan MI (2002) Optimal feedback control as a theory of motor coordination. Nat Neuroscience 5:1226-1235. Wessberg J, Stambaugh CR, Kralik JD, Beck PD, Laubach M, Chapin JK, Kim J, Biggs SJ, Srinivasan MA, Nicolelis MA (2000) Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature 408:361-365. Wickelgren I (2003) Power to the paralyzed. Science 299:497 Yoo SS, Fairneny T, Chen NK, Choo SE, Panych LP, Park H, Lee SY, Jolesz FA (2004) Brain-computer interface using fMRI: spatial navigation by thoughts. Neuroreport. 15:1591-1595.

Chapter 3

Brain-like computing on silicon

3.1 Introduction

Technological progress has led to the rapid increase in the number of transistors that can be included on a single chip. Nevertheless, most current computing architectures are based on a single powerful digital processing unit with a low degree of parallelism. This ’single processing unit scheme’ is limiting because waits for remote memory access increasingly dominate computation and most of the power is consumed by broadcast of synchronous clocks. Large numbers of idle transistors consume increasing amounts (nowadays > 25 %) of power. Logic designers struggle to enhance its computing power by implementing multiple cores on a single chip but scaling of this approach will be very difficult (Gibbs, 2004). One alternative is to implement bio-inspired processing schemes, based on massively parallel, scalable architectures and providing the kind of dynamic adaptation and reconfiguration found in the brain. Such schemes could be based on current silicon technology. In the short term they are unlikely to outperform conventional solutions except perhaps at the sensory periphery. But they could be the first step towards the development of novel high performance computing strategies and perhaps provide new insights into the working of the brain. This is the focus of our Grand Challenge.

3.2 Objectives

There are many reasons for implementing bio-inspired processing systems: a) bio-inspired strategies can provide more efficient processing; b) they can inspire novel computational primitives; c) they can be embedded in robots and other artificial devices, facilitating exploration of closed loop perception-action schemes, sensory-motor integration mechanisms, etc.; d) they can provide computing and communication primitives that are compatible with the nervous system, making it easier to build Brain Computer Interfaces.

3.3 Examples

Hardware implementations of bio-inspired processing schemes can be very diverse:

- **Implementation of vision systems on silicon**. A number of bio-inspired hardware implementations are already in use (e.g. Mota, Ros, Origosa, & Pelayo, 2004; Ros, Ortigosa, Agis, Carrillo, & Arnold, 2005; Anafocus, 2005) and several bio-inspired VLSI sensory and processing systems have already shown promising results (Liu & van Schaik, 2005; Lichtsteiner, Posch, & Delbruck, 2006; Indiveri, Chicca, & Douglas, 2006). There have already been several single-chip implementations of motion processing (Stocker, 2002; Delbruck, 1993) and stereo processing (Mahowald & Delbruck, 1989). But we are just starting to be able to build complex visual systems capable of emulating the early and middle stages of visual processing. This kind of computation requires dense processing and parallel execution of different algorithms. This means that in many cases the main problem is not computing but information transfer. A promising strategy is to separate sensing from computation and distribute the processing across several chips that communicate asynchronously using pulse-frequency modulation to encode information, like spikes do in biological neural systems (Serrano-Gotarredona, 2005; Chicca, Lichtsteiner, Delbruck, Indiveri, & Douglas, 2006). Alternatively, it would be advantageous to implement integration and disambiguation strategies on-chip. This could especially useful for active vision, when the results of visual processing feeds back to the image acquisition process.

- **Embedded processing in robots** (embodiment concept). 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 (Reichel, Liechti, Presser, & Liu, 2005). Compact devices that can process sensory signals locally, in real-time, perform data reduction and transmit a reduced amount of signals to the robot processing centers are essential for enabling the embodiment concept (Indiveri & Douglas, 2000).

- **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, requiring new real-time processing strategies (e.g. dynamic allocation of computing resources to different tasks). In hardware implementations, processing time is defined at a very basic level. In this setting, bio-inspired computing strategies that use time as a resource, could present interesting opportunities for designers.

- **Brain interfaces**. Current technology already allows real-time signal acquisition, processing and feedback. But in the near future, it will be possible to go one step further, simulating brain areas and interfacing them to areas in real brains. This approach could help, not only to understand basic functional primitives in the brain but also to build: a) simulations providing insights into the working of these primitives, b) brain prostheses to support sensory or motor peripheral 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.

3.4 Current state of technology

Progress in analog VLSI and in digital programmable circuitry, has allowed designers to build application-specific systems of ever increasing complexity. Examples include real-time visual processing devices for artificial perceptive systems, robots, implants (Implants, 2005; Seiingwithsound, 2005; Dobelle, 2005; Delbruck & Liu, 2004; Reeve, Webb, Horchler, Indiveri, & Quinn, 2005; Indiveri, 1999), and massively parallel processing architectures emulating spiking neurons (Indiveri et al., 2006; Liu & Douglas, 2004; Schoenauer, Atasoy, Mehrtash, & Klar, 2002; Ros et al., 2005), etc. Other researchers are working on strategies to exploit the inherent parallelism provided by FPGA and analog devices (ITRS, 2005; Xilinx, 2005). For these designers, the rapid evolution of technology is a challenge: it is not easy to devise scalable computing schemes that can make efficient use of the millions of logic gates now available on a single device. To achieve this, we need novel design strategies that operate at a higher level of abstraction than has been necessary to date. This is where bio-inspiration comes into play

Biological nervous systems are composed of relatively simple neural structures which can be composed into larger subsystems and systems. The performance of these systems on specific tasks improves during development, allowing them to adapt to the peculiarities of specific tasks, bodies and input/output systems. This is the kind of mechanism designers would like to incorporate in their circuits. Another topic of great interest is the way the brain dynamically allocates resources to different tasks (through attention mechanisms). Yet another is evolution, which we can now emulate in evolutionary hardware, using on-the-fly reconfigurable circuits (FPGAs) (Sipper, Mange, & Sanchez, 1999).

FPGA technology allows the design of complex systems. Modern FPGAs include not only general purpose reconfigurable circuitry, but also embedded blocks, such as parametrizable multipliers, processors, memory blocks, I/O transceivers, etc. (Xilinx, 2005). The presence of these resources makes it possible to implement high-performance, designs for real-time information processing, suitable for applications such as vision (motion processing, stereo, object segmentation, etc), multi-modal integration (with multiple sensory inputs), robot control, etc. One role in which FPGAs could prove particularly fruitful is as controllers for robots in experiments to explore closed loop Perception-Action schemas.

Nevertheless, the need to use existing technology forces engineers to adopt an 'opportunistic approach' by implementing computation principles from biological systems while still taking full advantages of recent technological advances (high clock rates, high communication band-width, etc). Examples include proposals to use bus multiplexing with AER (Boahen, 2004; Maass & Bishop, 1998) to address communication topologies far more complex than the limited 2D physical connectivity provided by current technology, or to use time slicing to simulate multiple neurons with a limited number of processing units.

Different classes of devices offer different opportunities. For instance designers of analog VLSI circuits can embed certain processing tasks within the signal sensors themselves (as in artificial retinas and cochleas). In these examples, sensing and processing are interlaced from the very first stage. This allows the design of sensors with a high dynamic range and local adaptation capabilities. Analog VLSI technology makes it possible to study processing schemes at a very basic level, manipulating the electrical properties of transistors, resistances and capacitances, the continuous signals these devices produce and the way they change over time. By contrast, modern programmable technology (FPGA devices) is mainly digital: analog signals are converted into discrete values and time is managed in slots (clock cycles). Nevertheless, this technology allows a very high degree of flexibility, time (though sampled) can be controlled at a very basic level, and the precision of computing operations can be defined in every stage. The same technology makes it possible to explore a broad range of very different computing schemes -though for true bio-inspired systems we will also need new design paradigms.

3.4.1 Building brain-like systems: adapting the technology

The adaptation of bio-inspired processing schemes for implementation with available technology is not trivial. The most promising solution in the long term is to develop dedicate VLSI solutions using both analog and digital circuits, exploiting the physics of silicon to emulate biophysics, and building massively parallel multi-chip neuromorphic systems. Alternatively, a parallel solution is to use the computational resources (multipliers, embedded processors, embedded memory blocks, etc) provided by modern digital FPGAs to implement massive parallelism in programmable devices. 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. Moreover, current design methodologies, the techniques used to embed computational primitives and other technological constraints may make it difficult to exploit their advantages. One first step should therefore be to develop optimal strategies for simulating neural structures on FPGAs. This will allow us to gradually develop computing strategies and design methodologies which are better adapted to the goal of implementing brain-like processing. This cannot be achieved in a single step. But reconfigurable technology will make it easier to develop new processing and design strategies (e.g. by allowing the design of asynchronous circuits).

3.5 Problem areas

There are many problem areas in which hardware implementations of bio-inspired hardware could be useful:

- Bio-inspired processing schemes requiring real parallel processing.
- Asynchronous systems.
- Systems using spike-based processing and communication.
- Real-time processing systems for vision, hearing, olfaction, complex robot control, etc;
- Brain-machine interfaces.
- Brain-like processing primitives in silicon.

3.6 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, and better understand how to incorporate massive parallelism, adaptation (including self-adaptation to the environment), learning, evolvable architectures and real-time processing in VLSI devices. In the longer run, it will be necessary to adapt these schemes to simulate specific biological systems: current schemes are handicapped by having to use computing primitives developed for other purposes. There are many, partially overlapping, areas where circuit designers should play an important role in Neuro-IT. These include:

1. Bio-inspired solutions to real world problems: artificial cochleas (Watts, Kerns, Lyon, & Mead, 1992; Liu & van Schaik, 2005), artificial retinas (Boahen, 2002), neuromorphic vision chips (Indiveri & Douglas, 2000; Delbruck & Liu, 2004), neuro-cortical implants, image processing, eg. motion processing (Kramer & Koch, 1997), stereo (Diáz, E., Sbatini, Solari, & Mota, 2005), etc.

2. Neuromorphic approaches: opportunistic design strategies taking new ideas from biological processing and combining them with techniques (such as time multiplexing) that take advantage of the implementation technology, while working round its constraints (such as limited connectivity). Possible strategies include: Address Event Representation (AER) (Boahen, 2004; Indiveri et al., 2006; Serrano-Gotarredona, 2005), event driven communication, and event driven simulation schemes (Schoenauer et al., 2002), etc.

3. Development of simulation platforms: there are several fields (such as spiking neural networks) that could benefit from ad hoc circuitry allowing efficient simulation of large scale systems (Schoenauer et al., 2002; Ros et al., 2005).

4. Real-time processing systems: machines that close the perception-action loop, interacting with the real world.

5. Evolving, self-adapting systems in real world environments: reconfigurable circuits (Sipper et al., 1999) that adapt themselves to the changing characteristics of the environment. From an evolutionary point of view, such systems would exploit information about task performance and resource use to generate a new generation of more effective machines. This is not the kind of problem that can can be resolved with simulation: it needs physical agents that evolve in physical environments.

6. Robotics: on-board processing resources (for perception, self-moving, communication, etc) making it possible to build autonomous machines with perceptive capabilities (Liu & van Schaik, 2005). Such machines could be used in studies of swarm intelligence or robot societies. Different machines with different morphologies and computational capabilities could specialize in different tasks, with each contributing to the goals of a robot society.

3.7 Immediate goals

- Implementation of bio-inspired parallel processing schemes, providing higher performance than is possible with simulation. For instance smart vision systems.
- Pushing the technology to include primitives and design methodologies suitable for brain-like computing.
- Hardware implementations for real time processing. Providing robots with powerful sensory-motor capabilities.
- Using advances in current technology to implement neural-like processing schemes.
- Using attention mechanisms for dynamic reconfiguration of hardware devices and on-the-fly allocation of computing resources.

3.8 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. The experience helps them to understand how biological systems work thus providing insights with practical implications, in medicine and elsewhere. In the long run, 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. This will represent a technological revolution -making it possible to perform tasks which are not currently possible. For instance, it might be possible to create new aids for handicapped people or to provide sensory augmentation for use in dangerous situations. If we really succeed in implementing brain like computing this will have a huge impact on industry and on human society as a whole -allowing the creation of devices with currently unimaginable capabilities.

Chapter 4

Mind-Body Co-Evolution: The Factor-10 Project

4.1 Introduction

Both the emerging fields of epigentic robotics and "smart" materials science offer a wealth of innovative research opportunities and promise a large spectrum of new products to become feasible in the mid to long term. However, a completely new discipline may develop by combining key research in these fields to work towards new types of artefacts. We envision three types of such artefacts to emerge from this combination

- **Type I:** artefacts evolving their cognition and motor control autonomously based on multimodal/multisensory feedback in a predefined and fixed body, whose structure may be optimised to perform a certain class of tasks (designed to a certain "ecology" – as is frequently the case for living organisms), e.g. the "dancing robot" of the AI lab at Zurich University 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.

  - robots with effectors made from material with mechanical plasticity, such as shape memory alloys and/or autonomous control intelligence (peripheral nervous system) in these limbs, like truly dextrous "hands" with a highly developed sense of touch, or
  - fine grained versions of the current attempts to design "modular robots" that may change their body shape to some extent by combining basic actuator modules into different shapes,

  and

- **Type III:** artefacts that co-evolve their (possibly distributed) brains (system) and their body in permanent interaction with the enviroment over an extended period of their lifetime (embodied artificial ontogenesis).

While the implementation of the first type of artefacts largely depends on progress in control architectures of cognitive systems, the latter two will also draw heavily on methods and technology developed in materials science for self-assembling materials and structures, "constructive chemistry" and – most probably – proteomics. In particular, the third type may be 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 in that it can only be realised as an *autopoietic* machine built from cells, each of which implements a full recursive plan for bodily development and behaviour in a given enviroment similar or identical to the genetic code in the cells of living organisms.

Following these lines of thought, we propose to define a long-term research project called "Factor-10" or Factor-X, which aims at fully functional physical artefact (i.e. not a computer simulation), which, during an extended but limited period of time (e.g. 10 months) autonomously grows

- *the volume of its body by at least a factor of ten, thereby differentiating out "organs" and "effectors" as well as*
- *its cognitive abilities (its "IQ" and its repertoire of sensorimotor behaviours), also by at least a factor of ten.*

This vision is obviously largely inspired by the development of living organisms and the theory of "enactive" or action-centred cognition by Varela (Varela, Thompson, & Rosch, 1991) – intelligence and autonomy in can only emerge in embodied creatures (and artefacts) as the result of their permanent interaction with a real enviroment. Based on this theory, one may even argue that eventually the implementation of such artefacts would be based on (modified) biological substrates (an hence become a case for genetics) because nature has solved exactly the same problems of survivial on the earth in the optimal way – through creating living organisms for the most diverse ecological niches. This would entail, however, not only massive ethical problems, it would also delimit the range of size of the basic building blocks (i.e. biological cells) and – depending on their machanical stability – the size and properties of the artefacts. It would also require the problems of artificial metabolims to be solved.[1]

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 and – only due to technological deficits – treat mind development and bodily adaptation through (and for) interaction with the environment as two separate problems. Being aware

---

[1] After all, this should be considered an additional challenge – not an impediment. There can be no doubt that a highly efficient and lasting energy supply is an indispensable constituent of autonomous artefacts. It is quite likely that the solution to the totally inadequate energy cycles based on electrical batteries may be found in copying the chemical pathways found in life. One may even argue that the search for food (not to be confused with simple search for a battery loading dock), which is a special kind of interaction with the environment under time pressure that has direct consequences on the constitution of the body, is an essential driver for mind development and cannot be separated from the artefact growth process. Whilst of high importance in its own right, research into adequate energy cycles for autonomous artefacts is definitely out of the scope of this challenge.

of this conceptual deficit, we should permanently aim at overcoming this artificial separation as soon as possible and capitalise on every technology advance that offers a potential to do so. In particular, research in (molecular) biology should be constantly monitored and and regularly be evaluated for progress made there for applicability to any of the research areas contributing to Factor-10.

4.2 Motivation and objective

For at least the last five decades the general public has been promised the advent of universal robots or even human-like artefacts that would be of real help to us in our daily lives and/or possess super-human capabilities. However, as expectations rose, science consistently failed to deliver robots that can be compared to biological creatures, not even to those with very low-level intelligence.

Notwithstanding this failure, enormous progress has been made in many fields potentially contributing to the design of truly autonomous artefacts of types II and III outlined above, such as brain and cognitive science, information technology and artificial intelligence, molecular biology and chemistry that the time is ripe to combine/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.

Moreover, apart from being one the most exciting research goals to pursue, artefacts that can – at least to some modest degree – develop an autonomy of their own in the literal sense of the word[2], would also be an economical market that cannot be underestimated.[3] Despite current wave of euphoria for humanoid robots – largely fueled by industrial companies like Honda and Kawada but also by applied and basic research projects as the Japanese HRP program and the Kawato Dynamic brain projects and its continuations –, it will soon become clear that these machines (type I according to the above classification) are impressive feats of engineering and highly interesting platforms for developing basic technologies, but they hardly lend themselves to any practical use outside of robot labs. Only when the qualitative transition to type II artefacts can be achieved will we see practical solutions that will find acceptance by a broader public for many interesting applications (see (Knoll, 2002) for an incomplete overview).

However, issues central to the development of living creatures that would have to be followed to a higher or lesser degree for type III artefacts, i.e. synchronous evolution of morphology and mind, have hardly been formulated, let alone been tackled. Fortunately, due to the need for a qualititative breakthrough, already from type I to type II artefacts, and the high quality of European research in the aforementioned disciplies contributing to type II and – in the longer run – type III development, there would be a window of opportunity for Europeans to compete with

---

[2] Meaning "give oneself one's own laws of behaviour" through "living a plan" by evolving all aspects of one's being there, instead of just executing a designer's plan (Ziemke, 2001).

[3] While about 10 years ago the market for service robot and/or assistance systems (for both home and factory use) was projected to be larger than 1 billion EUR by the year 2000, only very few such service robots (less than one thousand) have actually been deployed so far. 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 increased substantially and their programming efforts were reduced just as drastically.

Japanese research at the next stage of development – the Japanse advantage in humanoids (type I) research will hardly be caught up to.[4]

Looking at the preconditions for embarking on this research journey, we note that a there is already a sizeable body of research in the diverse necessary disciplines represented in Europe (see the non-exhaustive list in the appendix), however with fragmentation across disciplines and countries.

Apart from the scientific objective of developing the basic technologies and actually designing as well as building prototypes of type III artefacts – via type II as an intermediate goal – along a far-stretched time line, it is also the purpose of the project to establish a commonly accepted paradigm for designing these artefacts. Initially, recent results will be collected and translating into a language common to all the disciplines. More importantly, however, is the development of completely new theories, methods and paradigms controlled by carefully studying how the methods from one field can guide the research directions in another (e.g. by evaluating research results on imitation from psychophysics to define paradigms of machine imitation learning that can be translated into computer-operational algorithms). In parallel, for every milestone reached, its application potential inside and outside of the artefacts will be studied so as to ensure feedback to the research community of the newly developed field as to what features would be particularly useful to have in real systems, e.g. robustness and safety after failure, behaviour stability, reaction-times, cross-modal sensory input processing etc. – all in dynamic, unpredictable, adverse and partly unknown or even completely unseen, uncharted real-world environments.

The goals of Factor-10 are indeed very demanding. Up to now, they have hardly been formulated as a common integrating challenge because of the deterring technological impediments in every single area involved. We believe, however, that in view of the progress achieved in many of the disciplines, particularly 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.

4.3 State of the art and projection from today's viewpoint

As of this writing, there is a small body of published work on experimental systems, design simulations, materials analysis and proposals for architectures that may serve as starting points for further research, in particular:

- **Modular robots** that are built from a certain number of identical motor modules and can be combined into different shapes and macro structures (e.g. (Kamimura et al., 2001); see (project, 2003) for an overview).
- **Evolutionary and epigenetics robotics** both in the sense that robot shapes are optimised (e.g. (Funes & Pollack, 1999)) according to certain target functions and that the principles of autonomous learning based on very basic instincts is concerned (e.g. (Nolfi & Floreano, 2000))

---

[4] Honda claims to have invested in excess of US$ 100 mio. into their humanoids development program, which started in 1986, and the Japanese Humanoids Research Program (HRP) received another US$ 25 mio. of direct funding. Other Japanese giants like Sony, Fujitsu, etc. have not disclosed their figures – but they may be just as high.

- **Microscale structures** that can be assembled according to external conditions and that can serve as filters, modulators, etc. for chemical reactions..
- **Nanoscale self-asembling structures** were proposed that can build up aggregates of macroscopic size, e.g. for "muscle tissue" (?), that can be made to grow and exhibit useful properties, such as need for joints without lubricants, etc. As far as this field is concerned, were are confident to profit from nanotechnology (including nanomanipulation) to provide us with materials that can be used in different functions in the artefact. Of particular interest would be the technlogy that enables nanostructures (e.g. nanoscale motors) to build up in a controlled way – as long as these technologies have the potential to be used in an artefact. This would rule out the use of processes that rely on extremely high voltages or extremely high external pressures for the structures to from themselves.

This is just a selection of the competences needed to be integrated for a first step along the type III developments. Obviously, all of these fields are only in their beginning as far as the use of their potential for specific contributions to our goals are concerned. There are a number of research areas that may directly contribute 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, but also the principles that led to the formation of both motor and sensor entities, e.g. what drives phylogenesis to get from a single photoreceptor to "open" insect facet eyes and then on to lense-based eyes, what are the driving factors behind the development of different locomotion principles, in particular muscle-joint constructs, etc.
2. *Genetics*: Contribute a set of rules that allow to encode a certain minimal set of "genes" which allow stable bodily development but also the control of the communication between the individual body cells so that they can – in interaction with the enviroment of the artefact – develop a certain desired behaviour. A controlled modification of these genes should also result in a predictable change of behaviour development of the artefact.
3. *Computational Neuroscience*: Given the freedom of growth and structural developments of information processing entities in the artefact (but also the severe technological constraints), develop *appropriate* basic processors (neurons) *along with their interaction principles and communication networks/mechanisms* that enable the parallel and interleaved emergence of motor skills and cognitive skills taking into account the hypotheses about structural coupling according to Varela. Clarify issues of "assemblies" and regions of the basic processors building up and structuring themselves according to the genetic code during the artefact's evolution and their dependence on the environment in which the arefact grows up.

It may be argued that there are good reasons to carefully discuss and review the size and functionality of the ideal basic block (and hence he variability) for the growing artefact: should this basic building block be the atom, the molecule, the constituent parts of a micromodule (analoguous to the internal parts of the biological cell), the micromodule itself (corresponding in functionality to the biological cell), assemblies of micromodules at the level of organs – or intermediate stages between these individual granules. Seen from today's perspective, the basic block of type III artefacts will probably have to have most of the properties of what is attributed to stem cells of animals (with or without the ability to cater for its own energy supply): with minor changes in its own reproduction program it can differentiate into cells for the most diverse requirements of the body, affording the different abilities for sensory, motor and processing tasks inside the complex whole of the body. It seems that there will be a natural transition in granule size between the cell-like basic unit of type III and the larger unit size for type II as we go from types II to III, but this cannot really be predicted now.

From today's point of view we see four essential threads of technology research (as opposed to the indispenable conceptual lines of work mentioned above) that should form the basis for an integrated research plan and should be pursued both individually and carefully interwoven to traverse the huge tree of possible individual actions, with (1) being the precondition for the practical implementation of (2)…(4), not, however, for the theoretical investigation of the latter three:

1. *Molecular Robotics*: exploration and design of useful materials and substrates (nanotechnology and chemistry) lending themselves to build cells that can be made to meet the different requirements in the variety of body areas/volumes, e.g. high mechanical stability (for "bones" and "joints"), ability of energy transformation (for "muscles"), for information exchange ("networks of nerves"), information processing ("neuronal assemblies"), etc. The emphasis should be on materials that have the ability to bridge the gap between the micro-level and large-scale macroscopic structures.

2. *Distributed growable sensors*: for distributed areas of sensing "cells" that are sensitive to different physical modalities (force, light, odour, temperature), it will be necessary to investigate how they can be coordinated and produce sensible results when they are located over large areas of the outer surface of the body and are physically connected together through a medium (i.e. the body) that shows a high degree of plasticity. Of equal importance is the exploration of the role of preprocessing sensor data, either directly in the sensor (such as the preprocessing taking place on our retina), over pre-structured nerve channels (such as the visual chiasm) or the early processing stages in the cortex – i.e. why/how these predetermined structures have evolved in the phylogenesis of creatures and to what extent it makes sense to mimicked this concept in the artefact.

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. This implies that it not only has to permanently control the artefact's evolving sensors and effectors, it also has to exert control of the interaction with the environment for exploration and task purposes so as to control its own development – while it is growing itself in physical size as well as complexity and is to develop cognitive and motor skills in parallel with the sensors' processing capacities. The challenge is hence

not only to achieve a stable learning and growth behaviour of the information system for body control but also to make the system develop its own new structural skills, e.g. the emergence of the concept of "memory".

4. *Growable motor entitites and spatially distributed actuators*: the actuators must also be controllable as they develop both their actuation part (the muscle 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 and they must be supported in the artefact body so that mechanical stiffness and stability is achieved along with a maximum of locomotion effectiveness, energy efficiency and durability.

Ideally, it will be possible to formulate – at an appropriately high level of abstraction – principles of growth (like the competition metaphors for selection of species – but also for the development of synaptic connections), which govern the growth processes in the artefact, i.e. a straightforward and easy-to-formulate principle in terms of a target function like entropy maximisation, energy minimisation, sparseness etc., such as the principles recently discovered for the development of the different types of neuronal cells.

From a technology development view point, we suggest to lay out a plan which initially centers about the basic building block (BBB) in view of the four aspects above:

- *Functional properties*: what are the components that the BBB consists of? What is the minimum amount of functions integrated into one BBB? How can BBBs arranged in such a way as to form a large area distributed sensor, a distributed actor or passive support structures, respectively? 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*: how can the individual parts interact over different communication channels, not necessarily only through electrical connections? Studying the interaction patterns 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. These have to be clearly defined with respect to achievable plasticity, networking parallelism, scaling from a few to millions of nodes and further parameters.

In parallel, the development of convincing application scenarios scenaria should be advocated. This not only pertains to useful deployment on the factory floor, in private homes, outdoor support etc., but it also involves the transfer of parts of the technology to applications that could profit from, say, microscale machinery with integrated sensing and information processing abilities for medical use.

4.4 Expected Results: What will it be good for?

In Table 4.1 we have listed some of the possible applications of spinoff knowledge of potential research carried out within the framework of 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 modules. It may turn out to be more useful to start with the development of the latter type of BBBs right away, but this will have to be cleared up in a separate step. We have also listed some of the spin-off applications that may be the result of partial aspects of the developments. In particular for type III artefacts, the range of applications that one can imagine is so huge that it would be beyond the scops of this roadmap-contribution to describe them all. Suffice it to say that given the basic blocks are cheap enough all kinds of “intelligent structures” of small to large sizes may build themselves and can also change their shape according to various user needs. However, only the future cann tell if such a vision may come true and if such potential applications, which are far beyond the current conception and understanding of robotics, are a desirable addition to our daily life in terms of cost/benefit ratios. On the other hand, it is clear that the small-size artefacts we shall be enabled to build can most certainly finally deliver what robotics science has promised for a long time.

Expected Result	**Application of the 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 and a combination thereof.	Handicapped 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 (management) 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 enviroments (ultimate goal of the Factor-10 developments).	Unlimited range of applications. From microscale (e.g. use inside blood-vessels) to creatures of animal-like shape up to free-form structures with intelligent behaviour and distributed sensing (e.g. for house or road construction purposes) to symbiotic human-artefacts use (e.g. for increasing stamina, cognitive skills, etc.).
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.

Table 4.1: Application areas and users of direct and spin-off results from research under the umbrella of Factor-10.

Chapter 5

Successful Thinking and Acting in the Physical World

5.1 Introduction

Peripheral devices as well as the methods (e.g., buses and protocols) used to integrate them into a system have been important areas of mainstream IT for decades. Strong trends towards embedded IT devices and pervasive computing are likely to increase the importance of periphery and system integration aspects even more in the future. Since NeuroIT is an approach to improving IT, it needs to address these issues as well. Furthermore, there are substantial arguments suggesting that these aspects will carry even greater weight in neuroscience-inspired IT than in mainstream IT. These arguments are based on the nature of NeuroIT's specific goals, on the nature of neural systems and on the principal obstacles which need to be overcome in order to achieve the goals.

The principal argument for placing emphasis on periphery and system integration arises from the still elusive reasons for the existence of many sorely-felt performance gaps between natural neural systems and man-made artifacts. NeuroIT can be seen as an attempt to close these gaps by learning from nature. If the causes for these gaps were known for certain, closing the gaps would — at worst — require an incremental improvement in IT capabilities (e.g., more computing power), but this seems not to be the case. Instead, what seems to be hampering progress is a lack of operable concepts as to how IT capabilities, especially computing power available today or in the foreseeable future could be put to use in order to successfully reproduce the wide range of capabilities seen in natural neural systems. To remedy the situation, a set of working hypotheses for possible causes is needed.

One such hypothesis (emphasized in the 'Conscious machines' grand challenges) is that brains (in particular, human and non-human primate brains) have computational capabilities for reasoning, planning, assigning saliency and priorities, etc., which are vastly superior to man-made algorithms devised to reproduce these skills. It may therefore be speculated that superior algorithmic/computational capabilities account for the major portions of the capability gaps. While this is certainly a valid hypothesis worthy of further exploration, it is not the sole possible explanation.

In fact, when engineering a system intended to perform 'intelligently' in the physical world, designers have a wide variety of possible options. These can be grouped into four classes:

Choice of Computation & Control Strategies plays 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.

Choice of Morphology can be significant in facilitating or complicating the task of a physical agent: the right kind of body can be crucial. For example, the body shapes of animals which live in the spaces between sediment particles ('interstitial fauna') are crucial to settling this environment. Likewise, the choice of wheels or legs for locomotion significantly influences accessibility and traversability of terrain.

Choice of Materials is also crucial in high performance sensory and actuation systems. For instance, sensing of mechanical waves, like displacement and acceleration detection by mechanoreceptors or hearing of sound waves is accomplished in large part by the materials properties of the sensors; the characteristics of animals' effector systems — speed, compliance, efficiency — are determined by the materials properties of muscles, tendons and connective tissues.

The Environment itself can also be engineered or exploited to facilitate the system's performance. Examples include making use of the pattern of polarization in light scattered by the upper atmosphere as a compass sensor, or marking the environment to store navigational information, such as pheromone trails or with street signs.

An alternative hypothesis is therefore that natural neural systems are superior primarily because they are better integrated and optimized 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 — often overlooked — the physical phenomena created by the interaction of the agent with its environment. Along these lines, it may be speculated that many of the tasks which neural systems excel in are not as hard to perform as it seems, if they are posed in the right fashion. Significant reductions in task complexity can be realized if each component in the control loop solves a simplified problem while relying on the other components to create the conditions that make these simplifying assumptions valid (Brooks, 2002). Therefore, an intelligent, well integrated periphery which exploits the possibilities implicit in all the dimensions of the design space, may be the a key to lowering task difficulty from 'impossible' to 'possible' or even 'straightforward' thereby alleviating the need for more powerful central algorithms.

Fundamentally, these two hypotheses are not mutually exclusive. Both are most likely part of the complete explanation of intelligent systems in nature. What remains to be seen is to what extent they contribute to the technical reproduction problem and hence would have to be part of the solution. As long as neither can be dismissed as a negligible factor, both hypotheses should be pursued as grand challenges. While these grand challenges are interrelated and should ultimately become part of an integrated solution, each of them will probably maintain a distinct emphasis for the foreseeable future. It is therefore in the interest of maintaining clarity and focus to mark them as distinct approaches.

One further goal of NeuroIT is to make IT artifacts sufficiently intelligent that they can interact with humans in a natural way (s. 'Conscious machines' grand challenges) or interface successfully with the human brain as a useful replacement (prosthesis) or extension (s. 'Brain interface' grand challenge). In either case, adequate periphery will be of prime importance, because it will be a determining factor in how humans will perceive the artifacts and hence judge their impact on the quality of life. For example, it is very likely that a robot which looks like a 'trash can on wheels' (as iRobot's B21r and similar designs) will be perceived by humans primarily as a 'trash can on wheels', no matter what computational skills it is endowed with. As a matter of fact, this issue goes beyond human perception, since an intelligent artificial being embodied in a trash can on wheels will indeed have great difficulty to be a part of human life, because it lacks the physical capabilities to do many of the things that humans can do (e.g., climb stairs, sit down at a table). A generalized version of the latter aspect, namely being able to perform useful functions in the human environment or any environment into which there is a scientific or economic interest is not limited to natural interaction with humans. It is likely that NeuroIT artifacts can be perceived as useful and hence gain wide acceptance in society, if they perform useful services, many of which will not require them to interact with humans in the way one human interacts with another. Design of NeuroIT artifacts for such purposes may hence turn to the vast knowledge pool made up by the designs of the millions of non-human neural systems found in nature.

5.2 Objectives

The objective of the 'Successful thinking and acting in the physical world' challenge is to build complete systems which make optimum use of distributed intelligence embedded in the periphery (sensors, actuators, body morphology and materials) and at a system integration level. Research on this grand challenge will emphasize

- 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

Distributed, embedded intelligence should enable the artifacts to master tasks known to be performed by natural (neural) systems but currently elusive to technological reproduction. It should have a significant effect on central neural computations taking them to until now unattained levels of efficiency. Despite neural processing remaining an integral part, the focus of this grand challenge is on making the periphery smarter and integrating it better with central computations, so that the whole system gets more powerful and efficient. In addition, knowledge about the tasks to be performed and the world that they are to be performed in should be integrated at every stage. Efficient ways to distribute the storage of this knowledge, i.e. 'environment models', over the different subsystems should be developed. Ultimately, it should be possible for designers to have confidence that each part of such a system exploits all reasonable available prior knowledge. The same should be true for the system integration level. This calls for advanced methodological achievements in gathering the relevant knowledge. Optimization processes in nature operates on large time-scales and vast numbers of prototypes for testing. In order to apply such optimization to every part, every integration level and every task, shortcuts need to be found which narrow the search space so that it can be managed within the scope of a engineering design process.

In following the outlined approach and addressing the involved issues, research on this grand challenge will produce novel, smart peripheral devices for NeuroIT systems and thereby promote the pervasive use of intelligent robotic systems. While individual projects will probably have to include many case studies, work on the grand challenge as a whole should establish general rules, how these objectives can be achieved for any tractable problem. Obtaining a set of design rules will enable tailoring solutions to IT problems without further need to study specific natural solutions. Consequently, the design rules may be applied even to problems for which no solution is known to exist in nature.

Research results should lead to the creation of universal standards (e.g., 'bus standards') for smart NeuroIT peripherals, which would enable closer cooperation between research projects (parts developed in one project can be reused by another) and also facilitate the inclusion of novel technology into product design. A pool of smart, readily available periphery should not only provide the building blocks for powerful individual systems (robots) but also establish new capabilities for robot interaction and collaborative behaviors, either between self-contained individuals or parts forming 'states' or super-organisms.

5.3 Examples

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

  Such artifacts could — for instance — be made so ubiquitous that they could coexist with humans in symbiotic human-robot ecologies in which they would enhance the quality of human life. This would not necessarily require the artifacts to be capable of 'social interaction' with humans, but they could rather maintain human environment quality in a pervasive, yet unobtrusive manner. One may envisage that the organisms adopt behavioral patterns from animals which coexist with humans as commensals or parasites (for example, mice, cockroaches, geckos), but perform useful services to humans. Such symbiotic ecologies could be established in a variety of contexts, for example:

smart home ecologies: Humans share their homes with unobtrusive creatures, which, for example, keep the house clean (not only carpets, for which there is prior art, but any surface that needs care, like bath/kitchen tiles, kitchen sink and plumbing, clean and iron clothes hanging in the wardrobe, etc.), establish and adapt wireless communication infrastructures (between household appliances as well as with the outside, smart teleconferencing equipment) and are rewarded with access to power.

public spaces ecology: Perform cleaning of floors and windows, remove litter. Maintain and repair wall/roof surfaces, 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: Optimize sensors to monitor patients' health, fault detection, provide better comfort by reacting to symptoms of patients' discomfort with changes of environment (for example, temperature, lightning, humidity, noise cancellation and music), adapt the environment to be best suited to patterns of emergencies, for example, diagnostic equipment which optimizes its diagnostic skills.

communication ecologies: Optimize — in particular wireless — communication channels to maximize transmission quality, efficiently share resources, minimize power consumption in order to increase battery life and decrease EMI health and safety risks. An example for EMI related safety risks would be interference with navigation or other vital systems. Future wireless communication devices carried by passengers on an airplane could intelligently adapt to the navigation/communication needs of the airplane, granting them absolute priority automatically and removing the need for absolute restrictions on their use and enforcement of these restrictions.

security ecologies: Perform and optimize monitoring and surveillance tasks, optimize sensor configurations and communication links according to current sensing conditions or patterns of security breaches or vandalism. For example, instead of installing surveillance cameras at fixed locations, they could be mounted on agile NeuroIT artifacts, which can continuously adjust position and orient them to maximize image quality under changing lighting conditions or in response to noises. If these agents are sufficiently agile, they could even escape attempts to disable them.

playground ecologies: Enhance the value of recreational facilities for both children and adults by making them more entertaining, more likely to practice valuable skills and also safer by reducing the probability of accidents. For example, slides and swings could keep track of acceleration forces and make adjustments to dampen excessive accelerations.

- The study of simple 'organisms', both natural and man-made, that allow detailed analysis of their entire neural system, i.e. periphery and CNS, while performing natural tasks in challenging, natural environments or faithful laboratory reproduction thereof.
- The study of non-human and possibly 'super-human' senses and actuation principles found in nature, in order to lay the foundation for artifacts which can not only replace human labor but also extend human capabilities, for instance to enable successful living in hostile environments.
- The study of how '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. Starting from simple organisms should allow to study the evolution of gradually more complex models.

5.4 Current state of technology

The periphery and system integration of current NeuroIT artifacts is in general still woefully inadequate and lagging behind natural systems. There are many examples where this is significantly hampering technology. For instance, in autonomous driving the limited dynamic range of cameras makes it impossible for these systems to cope with the same range of situations as human drivers (for example, driving through a tunnel). In the same context, clever use of eye movements has been employed already to improve performance in negotiating curves and intersections (Dickmanns, 2002).

The physical embodiment of most experimental/research systems is either inadequate for interesting tasks (for example, trash cans on wheels) or too conventional (RC car or RC airplane) to yield significant advances in understanding or performance. Researchers in the field need ready access to much better embodiments or components for tailoring their own embodiments according to their specific needs. 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 serve to turn NeuroIT periphery into a widely used, mainstream technology at first for research purposes and later for the mass market. In doing so, it will push the limits of what such periphery and integration schemes can do and what NeuroIT researchers whose main research goals are not periphery itself can do using it.

In the field of natural scene statistics, ways to obtain and utilize probabilistic knowledge about the environment have been explored already. This kind of work needs to be taken to a new level where it can provide a more powerful and general framework for development of technology.

Projects like the DARPA-funded ‘smart dust’ project have been addressing the issue of deploying many cheap sensor modules. However, most research efforts seems to been directed mainly towards the networking and mass-manufacturing aspects, the smart dust grains themselves could be a lot smarter (and yet remain cost-effective to produce).

Despite the extensive research that has been conducted in the cognitive sciences into how space is represented by brains (Paillard, 1991) very few concepts have emerged which have been successfully applied to artificial systems. Hence, independent from this line of research, mobile robotics has investigated representations of the environment that allow autonomous systems to perform useful tasks in natural, unstructured environments. However, in particular, outdoor mobile robot applications are increasingly based on the use of GPS instead of environment models, restricting their application to environments where GPS-like information is readily available and making such applications less likely to become good models of natural systems.

Subsumption-based architectures (Mataric, 1992; Brooks, 1986), have paid special attention to the distribution of ‘environment models’ over the different subsystems necessary to control an artificial agent. This approach however, is only loosely based on living organisms and not on a systematic study of principles gleaned from control in natural systems.

5.5 Problem areas

- Find ways to mass-manufacture and assemble the parts of advanced NeuroIT devices. For small structures, MEMS technology may be a solution, but many structures will have to be larger in order to perform the intended functions. Rapid prototyping technologies should be looked at in the context of using materials and creating shapes particularly well suited for NeuroIT artifacts.

- Find ways to analyze an organism’s (natural or biomimetic) 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 the execution of a natural task.

- Improve understanding of the trade-offs between the different kinds of design choice available, as a function of task, environment, cost and technology, and find ways to support rational design of the complete system.

- Develop novel sensor and actuator technology to support the smart, biology-inspired, peripheral systems.

- Application of non-linear dynamic systems theory for analysis of interaction organism and environment.

5.6 Future activities

A systematic effort should be undertaken to facilitate the development of next-generation NeuroIT periphery. Today, embodiments for research systems can be obtained from — typically small — companies which cater for the needs of experimental robots. Because they are limited to small markets, these companies lack the resources for bold innovation and consequently their designs are very conservative and leave a lot to be desired in terms of capabilities and performance. Alternatively, such embodiments are developed in research labs as one-of-a-kind systems which take a lot of man-power to develop but often enjoy very limited use beyond the scope of the research project they originated in. Remedies for this situation should be systematically worked on in order to give researchers access to peripheral modules and system integration frameworks with capabilities and performance levels which go far beyond what is generally available to them today. At the same time, technology and markets should be developed in the research stage already in order to pave the way for the development of NeuroIT devices for the mass-market. The activities necessary to achieve these goals include:

- Develop benchmarking standards to stimulate and monitor the improvements of NeuroIT systems. This could take the form of a broader set of 'Turing tests' for NeuroIT systems, which perform tasks other than armchair conversation with humans. For instance, artifacts could be tested by making them interact and specifically compete with the biological systems they are meant to reproduce. For example, a robotic fly could chase natural flies, a robotic tuna could capture natural squid successfully, ...

- Develop and standardize general, flexible protocols for interfacing NeuroIT periphery, both physically and in terms of data communication in order to foster exchange of modules between researchers and prepare the ground for industry standards needed for the future commercial use of NeuroIT devices.

- 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. Specifications for bus and protocol standards as well as benchmark scores will be registered in this data base and will form the basis for making a match between requests and offers.

- 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, which is available to researchers on loan for testing/evaluation purposes. For from systems from EU-sponsored projects, entry of the results into the repository could be made mandatory, for other research, strategies should be developed for providing incentives for providing information about systems as a physical reference 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.

5.7 Ethical considerations

Deploying capable, pervasive NeuroIT system within human society poses risks of failure and misuse. While possibly not urgent in the basic research stage, ultimately strategies will have to be developed to make such systems failsafe and limit the opportunities for misuse. Suitable concepts for tackling these issues may again be inspired by nature, looking at natural mechanisms for fault detection and repair.

Adequate NeuroIT periphery may be a remedy for ethical obligations towards conscious artifacts: Such artificial beings should not be placed in bodies ‘handicapped’ by insufficient periphery, if they are aware of and capable of suffering from these handicaps.

Chapter 6

Machine consciousness

6.1 Summary

The last ten 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. It is this second approach which inspires the Grand Challenge. The assumption is that subjective experience requires a subject, and that subjects require a physical instantiation. 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 they have many limitations. Current control strategies are inadequate for complex tasks; it is very hard to adapt robots for tasks which deviate from the functions for which they were designed – or to 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 the Grand Challenge is that to build machines with real autonomy and adaptivity and a genuine capacity to interoperate with human beings, we 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 to 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. Last but not least, consciousness will reduce the burden on programmers. Current machines have to be carefully programmed to accomplish their tasks. By contrast, animal brains and bodies *adapt* to different environmental conditions, during their development. There are many reasons to believe that this ’behavioural plasticity’ requires (and produces) consciousness.

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 the subject? In the first place, the dramatic increase of computer power over last decades has not solved all our problems, on the contrary, all the motivations given above are just as valid today as they were in the 1980's. If anything, they have forced us to admit that more of the 'brute force' approach will not help.

In the second place, we have a much more refined image of what goes on in the mind. Advances in non-invasive imaging techniques, but also a more refined application of single-electrode measurements are starting to reveal the neural substrate of processes that cognitive psychologists until recently could only describe in very abstract and generic terms. As a consequence, it has now become possible to relate subjective notions as 'awareness', to behavioural performance on the one hand and to specific neuronal mechanisms on the other. One of these neuronal mechanisms, attention, has several instantiations which sometimes have been described in great detail.

Attention plays an important role in autonomous agents: it helps to select those parts of sensory input which are behaviourally relevant, suppressing the rest. Moreover, many researchers has given compelling computational reasons for the existence of attention.

Consciousness is a more complex phenomenon than just attention, although attention plays an important role in consciousness. It is likely that in the next few years a taxonomy will be established which is as precise as the one which already exists for some forms of attention. Already several attempts have been made in this direction and although the results are very controversial, these attempts to define consciousness can be examined, critized, refined and revised in a way which distinguishes them from earlier philosophical ones.

As understanding the brain from a computational framework progresses it shows that what is called 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 trasferred into the machine domain in order to achieve a similar competence. A study of this transfer may be called Machine Consciousness.

Such a study will require *systematic studies* to analyze the principles of action of biological systems at the signal-processing and conceptualisation level, 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, creating a model of the world, a model of itself and mechanisms to exploit these models? In proposing a response, we have to bear in mind that human cognitive architectures are solutions to the problem of gene transmission and that 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 (an interesting opportunity for the consumer market); *intelligent 'situated artificial communicators'*, e.g. 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'). Many of the applications are unimaginable today. It is nonetheless clear that they will have a major impact in many different areas. Early products are likely to focus on the enhancement of 'classical applications' such as industrial robotics and driverless transport systems, service robotics (e.g. mobile service guides for customers in a supermarket), and edutainment.

6.2 Motivation and Objectives

Over the last ten years, there has been rapidly growing interest in the scientific study of consciousness (Tani, 1998; Taylor, 1999; Jennings, 2000; Buttazzo, 2001; O' Regan & Noe, 2001; Ziemke, 2001; Zlatev, 2001; Perruchet & Vinter, 2002; Rees & G. Kreiman, 2002; Taylor, 2002b; Crick & Koch, 2003; Gallese & Metzinger, 2003; Harnad, 2003; Zeki, 2003). The Tucson 'Towards a Science of Consciousness' Conferences (1996-2002) helped to create the right climate, playing a role similar to that of the Macy Conferences on Cybernetics (1946-1953), which prepared the ground 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 artifacts 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 artifacts 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. The underlying assumption is that subjective experience requires a subject, and that subjects require a physical instantiation. In brief, the scientific study of consciousness should be based on the investigation of physical systems. Its final goal should be the construction of conscious machines.

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 a broad range of useful task. They can transport materials and tools in factories, provide delivery services in hospitals, perform household cleaning chores, undertake underwater inspections. Yet despite these successes, achievements have fallen short of expectations. There are three main reasons.

- *Inadequate control strategies.* Current control strategies are inadequate for complex tasks.[1] It is not clear whether we should be using complex software intensive controllers or whether we should adopt alternative strategies.
- *Poor adaptivity.* It is very hard to adapt robots, or indeed any complex technical system, to tasks which deviate from the functions for which they were designed – or to changes in the environment. In many case, the adaptation has to be carried out by the robot manufacturer.
- *Poor communications with humans.* We lack effective techniques allowing robots to communicate and cooperate with humans and other complex systems.

---

[1] i.e. they are unable to provide effective real-time control, using a technically 'reasonable' quantity of computational resources.

Because of these failings, many people think robots are ”dumb” and do not believe they can provide useful services. To change this perception, we need autonomous, adaptive machines. The premise underlying ’Conscious Machines’ is that machines with real autonomy and adaptivity and a genuine capacity to interoperate with human beings, will require *consciousness* – a cognitive architecture that includes reflective control mechanisms akin to human introspection.

Consciousness can help machines to select the sensory information most relevant to their goal and motivations, reducing the amount of sensory-motor information needed to represent specific events. And it can 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 machine should possess this same self-teaching capability. Last but not least, consciousness will reduce the burden on programmers. Current machines have to be carefully programmemed to accomplish their tasks. By contrast, animal brains and bodies *adapt* to different environmental conditions during their development. At least in mammals, behavioural plasticity is enormously enhanced by the ability to develop individualised cognitive capabilities. It is plausible that this behavioural plasticity and consciousness develop from the same architectural foundations. This implies, not only that consciousness is the product of behavioural plasticity, but that behavioural plasticity actually *requires* consciousness.

Some of those who study machine consciousness lean towards the ’access’ side where action and behaviour is 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 that is appropriate for the machine being studied.

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. In declarative learning systems or plant-wide optimising controllers, this is obvious. But, in reality , self-awareness mechanisms are present even in the smallest and purest ICT systems: from power aware computing to adaptive, reflective middleware or feedback schedulers for high performance operating systems.

- *Trust.* Before we trust a complex software system we require *justification*: the system has to be capable of explaining what it is doing.[2] For ’justification’ the system needs introspection: access to its own ’thought processes’.

- *Robustness.* In many cases, the goal of adaptation is not to optimise system performance but to embrace change and provide increased resilience. 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 introspection –self-awareness– to identify behavioural mismatches, to diagnose their own mistakes (self-diagnosis) and to repair software and hardware breakdowns (self-healing)

---

[2] This was a major issue for Expert Systems during the 1980s.

- *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“. The only way to reduce personnel costs for complex ICT infrastructures is to make systems self-aware.

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

6.3 Examples of Applications

Conscious machines could take many different forms. For example:

- *Self organizing 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’*, e.g. 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 this 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.

- *Enhanced classical applications.* The implementation of self-awareness can lead to major improvement in industrial robotics (improvement of programmeming interfaces through the integration of images and language; learning of complex action sequences, e.g. in factory assembly), driverless transport systems (simplification of task specification), prosthetics (adaptation to variable environmental conditions), etc. The main marketing problem will be the creation of attractive laboratory prototypes, acceptable to potential customers.
- *Adaptive service robots.* Here there are many possible applications - 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. The artificial porter would register a passenger's desired destination (e.g. " the PanAm flight to San Francisco") through a multimode interface and then follows the passenger to the designated destination, continuously adapting its behaviour as it goes along. Another example might be a traveling 'sales guide' in a large supermarket. The guide registers what customers want to buy, *leads* them to the right shelf and provides any additional information they may need etc.. All the time it is doing this, it observes and adapts to customer behaviour. When customers walk slowly it slows down; if they want to stop and look, the guide stops with them. Advanced versions of the guide might be able to actually take the product from the shelf – a useful service for elderly customers who do not always want to bend down.
- *Edutainment.* Conscious machines could be used in animated displays in science centres and theme parks. They could be interactive partners for humans, at play, in the film industry, or in theatre. Last but not least, they could explain and demonstrate their own developmental principles, achieving a degree of interactivity which would never be possible for software on a computer.
- *Complex Technical Systems.* Mechanisms of self-awareness can be used to control complex technical systems, such as chemical plants, or electrical power grids. The systems monitor their own performance. Supervised learning can help the system to generalize from situations which were labeled by the supervisor, thereby enabling the system to learn from past experience. In situations where there are no saftey issues, exploration strategies could be allowed, which would help the system to learn about its capabilities to perform its tasks and possibly to improve beyond its initial programming.

But in general terms, there is one class of machine that could leverage 'machine consciousness' more than any other. This is *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, etc. In all of these cases – and many more – the machine's ability to exploit deep representations of itself, (body and mind), the task and the environment will have an enormous impact on performance and dependability.

6.4 A business case for Consciousness Machines

It may sound strange to claim for the existence of a business case for conscious machines when there is even disagreement on the role that consciousness plays in natural systems and its evolutionary value. This is clearly shown in the fact that there is a school of thought that claims that consciousness is an epiphenomenon, i.e. nothing we can't live without.

Obviously, we don't think that way, and the best proof of its evolutionary value is our own everyday perception of consciousness: What do you prefer? a conscious or an unconscious taxi driver? What do you prefer? a conscious or an unconscious neurosurgeon? It's pretty clear that consciousness do play an important role in the correct execution of tasks, in the exercising of adequate behaviour in the presence of uncertainty.

But beyond exploring ideas on what consciousness is and how can it be present in an artificial agent we may consider the question of real needs for this technology. Is there any real business case for it?

Indeed there is. 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; our laptops, PDAs and mobile phones are full of software and communications because today's computing environment is changing from the isolated isles of productivity around a productivity toolset to the open sea of web services and dynamic applications. There is no longer a central deity that decides when to release a new complete update of our suite. There's nobody in charge of our whole environment anymore. We must strive for keeping our working environment in line with evolving realities out there. And the task is not easy at all: 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 sand strategies. Cross-border interconnection seems necessary from many points of view (e.g. robustness, efficiency, policies, etc.). But, from a purely technical point of view, the task of controlling such a system is hopeless. There is nobody in charge anymore. Subsystems are built using different technologies and operated under different regulation agencies. While the technical issues may be solved relatively easily (standardisation bodies do help in this) the main problem remains: integrated, unified decision making. The local operation decisions can be contradictory at a system-wide scale. Autonomous distributed decision making processes mine the technically sound operation of the global network. 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. We are engineering systems that suffer butterfly effects. That old wildly free character of some natural realities is becoming a daunting fact of our infrastructures.

What do these two cases have in common? The answer is relatively easy to identify: the behaviour of the global system is driven by interaction of local systems that are no longer under a common change authority. It may look like the problem is that of proper physical integration but it is more than just that. The main issue, the really daunting thing, is that the bottleneck, or to be more precise, the key of the path to the solution is the capability of cognitive level metareasoning and integration. 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 that 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.

6.5 Current Technology and Knowledge

This section draws heavily on (Aleksander & Dunmall, 2003) In the near future, even hardened sceptics expect to hear claims that non-biological machines have attained conosciousness. On the whole this is good news. Finally, it is becoming possible to have a serious discussion about the object of consciousness. Why, for instance, should a bat or any other organism be conscious? what exactly might a bat be conscious of? how does the bat's consciousness differ from my own? Last but not least, 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 (see the Appendix to this chapter). Prejudices are being replaced by logical argument. And although there is a long way to go, there has been real progress, as seen in the 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'. Many laboratories are working to establish an common engineering background for an attack on consciousness (McCarthy, 1995; Grossberg, 1999; Manzotti & Tagliasco, 2001; Perruchet & Vinter, 2002; Taylor, 2002b; Haikonen, 2003; Harnad, 2003; Manzotti & Glatzeder, 2003; Aleksander & Dunmall, 2003).

An example of the kind of work currently in progress is research by Nobel Laureate, Gerald Edelman – who is very much a sceptic with regard to traditional Artificial Intelligence. Edelman's goal is 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 that the survival of biological organisms depends on their 'values': for instance 'light is better than dark' or 'fear is to be avoided'. Some values may be inborn, others may be acquired during development. In both cases, they are encoded in neural circuits in the organism's brain, biasing its behaviour to favor 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. During its interaction with the environment the robot cleverly learns that small bright objects are important and should be touched with the tip of the arm (finger). Is Darwin conscious? Of course not. 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 this space, incoming sensory information activates dormant memories - making the organism conscious of the input. Memory fragments activated in this way stuimulate the organism to look for more input, activating additional memories and so on (Baars, 1997). Interestingly, this somewhat abstract idea has been used by Stan Franklin of the University of Memphis to build a working system for the US Navy (Franklin, 2003). The system is called IDA (Intelligent Distributed Agent) and is used to exchange email with sailors to arrange their five-yearly job rotation cycle. Tests have shown that the sailors interact with IDA in much the same way as they would with a conscious human. So IDA can pass a kind of Turing test: users are not sure whether they are communicating with a conscious human being or a machine. In other words, Franklin, has used the idea of a Global Workspace to produce a machine that functions in the same way as living organisms we generally take to be conscious. We could say it is *functionally conscious*. This is the old story that 'If it looks like a duck, it quacks like a duck, and swims like a duck, it may as well be a duck'. 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.

The builders of IDA have no misapprehensions on this point. A *materially conscious* machine would have an identifiable mechanism responsible for its sensations. IDA, on the other hand, has no such model, though we can certainly find processes acting *as if* the machine had sensations.

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 us 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. Thhe 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. Dubbed 'axioms' (as they are intuited but not proven) they are:

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 implemntation 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."

6.6 Future Research

Achieving the goals of 'Conscious Machines' will require *systematic studies* to analyze the principles of action of biological systems at the signal-processing and conceptualization level, enabling the transfer of these mechanisms to artificial systems. Key issues to be investigated will include:

- *Attention* play an essential role in consciousness. Although an understanding of attention by itself is not sufficient for understanding consciousness, most researchers agree its role in selecting behaviourally relevant and suppressing irrelvant information is an absolute prerequisite for humanly simple, but computationally complex behaviour.

- *The coordinated development of sensory systems, cognitive activities, and effector capabilities* during the lifetime of the artificial system and/or the evolution of these systems over multiple generations (the equivalent to biological evolution);
- *The control and/or exploitation of the growth dynamics* of the sensorial system and external structures (morphology). Instead of full coding (specification) of the growth process, it might be possible to specify *dispositions,* using self-organization to optimise the physical structures required to perform specific tasks;
- *The creation of cognitive mechanisms* allowing the system to perceive and exploit the structure in different contexts. Such mechanisms may include attention control, complexity reduction, category formation, concept learning, object naming (e.g., through imitation), and transformation of knowledge into forms intelligible to humans;
- *The development of suitable substrates/materials and implementation technologies* (analogue, FPGA, hybrid, and biological hardware); In particular the possibility to emulate mechanisms (attentional and otherwise) in real time implementations of 'conscious machines' using suitable hardware (e.g. spiking neurons) is an avenue to be explored (e.g. see Taylor's CHIMERA chip (Taylor, 2004)).
- *Social and communicative interaction* among artificial systems and between artificial machines and humans; the development of dynamic ontologies based on these interactions.

The knowledge acquired from these studies will help to answer a number of key research questions:

- 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 to 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 human cognitive architectures are solutions to the problem of gene transmission and that artificial systems should have very different user-determined goals. To quote McCarthy, "The useful forms of computer agent selfawareness 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."

6.7 Relation to other Roadmap chapters

Clearly, there are strong interdepencies between this chapter and other chapters of the Roadmap. 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 of the Roadmap.

Several approaches are possible, depending on personal interest and expectations of return on investment. An idealized 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 our understanding of these principles, we implement them in hard, in technical systems, which should function precisely in those situations were adaptive, flexible, intelligent behaviour is required.

There is nothing wrong with approach, and 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?

In the second place, 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 attenional structures and cognitive architectures, such as, for example, CODAM, and EU project ALAVLSI, or various vision projects which take inspiration from visual cortex, suggest that 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. The research which is described in 'constructed brain' (chapter 8), tools for neuroscience (chapter 9). For an understanding of the fundamental abstract principles of natural computing, assuming they exist, modelling and understanding of dynamical systems is important and also the development of mathematics which allows an economical expression of these principles.

Research that would take an ’inspired approach’ may roughly be divided into two categories. It is to be expected that 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 organization 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 chapter 2) 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 approch, ’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 by building ever more complex technical systems, which are 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) one may discover machine equivalents of awareness. Research that is relevant for such an approach might be Factor-X (chapter 4), the chapter on peripheral processing (chapter 5) and the chapters on bio-inspired hardware (chapter 3 and biological complexity (chapter 7).

6.8 Appendix: Philosophy, Psychology, Neuroscience and consciousness

6.8.1 Philosophy and Consciousness

The philosophy of mind - from Socrates and Plato to the brain-based approaches of some modern philosophers, is too vast to be summarised in a few lines. Here we will limit our discussion to two ’strands’ of thought: Western cognitive science, as discussed in (Crick & Koch, 1990, 1998, 2003; Crick, 1994; Damasio, 1994; Dennett, 1991), and what could be called ’the other side of consciousness’.

6.8.1.1 Western Cognitive Science

For many decades, Western cognitive science rejected the idea of a ’Ghost in the Machine’ (Ryle, 1952). From a scientific point of view, the only ”observables” were bodily functions, including the subtle activities of the brain. This left no room for a distinct, inner or mental world (the ghost in question). (Cotterill, 1989; Crick, 1994; Crick & Koch, 2003; Damasio, 1994; Dennett, 1991) . Conflicting claims and theories could all be traced back to the same shared doctrine, namely, that consciousness is nothing other than ’intentionality’. As the philosopher Brentano put it,: “all is content in consciousness; there is nothing else”. There can only be ’consciousness of’ something; ’consciousness’ in itself does not exist. This matches the views of the David Hume, the British empiricist. When Hume examined his ’inner self’, he could find ’nothing other than a bundle of perceptions' . For Western cognitive science and the authors cited earlier, this bundle of perceptions *is* consciousness. Anything else would smack of dualism. So all we can really study is the content of consciousness.

This approach leads to a deep and basic problem: if it is correct, where do we get our sense of self, our feeling of what 'what it is like to be' a certain kind of organism? Some authors deny that this is an issue (Crick, 1994; Crick & Koch, 2003; Damasio, 1994; Edelman, 1992) or that it is relevant (Dennett, 1991). Others realise there is a problem (Chalmers, 1996; Levine, 1983; Nagel, 1974) and try to resolve it with 'dualistic' explanations involving non-material 'stuff' such as information (Chalmers, 1996). But even if they accept the 'explanatory gap' (Levine, 1983), the challenge of consciousness, and the reality of 'the what it is like to be' problem (Nagel, 1974), they are shackled by their denial that consciousness can be anything more than the content of consciousness. Under these circumstances, the best they can do is add extra elements to those of the material world – becoming dualists in all but name. Researchers in the 'consciousness explained' camp never really bite the bullet. Either they claim that there is no inner self at all (Crick & Koch, 2003; Dennett, 2003) , or they are left floundering as to what is still missing.

An alternative would be to adopt the view put forward by American philosopher Block (Block, 1995)'s In Block's view, consciousness has two parts: *phenomenal consciousness* (the 'what it is like to be' component) and *access consciousness* (the component 'poised for rational action and report'). Block's view matches recent findings on attention and consciousness. Brain studies have shown that it is possible to dissassociate *sensory attention* and *motor attention.* This suggests that each produces its own form of consciousness. If this is so, Block's phenomenal consciousness could be the result of an emerging representation of a *sensory attended-state* – created by attention control mechanisms within the organism. These mechanisms would ensure that this state is created as rapidly and reliably as possible – giving the organism the experience of being *centred* and generating the irreversible sensation that it *owns* the content of its consciousness.In this account, the role of access consciousness – Block's second component – would be to broadcast the phenomenal content of consciousness to other brain components, allowing them to prepare for, or carry out, actions to achieve goals. This is achieved via *motor attention*.

6.8.1.2 The other side of consciousness

When addressing consciousness, cognitive science pointedly ignores the wealth of insights offered by Western phenomenology and Eastern meditation. Again, it is hard to summarise what amounts to a huge body of thinking. But the important debates cited in (Gallagher & (Editors), 1999) deserve at least a few brief comments. From a phenomenological study of his own consciousness, Strawson (1999) reaches the conclusion, that “there are eruptions of consciousness from a substrate of apparent non-consciousness” (p. 21). (Hayward, 1999) cites his experience of meditation: “Between moments of experience of self are gaps in which there is no sense of self or of separateness from what is being experienced” (p. 390). Both authors point to the 'gappy' nature of the inner experience .

Gallagher (1999) makes the point even more strongly. He discusses Kant's demonstration that the self cannot be experienced or even defined in terms of empirical qualitities, going on to argue that this view matches descriptions from Eastern writings, in which training for meditation leads ultimate to what is termed a 'pure' state of consciousness. To cite the words of a Japanese Zen mystic, this Pure Consciousness Experience or PCE , "....is not vacant emptiness. Rather it is the purest condition of our existence" (Gallagher, 1999, p. 413).

According to ancient Indian writings (Galllagher, 1999, p. 413) PCE is : "...unseen, incapable of being spoken, ungraspable, without any distinctive marks, unthinkable, unnameable... He who knows it thus enters the self with his self".

This aspect of conscious experience, which has so much experiential evidence in its favor, should be a natural component in any model purporting to provide an understanding of consciousness.

6.8.2 Psychology and Consciousness

6.8.2.1 General Features

Psychology uses consciousness all the time. This is what it is happening when researchers ask experimental subjects to describe their subjective experience when they are exposed to particular stimuli or report what they remember when the stimuli are no longer present. In other words, consciousness is the pre-condition for any psychological experiment, even when it is not actually trying to probe consciousness. A number of experiments have attempted to dissociate unconscious from conscious processing. In experiments on subliminal processing, for example, stimuli are presented for so short a time, the subject is unaware they are present. Subsequent tests determine how far the subliminal stimulus has been processed. Techniques such as attentional blink allow controlled dissociation of unconscious and conscious levels of processing. The results suggest that before the subject is aware that they are present, stimuli can be processed right up to the semantic level, .

6.8.2.2 Psychological Models of Consciousness

Psychologists have proposed numerous models of consciousness, though few have make any attempt to relate their models the structure of the brain. In what follows we describe two exceptions.

- *The Global Workspace*
  The Global Workspace (GW) is a psychological theory emphasizing particular aspects of consciousness (Baars, 1997). It is based on the computer science concept of a global workspace. According to the theory, the global work space provides shared storage for material which needs to be processed by several different processors. This processing, it is claimed, produces consciousness. The GW approach has been very fruitful in bringing together ideas from psychology and computer science. Nonetheless it leaves a number of issues unexplained. Despite attempts to match the theory to results from brain science (Baars, 2002), GW has nothing to say about recent imaging studies, which suggest that different executive and higher level functions are performed in separate areas of the brain (Corbetta & Shuman, 2002). More specifically, the prefrontal cortex does not seem to behave as the kind of 'shared storage' predicted by GW. On the contrary, its main components seem to be executive processors, each of which performs its own highly specific task,

maximizing the reward to the organism. The specific prefrontal brain circuitry responsible for these functions has yet to be identified. Another issue that Global Workspace is unable to explain is the nature of 'inner experience'.

- *The Relational Mind*
  A second incomplete theory emphasises the role of memory in creating the content of consciousness (Taylor, 1999). The theory suggests that the meanings attached to specific internal content depend to a large extent on their relationship to other content. In a sense this must be so. If a subject has no experience with a given object it can acquire no meaning. Here again, the theory emphasises one aspect of experience, to the detriment of many others, especially the inner self.

Taylor (2002b, 2002a) gives a detailed discussion of how a Pure Conscious Experience (see also section 6.8.1.2) can arise from attention attending to itself (through the the CODAM model). Taylor suggest that such explanations begin to allow a brain-based explanation of what some people might call a soul.

6.8.3 Neuroscience and Consciousness

6.8.3.1 New techniques

With the advent of brain imaging technology such as PET, fMRI, EEG and MEG, neuroscience has made great strides forward. The most important recent advances have come from research strategies which combine these techniques. Other important results have come from single cell recordings from animals (mainly monkeys) and in a few cases from intercranial EEG or deep electrode penetration in human patients.

Using these tools, neuroscientists are discovering the network of areas involved in conscious and unconscious processes of sensory information. Experiments show that some regions are active even when awareness is lacking, e.g. in semantic level activations during the attentional blink or during the presentation of subliminal stimuli. Other regions are inactive in these conditions. The lack of activity is particularly striking in sites associated with P2 and P3 ERP waves, about 200 and 300-500 ms after stimulus onset.

All this suggests that specific brain regions are responsible for elevating neural activity from subliminal or unconscious processing to awareness. The same sites, in particular the posterior buffer sites in parietal and medial temporal cortex, seem to be involved in working memory. By contrast, prefrontal executive sites – also involved in working memory – are not essential for awareness. Patients with pre-frontal lesions are fully conscious, even though they suffer from behaviour deficits and poor decision-making.

The findings just described have inspired a number of models. Below we review three of these.

6.8.3.2 Quantum approaches

Quantum models of consciousness suggest that that consciousness depends critically on quantum effects such as wave function coherence, collapse of the wave function or quantum gravity. Such effects, it is suggested, spread coherently through the brain thanks to sub-neuronal processes. According to Roger Penrose and Stuart Hameroff, this involves a Planck scale “decoherence of quantum superpositions” (Penrose, 1994), “... shearing off into separate, multiple space-time universes as described in the Everett ’multiworlds’ view of quantum theory...”.Hameroff (1997) They suggest a critical role for so-called ’microtubules’. These are protein structures internal to the axon, soma and dendrites of the neuron. The main known function of ’microtubules is transport of material through the cell. But Penrose and Hameroff argue that their diameter – barely 20 nm – is sufficiently small to support a coherent quantum state (Hameroff, 1997). This state could be transmitted from one neuron to the next, without loss of quantum coherence, through gap junctions (perhaps 10-15% of synapses connecting neurons in the brain). Finally, the quantum state would lose its coherence via a Planck-scale decoherence process.

Since Penrose and Mameroff’s work was published, several authors have proposed extensions of their idea, exploiting the properties of superstrings or through quantum electrodynamics. These models predict different patterns of temporal and spatial correlation of brain activity during conscious and unconoscious processing. A number of studies are trying to match these predictions against observed brain activity. But regardless of the results of these studies, quantum models have very little to say about the ’inner self’.

6.8.3.3 40-Hz and gamma synchronization

For a ten year period which ended only a few years ago, ”40-Hz” was the favorite brain-based explanation of consciousness . Francis Crick (Crick & Koch, 1990) suggested it could be the basic mechanism, binding different features across different regions of cortex. Experimental work has found that 40 Hz does indeed play a role in binding, as well as in attention. . This supports the idea that 40 Hz is *necessary* for consciousness. But the experiments do not support the claim that 40 Hz is *sufficient* for consciousness. Indeed the original evidence for 40 Hz’s role in binding was obtained on anaesthetised cats, which were not conscious at all. Neither of the authors of these experiments has backed the claim identifying consciousness with the 40 Hz signal (Crick, 1994).

The tie to other features of consciousness is even less clear. Synchronisation of brain activity may help to explain the origins of intentionality, through feature binding for object experience, but an explanation of conscious experience has to explain more than just the binding of content. In particular, conscious experience has an *owner*. It is the owner who has the inner experience of ’what it is like to be’ the owner. Synchronisation on its own is not enough to explain this experience. A full explanation will have to take account of other aspects of brain processing. More generally, synchronisation cannot explain how the self exerts control over brain activity. Every component of global synchronised neural activity is equipotent. Explaining specific forms of control forces us to consider other functional components .

In conclusion, synchronisation undoubtedly plays an important component of brain processing, and especially in the creation of consciousness through binding. It must thus play a role in any final theory. But the final theory will also include many other components.

6.8.3.4 Dynamical systems and complexity

To some extent, the dynamical system and complexity approach to consciousness has been a bandwagon similar to the 40 Hz endeavor. Many groups have followed general dynamical systems approaches, and on occasions their work has overlapped with research by the 40 Hz school. The same groups have dedicated a major research effort to the role of chaos and of complexity in brain processing, examining this prospect from an experimental and a theoretical perspective (Freeman, 1975, 2000).

At the most basic level, it is clear that a general dynamical system analysis of brain activity plays a complementary role to detailed experimental data. The equations governing the dynamics of interacting neurons, or those of their synapses and of NMDA, AMPA and GABA channels, are mathematically complex. Concepts such as attractors, basins of attraction, bifurcation, fixed points, and stability analysis could well provide new insights. What is not clear, from a dynamical systems viewpoint, is how the neural dynamics underlying conscious activity (the neural correlates of consciousness or NCC) differ from those of unconscious states. Experiments show that some brain regions are more active during conscious than unconscious processing. This helps to identify the NCC. But it cannot explain how this dynamic activity leads to conscious experience, or identify a set of dynamical systems components sufficient for consciousness. To do this we need a theory which assigns specific functions to specific brain sites and dynamical components.

Despite the absence of such a theory, there can be little doubt that dynamical systems theory will play an important role in future models. A number of detailed proposals have suggested how dynamical mechanisms, such as strange attractors, could lead to consciousness (Freeman, 2000). In these theories, the brain follows a trajectory across a dynamical landscape occupied by strange attractors. Shifts between attractors correspond to rapid transitions in brain state, similar to those often observed in experiments. This kind of concept, together with detailed simulations, is a useful tool for understanding how the brain processes information. What is less clear is the role of this kind of dynamic process in supporting consciousness.

Edelman and colleagues have examined brain activity from the point of view of complexity theory, examining whether there exist specific regions of the brain with higher connectivity during conscious than unconscious processing (Edelman & Tononi, 1998). The results suggest such regions exist. But although some of Edelman's earlier work produced suggestions for modelling (Edelman, 1992), a specific model of consciousness has yet to emerge. This is another area of research which will play an important role in a final theory of consciousness, even if cannot build such a theory on its own.

6.8.3.5 Centre of narrative gravity

In his provocatively titled 'Consciousness Explained' Dennett (1991) argues that consciousness is created as a 'centre of narrative gravity', bringing together narratives created in real time, as the brain processes information in multiple working memory areas. For Dennett there no Cartesian theatre, where consciousness is finally created.

At a general level, there is much to support this picture. But it is not clear how the ”centre of narrative gravity” generates the known features of consciousness. Dennett is not worried. In his view, the ”centre of narrative gravity” explains everything about consciousness we can actually observe . If there is no place in the brain where consciousness ’all comes together’, there is no need for a first person perspective. The objection to this is that the first person perspective *does* exist: our inner experience tells us so. But all need not be lost. Maybe the first person perspective comes from some specific characteristic of the brain, such as inter-module connectivity for attention control. In sum, Dennett's is another partial theory – with gaps waiting to be filled. One of these is its lack of an explanation for attention and other higher level functions.

6.8.3.6 CODAM

CODAM (COrollary Discharge of Attention Movement) has been considered a very important component in models of motor control by the brain. (Taylor, 2005) extends the concept to propose an engineering control model of attention, using efference copy. Corollary discharge is a crucial component of this process. In Taylor's model, efference copy creates a sense of ownership of the about-to-be-experienced content of consciousness. Simulations of the model have been used to successfully explain the major components of the attention blink (N. Fragopanagos & Taylor, 2005). It has since been extended to include rewards, and emotions generated by the limbic system (Taylor & Fragopanagos, 2005).

6.8.3.7 The elements of consciousness

Aleksander and Dunmall (2003), have argued that consciousness may be decomposed into perception, imagination, attention, planning and emotional evaluation of plans. They have shown through digital neuromodelling that these elements are supported by specific neural architectures, both in the brain and in models. These can be modelled separately but as part of an overarching architecture. This work has been used to produce a comprehensive model of attention-driven visual awareness which has been implemented both in virtual and real robots and used to understand pathological situations such as vision deficits in Parkinson's disease.

Chapter 7

Biologically inspired techniques for the automated design of Artificial Cognitive Systems [1]

Glendower: *I can call spirits from the vasty deep*
Hotspur: *Why, so can I, or so can any man; but will they come when you do call for them?*
W. Shakespeare, Henry IV Part I

7.1 Introduction

It has long been recognized that traditional Artificial Intelligence (AI) has not lived up to its initial promise: systems that performed impressively in ”toy worlds” failed to scale up to real-world applications (Dreyfuss, 1972). Today, similar problems face so-called new AI. While biologically inspired approaches such as neural networks (Anderson, 1995), evolutionary computing (Mitchell, 1996), and evolutionary robotics (Nolfi & Floreano, 2000) have produced interesting results, applications (such as OCR and data mining applications of neural networks) have been highly specialized and have 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 where cognitive sophistication is more important than number-crunching, artificial systems have yet to come close to the routine performance of humans and even relatively simple animals. The fact that biological organisms successfully perform sophisticated cognitive functions is proof that, at least in theory, there exist solutions to the problems we are trying to solve. Molecular and neuroscience investigations of real-life systems suggest, however, that the mechanisms required will be highly complex - in Chaitin’s technical sense that they cannot be compressed into a compact piece of code (Chaitin, 1975). In other words, if we want to implement artificial systems that emulate human and animal cognitive competencies - as implied by the other grand challenges

---

[1] The idea that the development of naturalistic models for Artificial Evolution represents a ’Grand Challenge’ had been independently suggested in (Kendal et al. 2002)

in the Road Map - we will have to implement very complex algorithms. Current techniques in software engineering are almost certainly inadequate to this task.

Paleontology has shown that nature can, when required, evolve complex structures and behaviors in what, in geological terms, are very short periods of time. Examples include the evolution of nearly all currently extant animal baupläne during the so-called Cambrian explosion (565-525 million years ago) and 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. For many years researchers in Artificial Neural Networks and in Evolutionary Computing[2] have attempted follow precisely this strategy. Much of this work, however, has used models, so highly stylized as to be almost unrecognizable to biologists. At the same time, whole areas of evolutionary biology (e.g. evo-devo, gene regulation and expression in the evolution of novel phenotypes, multi-level evolution, the evolutionary role of structural and mechanical constraints, the evolution of evolvability, the roles of gene duplication, redundancy and degeneracy) have yet to be adequately mined by engineers. In brief, current attempts at biologically-inspired design have only scratched the surface of what might be possible.

7.2 Objectives

The strategic goal of the Grand Challenge is to contribute to the long term goals of traditional and 'new' AI: namely the production of artificial cognitive systems capable of performing useful tasks effectively performed by human beings and other animals. The ability to design and build such systems is a pre-requisite for responding to the other Grand Challenges described in the Roadmap. 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 behaviors. In short, we have to construct **a theory of the evolution of complex systems in nature and** we have to apply the theory to generate **biologically inspired techniques for the automated design of Artificial Cognitive systems.**

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), identifying models capable of performing in low-order polynomial time and characterizing tractable and intractable problem sets;
4. design and implement tools for the automated design of Artificial Cognitive Systems based on (possibly simplified) versions of these models;

---

[2] In this proposal, 'Evolutionary computing' is used as an 'umbrella term' describing a range of different techniques, broadly inspired by Darwinian natural selection. These include (but are not limited to) Evolutionary Strategies, Genetic Programming, and Genetic Algorithms

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

7.3 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 realizations include:

1. 'elastic' designs for autonomic 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, 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-computerized development environments, for the artificial evolution of complex 'wetware' e.g. bio-electronic sensor, actuator and cognitive systems

7.4 Current state of technology

The concept of biologically-inspired automated design is not new. **Artificial Neural Networks**, first conceived in the 1940s (McCullough & Pitts, 1943), attempted to model animal and human learning as an alternative to explicit design; **Evolutionary Computing**, originating with Rechenberg's Evolutionary Strategies of the 1960s (Rechenberg, 1965) adopted an alternative approach, drawing inspiration from Darwinian models of evolution in **populations of organisms**. More recent work, in the 1990s and the 2000s, has moved beyond the symbolic world of computers towards the automated design of physical, chemical and biological artefacts: electronic circuitry, autonomic robots, and even 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 it is not necessary to explicitly program an ANN to perform desirable tasks: ’learning’ algorithms (e.g. **back-propagation**), allow human operators to ’train’ networks to recognize patterns belonging to specific categories. 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 (CRM), and medical applications. ANNs are used intensively in **evolutionary robotics** (see below)

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. As in biological evolution, ’copy errors’ during reproduction create new variants of parent models. Multiple iterations of the selection-reproduction and variation cycle, produce a population of organisms with the ability to effectively perform the task assigned by the operator .Popular approaches to evolutionary computing include **Evolutionary Strategies, Genetic Algorithms and Genetic Programming.**

- In Evolutionary Strategies (Rechenberg, 1965; Schwefel, 1981) 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, in which a (small) random value is added to one of the elements in the vector. Applications of Evolutionary strategies have focussed on hard optimization problems, including the design of aerodynamic surfaces, the design of road networks and the solution of the vehicle routing problem.

- Genetic Algorithms, introduced in the 1970s (Holland, 1975), represent competing problem solutions as strings of bits - a rich and flexible coding scheme which allows a natural implementation of ’genetic operators’ such as point mutations, insertions, deletions, ’cross-over’ etc.. The literature includes 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 circuits.

- Genetic Programming, (Koza, 1992) adopts an alternative approach, in which the evolutionary process constructs a computer program, represented as a ’tree’ of functions and values, which, in many cases, are forced to respect a pre-defined set of grammatical constraints. Like biological evolution, Genetic Programming is open-ended: in theory, though not in practice, the programs constructed through this approach can attain arbitrary levels of complexity. Applications have included the automatic generation of programs to resolve complex problems in inverse kinetics, the automated development of control mechanisms for prosthetic limbs, and the discovery of classification rules for medical diagnosis (Bojarczuka, Lopesb, & Freitasc, 2001)

Recent research has moved into new areas. In particular a number of researchers have developed evolutionary methods of design which have moved beyond the purely symbolic and numerical representations typical of computer-based approaches. Areas of current interest include **Evolvable Hardware, Evolutionary Robotics, and Combinatorial Chemistry**.

Evolvable hardware (EHW) (also known as Evolutionary Electronics, EvolWare, bio-inspired electronics) (Thompson, 1996, 1997) combines computer-based techniques (automatic generation of random circuit configurations, generation of variant 'offspring') with physical tests designed to measure the performance of the resulting circuitry and to select the 'fittest' circuits for further rounds of evolution. Proponents of EHW suggest that "Evolved circuits can have a richer spatial structure and internal dynamics than normally envisaged, and can extract unusual leverage from the physics of their medium of implementation — be that microelectronics in simulation, physical silicon reconfigurable chips (FPGAs), or even proposed future technologies for nano-scale systems." (Thompson, 2004)

EHW's use of physical hardware is paralleled in **Evolutionary Robotics** (Nolfi & Floreano, 2000) where researchers again combine computer-based methods and physical testing (or detailed simulation of robot physics) to evolve not only robot control systems (often managed by **Artificial Neural Networks**) but also robot morphology (body plan, motor and sensor placing etc.) As in EHW 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 developed the ability to synthesize large numbers of different, compounds by simultaneously reacting a set of simpler components in thousands of different combinations and to screen the resulting **combinatorial libraries** for molecules with a desired chemical property (usually the ability to bind to a particular molecular receptor). This **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, electronics, photonics, advanced packaging, and self-assembled materials. A number of laboratories are working on methods in which single-step generation of combinatorial libraries is replaced by an iterative, evolutionary process, allowing the selective 'breeding' of molecules with desirable properties.

7.5 Problem areas

Each of the technologies, described in the previous section has produced impressive results in the laboratory. In the majority of cases, however, industrial applications have been highly specialized and of limited economic significance. The difficulty of moving from the laboratory into the field is at least partially due to a number of intrinsic weaknesses in current technology, which in many cases coincide with areas where artificial models have little resemblance to natural processes. Some of the weaknesses affecting current techniques of Artificial Evolution 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. ANNs, as a rule, consist of tens, hundreds or at most thousands of artificial neurons (compared to the $\approx 10^{12}$ neurons

in the human brain); Evolutionary Strategies, Genetic Algorithms, and Genetic Programming represent problem solutions with strings or vectors containing tens or hundreds of elements (compared, for example, to the $\approx 3*10^9$ bp in the human - and mouse (genomes); 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 for individual molecular receptors.

2. Experimental experience shows that the time required to ’train’ ANNs or to ’evolve’ other classes of evolutionary system, increases rapidly with the size of the problem to be resolved. In at least one case (the training of feed-forward ANNs to perform arbitrary tasks in pattern recognition) the training task has been shown to be NP-complete (J.S., 1990). The ’No free lunch’ theorem (Wolpert & Macready, 1997) demonstrates that for arbitrary tasks, Darwinian procedures of evolution are no more effective than random search. These results suggest that even rapid increases in computing power will have a limited impact on the size of the problems that can be attacked under current approaches.

3. In the majority of models (grammar-based approaches are again an exception, as is Combinatorial Chemistry) there is no differentiation among the sub-units. Such approaches fail to model the complex interactions among heterogeneous agents (molecules, genes, transcription factors, neurons, organs, organisms etc.) which are an essential characteristic of biological systems. As a result, there are no internal constraints on the way in which the CAS can develop during training or evolution. This makes for a much larger search space (relative to the size of the system) than the space explored by biological organisms during their evolution and development.

4. The majority of current techniques in ANNs Artificial Evolution (grammar-based approaches are an interesting exception) pre-define the basic size, architecture and (where relevant) morphology of the ’organism’ they are trying to train or evolve. Such techniques leave no room for the open-ended evolution which has characterized the development of life on earth.

5. Virtually all current ANN or evolutionary algorithms emulate a single period (or a very small number of periods) of ’training’ or ’evolution’. In nature, on the other hand, organisms and species acquire complex competencies and behaviors, though long sequences of learning and adaptation, in which later acquisitions build on the results of earlier phases. The failure to model the multi-stage nature of learning and evolution places critical limitations on the ability of artificial evolution to design genuinely complex systems.

6. In the majority of evolutionary models, there is no distinction between genotype and phenotype - and thus no room for artificial organisms to ’develop’. Rare attempts to develop ”growing networks” (Nolfi & Parisi, 1995) have had little influence on standard practice. The failure to distinguish between genotype and phenotype makes for highly inefficient coding schemes (in evolutionary models based on ANNs, for instance each individual synapse is represented in the code) that contrast with the coding efficiency of animal

genomes[3] . In biological elements genetic networks can code for growth processes which can involve very large numbers of differentiated cells. No equivalent process occurs in Artificial Evolution.

7. In line with the failure to model development, the majority of evolutionary models make no attempt to model the ”regulatory” elements which govern gene expression in biological organisms. The absence of regulatory elements, together with the failure to model development leaves little room for the developmental plasticity that characterizes biological organisms. At the same time it prevents macro-evolution through modification in regulatory code - a mechanism which many evolutionary biologists believe to be a defining feature of biological evolution.

8. In reality, current models of Artificial Evolution fail, not only to model the regulation of gene expression but to include any biological realistic mechanism for macro-evolution. Despite interesting laboratory investigations (Calabretta, Nolfi, Parisi, & Wagner, 2000) 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 (Ohno, 1970). Little attention has been paid to genetic redundancy (Walker & Miglino, 2002). The inability of current models to models macro-evolutionary mechanisms imply that what is being modelled is micro-evolution: the ’tuning’ of a pre-existing architecture to meet the specific requirements of a specific environment; there is no representation of the origins of the architecture or of major architectural change

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-organization providing ’order for free’ (Kauffman, 1993). Research into self-organizing systems (Wolfram 2002) often pays little attention to its potential role in evolution

10. Most evolutionary models make no provision for the transmission of information between individuals (for a model of how novel signalling schemes could evolve see Steels 1998). This theoretically precludes the emergence of the ’major evolutionary transitions’ identified in (Száthmary & Maynard Smith, 1995), which depend on the evolution of novel mechanisms of communication among organisms and which subsume the symbiotic route to biological complexity proposed in (Margulis, 1998)

11. The majority of ANN and evolutionary models are single level, making no attempt to model the hierarchical organization of real life biological systems into genes, gene networks, cells, organs, organisms, demes, species, higher-level taxonomic units (S.J., 2002). Evolutionary theory predicts that in the absence of this kind of hierarchical organization

---

[3] The whole human genome is only slightly larger, in terms of bits, than current ”office” software suites. The genome of E. Coli is no larger than an average utility

relationships between genetically unrelated systems will be purely competitive, precluding the emergence of integrated systems from smaller sub-units (Sober & Wilson, 1998).

As has been shown, current models of ANN and Artificial Evolution fail to represent key aspects of biological evolution and of biological organisms. It is important to realize, however, that many of the limitations of current models concern implementation rather than theory. There is no *theoretical* reason preventing models from achieving a higher degree of biological realism. In fact, as has been shown, nearly all of the issues described in the previous section have been explored by specific approaches to Artificial Evolution, or at least by individual research teams. The key factors that have so far hindered progress are partly cultural but primarily technical.

From a cultural viewpoint, there can be little doubt that studies of Artificial Evolution have suffered from inadequate communications between different disciplines. Within Artificial Evolution itself different approaches have their own conferences and their own separate literatures; applied research often has only limited contacts with more theoretical approaches; interesting lines of research pursued by individual groups fail to develop into sustained programs of research. And at a higher level, approaches whose founders drew inspiration from biology have tended to cut their ties with these disciplines: in the technical literature on ANNs or Evolutionary Computing references to recent biological research are surprisingly rare.

From a technical point of view, groups pursuing interesting lines of research have often lacked the computational resources to build large-scale simulations of biological processes which often involve large numbers of units, evolving and interacting over very long periods of time.

However, the real reasons underlying the lack of progress are much deeper than this. The biological processes Artificial Evolution needs to model and emulate - in particular the processes regulating ontogenesis - are immensely complicated and appallingly hard to model. As a result, and despite recent interest in ”systems biology” (Kitano, 2003), the main thrust of recent biological research is towards the investigation of specific systems and organisms rather than broad theory. This creates a fundamental problem for Artificial Evolution: traditional evolutionary theory lent itself naturally to mathematical modelling; more recent research, on the other hand has generated a vast wealth of disjointed information which has yet to be adequately organized. What is missing today, and what Artificial Evolution requires, is a 21ŝt century **theory of the evolution of biological complexity**.

Among other things, such a theory would make it possible to achieve a goal of vital importance, namely to identify the **intrinsic limitations** of evolution, identifying and characterizing problems that no evolutionary process -natural and artificial - will ever resolve.

7.6 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 the strategic goals of the project, future research will have to develop a broad range of innovative techniques, each of which represents a significant scientific challenge in its own right. These might include, among others:

1. Models incorporating mechanisms for macro-evolutionary change e.g.: modifications of regulatory mechanisms with macroscopic phenotypic consequences, symbiotic relationships and horizontal gene transfer; gene, chromosome and genome duplication.
2. Techniques for the automatic design of highly evolvable structures and behaviors with the ability to rapidly adapt (via micro-evolutionary processes) 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 (e.g. organisms, demes)
5. Models which combine evolution and development, in which ”artificial genomes” include both regulatory and structural elements, coding a (flexible) 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. A key practical goal be to maintain communications between teams working in different areas and employing different approaches. To achieve this it is essential that the project identifies **benchmark problems**: problems which are sufficiently complex to be insoluble with current techniques yet which are sufficiently simple to give developers real hope of success.

7.7 Ethical considerations

The attempt to design design processes, with the capability to produce Artificial Cognitive Systems with levels of complexity (and performance) comparable to natural systems, is - like most of the challenges in the Roadmap - very close to an attempt to artificially create life: Inevitably, it contributes to the ”disenchantment of the world” (weber1919, 1919), to what Bill McKibben has called ”The End of Nature” (McKibben, 1990) - the illusion of a human-generated, human-managed world, in which civilization creates itself, independently of nature. As such there can be little doubt that it is ethically ambiguous, even dangerous.

There is, however, another side to the challenge. For 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 characterization of problems which are by their very nature intractable. Such knowledge is perhaps not without ethical value.

Chapter 8

The Constructed Brain

8.1 Summary

During the last decade the amount of experimental data coming from the brain sciences has exploded. This is a result of the rapid development of new experimental techniques and an enormous investment in equipment. At the same time there is the feeling that 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 at a level which is so basic that we can apply them to artefacts. There are several reasons for this. The brain is obviously a huge and complicated structure and it will take time to understand the way it is working, even if we have all pieces of the puzzle together. But even so, we are not using the data that are available to us now very well. Data in brain science is very heterogeneous and scattered over thousands of books, and journal papers, in which form they are hardly accessible for modellers. And much of modelling in brain science and cognition is done on the basis of home-grown software. There is no established software repository for cognitive modelling, which leads to an enormous duplication of effort and a limitation on the complexity of cognitive models. We only have to look at CERN, which has invested several person millennia in the creation of a well maintained software infrastructure to see that there is nothing comparable in brain science. This imposes a severe limitation on the usefulness of the experimental data. These problems are aggravated by the fact there is still not much communication across traditional scientific boundaries. Cognitive neuroscience, for example, replicates many findings which have been known in psychology for decades.

The lack of coordination in the interpretation of experimental data is a world-wide problem. There is an enormous potential for Europe to make better use of experimental data in brain science, even without investment in new experimental equipment, if it undertakes an effort to improve coordination of theoretical efforts to understand the brain. There are many ways to do this, but a project which involves a simulation of brain processes as detailed as possible is one of the obvious way to start. Such a project may be the core of a much needed theoretical infrastructure for brain science.

Moreover, within one or two decades computers will be sufficiently powerful to implement brain-like computing on a scale and with a complexity that matches the brain of large mammals, primates and eventually humans. New molecular-scale, non-deterministic, computing technologies will make it possible to implement real time, brain-like computing on compact, low power hardware. By applying principles from biological information processing, it will be possible to build a new breed of "computers", whose cognitive abilities and mobility would be comparable to those of humans.

These new truly intelligent machines will have important applications in many sectors of society and they will be of great commercial interest. In parallel our understanding of the healthy and diseased brain will increase dramatically, due to the availability of more powerful brain modelling tools and an increased use of models in the experimental brain sciences, also allowing for a better utilization of data from animal experiments. This development is now in its infancy but it is already of strategic significance for Europe. To keep a leading edge in this field it is crucial to strengthen the research basis in the directions of computational neuroscience, parallel and distributed brain-like algorithms and architectures, and unconventional molecular level computing substrates 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, which focuses on the cognitive abilities as they exist in humans and animals and investigate what kind of neuronal networks how these abilities can be realized in neuronal networks. The approaches complement each other, and most likely will converge.

It is highly likely that understanding the brain implies understanding computing on a massively parallel substrate. Combining such an understanding with the current developments in both conventional and novel kinds of hardware is a truly exciting prospect.

8.2 Motivation and Objectives

Today we are close to the point when we will have sufficient computing power to simulate a complete brain in considerable detail. Depending on available computer power, the brain we simulate could be a human brain, or a smaller monkey, cat, or insect brain. By analogy with the 'virtual cell' we could call our simulation a 'virtual brain'. If we integrate hardware devices into the computational architecture (see examples), we would have an 'incorporated brain'. Alternatively 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. Unless we assume that nature is tremendously suboptimal, artificial neuronal networks with comparable performance will need comparable numbers of units and connections. Today, most neuronal network simulations 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. If current trends continue, it is likely that sometime between 2015 and 2020 performance will reach 10^{10} GFLOP per second (see Figure 8.1). In a neuronal network the size of the human brain, i.e. with about 10^{11} neurons and 10^{14} synapses this would amount to performing about one hundred synaptic computations per millisecond per synapse. With this capability, it would be possible to simulate networks the size of the human brain, including online learning at 1000 updates per second. Computing power on a similar scale could also be used to perform brain-scale biophysically detailed simulations in real time.

It is, of course, not at all clear to what degree networks of this kind could emulate the actual information processing going on in the brain. But what this simple example shows is that real-time simulation of brain-sized neural networks are not very far in the future. Although the brain has a complex and intricate multi-network structure, the amount of raw computational power required to simulate it is independent of complexity at this level. Today the key obstacle to the development of 'intelligent' applications is not raw computational power but 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 (see chapter 9 and below)? Or could the crucial factor be *spiking* neurons and the way they code information?

To answer these questions, what we need is 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. This would allow us to develop engineering principles for bio-inspired hardware in a systematic way. Using these principles, we could identify cognitive functions which could be implemented in existing hardware, and specify what kind of hardware we would need to implement functions which cannot be implemented in this way. In a ten to twenty years 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-style and brain-size computation;
- Design artificial nervous systems and brains capable of perception, attention, decision-making, action and movement control, and learning.

8.3 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 increase our ability to simulate large-scale and full-scale brain networks, allowing us to implement and investigate very-large scale, brain-inspired architectures and algorithms. Looking even further ahead, the robustness of brain-inspired algorithms will provide the logic needed to tame – and exploit – the inevitably stochastic nature of molecular scale computing

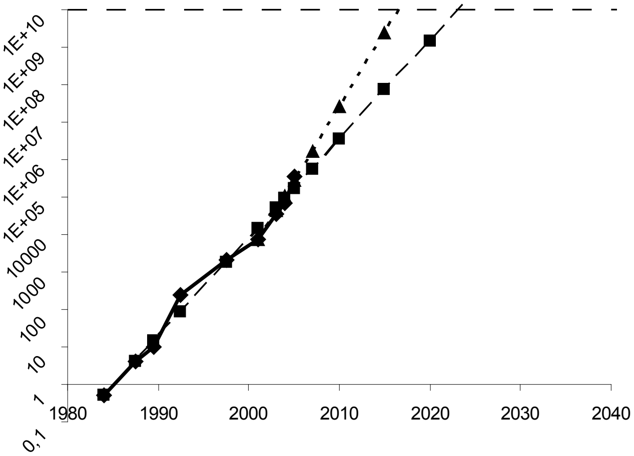

Figure 8.1: Development of computer capacity with time. Vertical scale is in Gigaflop. The solid line shows historical data, the dashed line shows Moore's law and the dotted line shows the current trend, extrapolated from the recently released IBM BluGene family of computers and the STI Cell processor family. The horizontal dashed line corresponds to 100 synaptic computations per synapse per millisecond for the human brain.

At the same time, it is possible to study cognition at a higher level, and investigate how neuronal architectures can implement this performance. Language, for example, has been notoriously difficult to implement in artefacts. It turns out that even basic aspect of language representation provide constraints at the neuronal level (van der Velde & de Kamps, 2006). So, rather than just focussing at detailed, realistic neuronal simulations, another approach can be to model behaviour in great and realistic detail and move downwards to the neuronal level.

Of course, these approaches complement each other, and are by no means mutually exclusive. In this section we will give a review of the state-of-the-art in large-scale realistic simulations of neuronal networks, but we will also discuss other activities which are relevant for 'constructed brain'.

8.3.1 Large-scale biophysically based brain simulation

Quantitative 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, , e.g, GENESIS (Bower & Beeman, 1998), NEURON (Hines & Carnevale, 1997), CATACOMB `http://www.compneuro.org/catacomb`, can be found at `http://www.neuroinf.org` (links)).These packages are powerful and useful tools, which make it possible to set up sophisticated simulations. But with the exception of CATACOMB, they are very much 'stand-alone' tools, with few interfaces to other programmes and databases. A number of publications have described sophisticate algorithms for the simulation of large groups of neurons (e.g. Hansel et al., 1998; Mattia & Giudice, 2000; Djurfeldt et al., 2006). But with some exceptions (e.g., SPIKENET (Delorme et al., 1999)) the programmes used to obtain key theoretical results, are not publicly available. On a higher level, the most common simulation tools are neural network simulation packages. For example PDP++ (OReilly & Munakata, 2000) and SNNS (Zell, 1995). In general, these packages 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. Most of this software is written with a very specific problem in mind. In general there has been little effort to standardise data formats, coding practices and software distribution. Little or no attention has been paid to designing interfaces allowing different simulation tools to be used together.

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 a lot of detail of the system under study and allow an open ended investigation of its properties. 'Top-down' approaches use abstract models or pure mathematics to cast the general principles of the system under study into a minimal model, describing its essential properties with as few parameters as possible. The central role of modelling is to bring together experimental data from different sources to build a coherent picture of the system under investigation. The resulting model can then be used, for instance, to demonstrate how seemingly unexplained phenomena are, in fact, a consequence of what is already known. Exploration of the model can also produce truly unexpected findings, which then provide important input for the planning of new experiments. In this manner, quantitative modelling allows us to extract maximal knowledge from existing data and to find the most promising way forward. Increased use of modelling will speed up the build up of knowledge and will reduce the need for animal experiments, by maximizing the knowledge extracted from the experiments performed.

It is sometimes claimed that 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. The brain is a giant and complex, non-linear dynamical network. Thus, 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, (e.g. a cortical minicolumn with a few hundred neurons) or by dramatically sub-sampling the global network (e.g. by letting one model neuron represent an entire cortical column or area). In sub-sampled network models, in which the parameters for model neurons are based on their real counterparts, one significant problem is how to provide enough synaptic input current to activate model neurons. In a small network there will be very few pre-synaptic 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. This kind of model significantly distorts the dynamics of real cortical networks where many weak signals interact. As computing power increases, researchers will no longer need to make this kind of compromise.

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. Even a rather simplistic, conductance-based, multi-compartmental model of a neuron comprises tens of equations and hundreds of parameters. Large models have thousands of synapses per neuron, each with their own equations and parameters. Thus, any brain-scale network model would contain many billions of parameters. At first sight, it appears that there is not enough experimental data to constrain the model and that even with advanced automated parameter search it would be hopeless to find a reasonably adequate combination. But fortunately typical neuronal networks comprise a limited number of cell types, each with roughly the same properties – though there may be some variation within the population. Thus the parameters used for one neuron of a certain type are likely to do for the others as well, particularly if we know the distribution around the mean. This is also true for synaptic interactions, though now we need to consider pairs of cell types. Here, models can benefit from the fact that synaptic conductances are not determined arbitrarily but are the result of an original genetic specification and the action of plasticity and learning rules coupling them to historical network activity.

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 used to instantiate the network. A huge network model can use the same averages, distributions and learning rules as a tiny one. In this setting, the distribution of cell and synaptic properties becomes as important as their mean values. The realism of the model increases as the number of neurons and synapses approaches those of the system being modelled. In large- and full scale neuronal network models the key limiting factor becomes synaptic rather than neuronal complexity. Thus in really large networks there is little extra cost associated with complex cell models. Since overall compute time is dominated by synaptic computation the choice of methods for problems such as dendritic integration does not add significantly to the cost of computation.

Perhaps surprisingly, the time spent in parallel neural simulation is often bounded by local computation and communications with processor-memory, rather than by inter-processor communication. A key factor influencing the speed of communication is whether neuronal interactions can be represented by spiking events or whether they are graded. There is no doubt that graded communications play an important role in biological systems. However, it appears that in many brain systems the prevalent form of communication is spike-based. From a scalability viewpoint, it is crucial that action potentials should be represented as discrete events, making it possible to use so-called AER (Boahen, 2000) communication. As soon as we try and represent graded, millisecond-scale, cell-to-cell interactions, such as those involved in the gap junctions or in sub-threshold transmitter release, we have to completely change the organization of our parallel simulations. In this new setting, neuronal interactions are potentially performance limiting.

8.3.2 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. As we learn more about the basic computational principles underlying neural information processing, the number of theories compatible with experimental observations will shrink further – perhaps to the point where 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, addresing morphology, ion channels, receptors, cable theory and so on. 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 (e.g. various authors, 2001) (rate coding, precise inter spike times (e.g., Maass, 1997) etc.), 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, at least at a superficial level, this in itself is not very helpful. One would still have to extract higher-level cognitive 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 the way large groups of neurons respond to stimuli is very complex, recent work has made significant progress. For example, it has been demonstrated convincingly that large groups of neurons (e.g., Amit & Tsodyks, 1991; Gerstner & van Hemmen, 1992; Knight, Manin, & Sirovich, 1996; Omurtag, Knight, & Sirovich, 2000; Eggert & van Hemmen, 2001), can be described by powerful sets of equations, and that such descriptions can even incorporate

	human	macaque	cat	rat	mouse
minicolumns	$2.0 \cdot 10^8$	$3.0 \cdot 10^7$	$6.0 \cdot 10^6$	$5.0 \cdot 10^5$	$2.0 \cdot 10^5$
hypercolumns	$2.0 \cdot 10^6$	$3.0 \cdot 10^5$	$6.0 \cdot 10^4$	$5.0 \cdot 10^3$	$2.0 \cdot 10^3$

Table 8.1: The number of mini- and hypercolumns in cortex calculated for a number of mamals

a degree of neuronal detail (Casti et al., 2002). Although solving such equations is non-trivial, 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 (e.g., Amit & Brunel, 1997), attention, and the formation of orientation columns in visual cortex (Nykamp & Tranchina, 2000), 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 fact, synthesis of EEG from biophysical neuronal simulation is already done. Lansner's group is working on synthesizing MEG, BOLD/fMRI and Voltage-sensitive-dye signals from such simulations. This will soon be a very useful way to relate simulation results to real measurements of global dynamic brain function.

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 etc.) (e.g., Zipser & Andersen, 1988; Pouget & Snyder, 2000), attention (e.g.; van der Velde & de Kamps, 2001), and long-term memory formation in the hippocampus complex (e.g., Rolls & Treves, 1998), have all been modelled using Perceptrons 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 (see e.g. Gerstner, 1995; Maass, 1997; de Kamps & van der Velde, 2001). 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, e.g. 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 (see Tables 8.1 and 8.2) the network may still comprise millions of units and billions of connections.

This kind of model constitutes the key interface between brain modelling and brain-like computation. The complexity of the individual computational units is similar to that of Artificial Neural Networks. But thanks to careful, step-wise reduction of complexity we can be more confident that they represent neuronal computational mechanisms, actually present in the brain. Brains have very complex structures not only at the low level but also at the network-to-network level.

	human	macaque	cat	rat	mouse
corticocortical connections	$3.2 \cdot 10^{13}$	$4.8 \cdot 10^{12}$	$9.6 \cdot 10^{12}$	$8.0 \cdot 10^{10}$	$3.2 \cdot 10^{10}$
connectivity	$8.0 \cdot 10^{-4}$	$5.3 \cdot 10^{-3}$	$2.7 \cdot 10^{-2}$	$3.2 \cdot 10^{-1}$	$8.0 \cdot 10^{-1}$

Table 8.2: The number of connections and connectivity in the cortex. The connections are not synapses, but bundles of connections between minicolumns. The level of connectivity is computed as the fraction of full connectivity between minicolumns.

Connectionist style models have to represent this complexity. Learning rules will be mainly unsupervised and correlation-based, as in biology. It is likely that competitive learning in local modules, modelling cortical hypercolumns, or in different cortical layers, will be important for adaptive formation of unit response properties. 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 march a rat brain. The key issues these models have to face include:

- 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.

8.3.3 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. This suggests that the computational capabilities of the brain may rely on a relatively small set of cortical configurations. The idea is that, despite the astounding variety of complex computational tasks, performed by the cortex, they might all be based on a the same basic computational principles. For instance there is growing theoretical and experimental evidence that all cognitive performance is based on a small set of basic 'cortical circuits' (Douglas, Martin, & Whitteridge, 1989). It is at a higher level that computational architectures begin to emerge. There is relatively strong evidence, for instance, that the visual cortex uses a so-called 'blackboard architecture' (van der Velde, 1997; Bullier, 2001; de Kamps & van der Velde, 2006; van der Velde & de Kamps, 2006) in which different high-level features of visual stimuli, such as colour, form, motion, etc., are processed by high-level visual areas. Feedback information from higher to lower visual areas can lead (where necessary) to a re-evaluation of information in lower visual areas helping, for instance, to solve problems of binding. Several researchers have suggested that 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. We would also be in a position to decide if it is best to emulate such structures in existing or in some kind of future hardware, and to forecast whether such emulations could provide the desired performance. At this point, the road would be open for the construction of brain-like, genuinely intelligent artefacts. Obviously, this would be a technological breakthrough of great significance.

8.3.4 Databases

Theoretical modelling must be constrained by data. It is essential therefore 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` (links). Some of these databases are designed very professionally, making good use of modern database techniques. 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 (van Horn et al., 2001), 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. Web pages referring to these databases have many broken links. Few of the databases conform to the high standards set by e.g. COCOMAC. To date, there has been relatively little effort to standardise data formats and database design. Recently, however, there have been a number of attempts to address the problem. NeuroML (Goddard et al., 2001), for example, is an XML extension which aims to enhance the interoperability of databases, simulations and computational models. NeuroML is a relatively recent development and still has to establish itself. If we are to use it for the 'constructed brain' it will need a number of extensions to incorporate higher cognitive concepts. Nonetheless NeuroML and BrainML (`http://brainml.org` (its American counterpart) do look beyond single problem domains - something very few other initiatives even seek to achieve.

8.3.5 Dedicated hardware

Although reductions in size and increasing clock speeds may be approaching their limits, the number of processors in parallel computers will continue to increase dramatically. IBM's Blue Gene computer family (IBM, 2006/2006) is currently the most brilliant example of this trend. 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. Neuronal networks are fairly homogenous computational structures. As a result, parallel simulation and execution is relatively straightforward. Thus, in the near future, we will have sufficient computer power to build large, and even full-scale models of global brain networks. 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. In models on this scale, millions of biophysically detailed neuron models interact via billions of synapses (Djurfeldt et al., 2006). With currently available equipment, simulating one second takes on the order of an hour.

Traditional computers designed for high-precision deterministic computing are not ideal for running brain-style and brain-sized computation. Real neuronal networks use inherently low precision, stochastic computation, without losing their ability to perform useful computation. In the future, we should take advantage of this robustness to design compact, low-power dedicated hardware, perhaps using molecular-scale substrates with computational characteristics similar to those of real neuronal networks. Though these non-traditional compute substrates are unable to execute conventional code, their properties match very well to the needs of homogenous stochastic neural computation. Other non-conventional technologies, which are currently being explored include analog VLSI(Chicca et al., 2003), single-electron devices(Oya, Asai, Kagaya, & Amemiya, 2005) and nano-grids, possibly in combination with CMOS6. All these technologies have advantageous properties which make them very attractive for neural networks. These include very small size, low power consumption and potentially high speed. However several are only in their infancy. In the medium-term practical brain-like computing is likely to be based on more conventional approaches.

8.4 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.

But obviously this will take time. 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, etc.). The most important deliverables from this kind of project, should be software libraries. These should conform to high quality standards for code, interfaces, documentation, and maintenance and should be verified by independent experts. All libraries should be designed for compatibility with a pre-defined set of platforms (including high-performance computers). The requirement that models should not only be published, but should be based on well-designed software libraries 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. The requirement that hardware should have a well-defined software interface opens interesting possibilities: expensive hardware could be maintained locally in a laboratory set-up, with remote users accessing the hardware via the software interface. Special encouragement should be given to projects that promise to create new synergies, using common software interfaces to combine models or to integrate models with 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 interfaced and a model for multi-modal representations, etc. The emphasis on software libraries, would distinguish this 'start up' project from other projects in the same area of research. In the future, these software libraries could themselves become an important object for study. We will need to investigate how they can be maintained, and how they can be designed to facilitate future extensions, meeting the needs of projects which have yet to be conceived. These are important software engineering problems in their own right. But the 'constructed brain' gives them a new flavour: as we learn more about how the brain works, some of this insight might find its way back into our techniques for building reliable, fault-tolerant software which 'degrades gracefully'. In addition to the work just described it will be necessary to conduct a number of parallel activities, allowing iterative refinement of the initial plan.

- Aggregation of databases containing data for different levels of brain organization. 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. This will require the design of new interoperability techniques, including methods to load parts of large databases into local memory in a flexible way.
- Creation of an external environment for the brain to interact with. In the first instance, the environment would probably provide simulated sensory input and output, corresponding to

very simple, 'abstract' simulations of the 'real-world' . In later the environment could be extended to include 'real' sensory input and motor output. Later still it might be possible to consider sequences of sensory input and the consequences of motor actions on sensory input.

- The development of theoretical methods capable of bridging the gap between phenomena at different levels in the organization of the brain. One example might be statistical mechanical methods describing the collective behaviour of large groups of neurons. Another could be dynamical models for cortical circuits, capturing the essence of more detailed and realistic neuronal models, for instance the phase space portrait, but requiring less computing power. A successful programme describing cortical circuits could probably be the input for a similar programme representing higher-level cortical structures.
- Investigation of 'computational architectures', from a theoretical and an experimental point of view. If, as we have argued in section 8.3.3, there exists a 'cortical principle' of parallel computation, or at least a few relatively simple principles, , we could implement these principles in hardware. If we knew the principles we could select the most suitable hardware for the emulation of specific cortical functions.
- Creation of visualization tools providing 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. The core disciplines underlying computational neuroscience – brain science and information technology – are conservative disciplines. As a result, this kind of 'bridging activity' does not receive enough support from national academic establishments. Scientists who devote their careers to it are not always as successful as others in the competition for resources and scientific esteem. It is essential that FP7 should contribute to remedying this situation. In attempting to achieve this goal there is much we can learn from other sciences which have established large multi-disciplinary collaborations. One of these is bioinformatics - an area in which the Human Genome Project has provided an enormous impetus for coordination. Another area in which there has been a lot of collaboration is high energy physics. The design and operation of the world's few large accelerators brings together theoretical work in particle physics, heavy engineering (accelerators are huge), and electronics. High energy physics created the `WWW`. It has also developed software suites for detector simulation, data analysis and visualisation, which are used by virtually every high energy physics laboratory in the world. Database techniques and distributed computing techniques (the `GRID` project) from high energy physics have won much attention from other branches of science. 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.

Chapter 9

The Brainprobe project-Tools for Neuroscience

9.1 Introduction

Why do we need neuroscience?

Never has information technology realized so acutely that it's traditional ways of tackling problems fall short and has the quest for using smarter, more cognitive artifacts been more pressing. Computer vision is going cognitive at every occasion, IST has launched a call for cognitive systems, robotics dreams of cognitive robots, objective of the 'Beyond robotics' FET call, in this roadmap we are proposing self-aware complex systems etc. Thus at the technical side there is a tremendous need for facts and even more so for principles about brain functioning. This need does not only follows from the fact that we want smarter, more reliable, more flexible systems. The need equally proceeds from the observation that artifacts of whatever nature, in most cases have to interact with humans and have to be accepted by them and thus must somehow be tuned to the human mind. On the other hand neuroscience, and in particular the systems neuroscience which is the component most relevant to information technology, is making giant leaps forward due the introduction of functional imaging techniques of the brain. This has cumulated recently in the introduction of functional imaging in primates, which establishes the bridge between the human work and the knowledge from invasive techniques, accumulated the last forty years. That this last development has occurred in two European labs opens an extraordinary opportunity for the EU to lead the world in linking neuroscience and information technology, 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 has provided them with much needed support. There are two extremely compelling arguments for neuroscientists to collaborate with engineers. First, trying to build real world systems provides a much clearer picture of the problems an artificial system, thus also the brain, has to solve. The classical example in vision is segmentation. Neurophysiologists became aware of this problem only after engineers had made them realize that the single stimulus introduced by psychology hundred fifty years ago was a laboratory abstraction: what is present on the retina is a spatio-temporal distribution of light, not the image of an object. The second reason is that the brain is so complex that even models are insufficient to understand this complex reality and that it is even more difficult testing that the model captures all the facts. By building a real system according to the model and verifying that indeed this system solves the problem, provides evidence in favor of the model. Given the need for an increased cooperation between neuroscience and information technology, it makes sense to increase the potential of European neuroscience, so as to enhance the dialogue. It is indeed the case that to list the problems is easier than to solve them, and that the more efficient neuroscience gets, the more information technology will benefit from the dialogue. In order to know how to strengthen neuroscience, it is crucial to understand the complexity of the brain, an organ different from any other in the biosphere.

What makes the brain so special?

The brain differs from most other organs of the body because of the connections between neurons: each of the 10 billion or so neurons in the human brain is connected to 1000 or more other neurons. The cerebral function is heavily dependent on these connections: in fact knowing all the detail of the cellular equipment of neurons is insufficient to understand brain function. Brain function arises from the concerted action of anatomically organized groups of neurons.

These anatomical connections 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 subneuronal levels, these latter including the subcellular-part level (e.g. the synapse) and the molecular level. Although these latter levels are also found in other cell types, some of them, e.g. the synapse or certain molecules producing transmitters, are typical for the brain.

To understand the brain we need to be able to address these different levels, and integrate information across levels e.g. by modeling (see chapter constructed brain). While we have powerful techniques to address the neuronal level (single cell recording in awake animals), the whole brain level (psychology and cognitive sciences), and of course the subcellular (patch clamp etc), molecular and genetic levels, techniques to address the supraneuronal levels have only begin to develop recently. These supraneuronal levels are extremely critical for understanding brain function and are most relevant to neuro-IT, because they embody the computational principles we want to endow artifacts with. We propose that the combination of dense multiple recording with functional imaging can address these intermediate levels of integration and provide the data required for relating all integration levels from the single cell to the whole brain level. If Neuro-IT is to flourish these techniques have to be developed further at maximum strength. Only then will neuroscience be able to produce the data required by the modeling and computational studies, which are at the heart of neuro-IT. The different levels of possible interactions between systems neuroscience and robotics are indicated in figure 9.1

Although this roadmap is intended for Neuro-IT it is worth mentioning that most of the recommendations made here, hold for other programs such as Quality of Life. In fact the only justification for the experimental strategies recommended in this document, however useful for FET and IST, resides in the medical domain. These recommendations may be the more useful

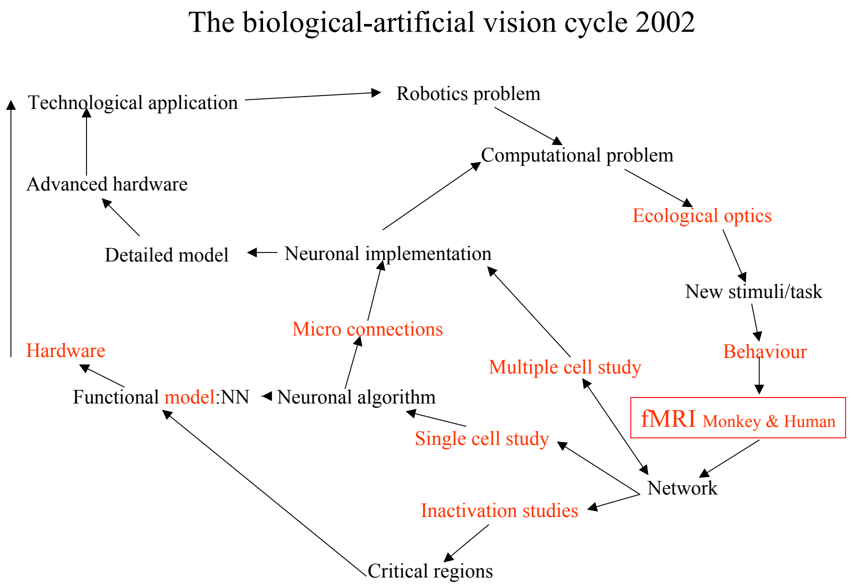

Figure 9.1: The different levels of possible interactions between neuroscience and robotics.

that it has recently become clear (EU conference on structuring the European brain research, 18/9/03) that the crucial intermediate levels of brain organization have been largely neglected in the first calls of Quality of Life under the sixth framework. This was the unfortunate consequence of treating the brain as any other organ of the human body.

9.2 Objectives

1. Strengthen the knowledge base of European neuroscience, to enhance the cooperation between information technology and the 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

9.3 Examples of realizations

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

9.4 Current state of technology

9.4.1 Brain imaging technologies

Positron emission tomography

Positron emission tomography (PET) uses radioactive tracers to visualize brain function. With modern scanners the amount of tracer to be injected is minimal and studies are ethically readily justified, but as a rule 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. In studies of regional cerebral blood flow one compares levels of activation in different conditions, since regional blood flow correlates with neuronal activity. During the 1985-1995 period this was the main avenue for functional study of the human brain (Fox et al., 1986; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Dupont et al., 1993; Dupont, Orban, De Bruyn, Verbruggen, & Mortelmans, 1994; Paulesu, Frith, & Frackowiak, 1993). Spatial resolution of PET activation studies was limited by the need to average across subjects and also by the physical process of positron emission, typical values were full width at half height (FWHH) of 16 mm. While this resolution was plenty to discern coarse localization in the brain, it was inadequate to study neighboring functional maps, some of which may only be 10 or 15 mm in size. Therefore activation studies have been taken over by functional Magnetic resonance imaging (fMRI) which has better spatial resolution: FWHH of 7-10 mm for group studies and 2-4 mm for single subject studies. FMRI allows repeated testing of the same subject and comparison between different activation regions in a single subject. FMRI is subject to susceptibility artifacts, especially in the temporal and prefrontal cortex. Therefore in particular studies e.g. of language where these regions are crucial, PET activation studies remain an option, as they are also in Fluoro-deoxyglucose (FDG) PET correlation studies with behavioral deficits in patients.

PET remains unsurpassed for its other main application: studies of neuronal transmission. In fact, with the advent of new more powerful and higher resolution cameras (micro-PET), claimed to reach mm resolution, this is presently the main avenue for so-called metabolic imaging, which is outside the scope of the this review. This will be of interest for animal experiments where the system has to be manipulated by local pharmacological injection. These experiments will also be important complements for fMRI studies with pharmacological (systemic) challenges. A final remark about metabolic PET: this technique depends heavily on the development of tracers and on radioactive chemistry laboratories (and cyclotron) to produce these tracers locally.

Functional magnetic resonance imaging (fMRI)

fMRI is based on the BOLD (brain oxygen level dependent) effect reflecting the different paramagnetic properties of oxy- and deoxy-hemoglobine (Belliveau et al., 1991; Kwong et al., 1992). In fact the BOLD effect is dependent on three hemodynamic variables: blood flow, blood volume and oxygen extraction. This effect increases with diameter of the vessel explaining why BOLD imaging necessarily suffers from a draining vein problem (Mandeville & Marota, 1999). Many of the new sequences and one of the reasons to move to higher field strength is to minimize this localization artifact. The other reason for higher field strength is a better signal to noise ratio, which can then be traded for resolution. The typical voxel size in 1.5T magnets is 3x3x4.5 mm, which is gradually replaced by 2x2x2.mm in 3T magnets, which are becoming the new standard.

Just as PET activation studies, fMRI measures neuronal activity indirectly and needs to compare MR signals in different conditions. In the simplest design, two conditions are presented in alternative epochs and mean MR activity in the two epochs is compared, either using General linear theory, as in Statistical parametric mapping (SPM), or by correlating the time course of stimulus alternation with that of the MR activity, as e.g. in AFNI. To enhance the interpretability of the findings it is very useful to add a third low level control condition (e.g. simple fixation in visual studies) to the two other conditions which are matched as closely as possible to extract the effect of single factor. By adding a low-level control we can disentangle small differences in activation of in deactivation (relative to baseline) between the main conditions.

This subtraction design has been criticized 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 (Friston, 1997). Factorial designs have the additional advantage that interactions between factors can be studied. While it may be difficult to isolate a cognitive factor in a single subtraction this might be obtained by taking the conjunction of different subtractions (Price & Friston, 1997). The requirements are that each of the subtraction isolates the factor in question even if mixed with other irrelevant factors, and that these other factors differ between the subtractions. Conjunctions are also useful to make sure that higher order effects such as interactions are studied in relevant regions, e.g. those showing a main effect of a single factor.

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. Simple fixed effect models allow one to derive conclusions only about the population scanned. To ensure general conclusions, one needs to use the random effect model in which both within and between subject variability are modeled. The prize to pay for this generality are very high thresholds if one corrects fully for multiple corrections. Classically, one accepts as significant voxels reaching $p < 0.05$ corrected for multiple corrections, unless a priori information is available, in which case $p < 0.001$ uncorrected is acceptable for significance. The risk of false negatives in random effects can be alleviated by restricting the analysis either spatially (e.g. to the posterior part of the brain) or functionally by a low level subtraction (e.g. only visually responsive voxels) or by lowering the threshold, e.g. to $p < 0.0001$ uncorrected, taking into account the number of false positives. On the other hand, are studies in which only single subject analyses are performed. This is especially attractive when different functional activation patterns have to be compared spatially, i.e. when overlap or juxtaposition between the activated regions has to be established.

In between we find the ROI analysis in which the ROI can be defined anatomically but more generally functionally by a so-called localizer scan. Much of the results of these studies depend on the definition of the ROI. The better established the identity of the ROI, the more the localizer paradigm is standardized, the more reliable the localization of the ROI in a given subject will be. Examples of well-defined ROIs are the topographically - defined regions in sensory cortex (e.g. retinotopic regions in the visual cortex). Their identity is well established, at least the early ones, the paradigm to define them is well established (Shipp, Watson, Frackowiak, & Zeki, 1995; Sereno et al., 1995) and a good experimenter will be careful to sample enough images so that the area or subarea (Heeger, Boynton, Demb, Seidemann, & Newsome, 1999) is unequivocally defined. The differences between magnitudes of the MR activity averaged over the ROI in different conditions can then be tested statistically (e.g. with ANOVA) across a number of subjects.

The time course of the BOLD effect is slow, yet fast enough to be convolved with brief trials or with different subperiods 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 rather than over a whole block of trials including the intertrial intervals as done in block designs. In addition, this technique allows the comparison between different types of trials, e.g. correct and incorrect trials, trials with targets and without distracters, trials with stimuli in different parts of the visual field etc. The cost of these more specific activation patterns is the loss of statistical power: event related activation levels are weaker than block design activation levels. This lack of power can be offset by increasing the number of subjects. 47 subjects participated in a recent event related fMRI study of remembering (Wheeler & Buckner, 2003). An increasingly used application of event-related fMRI is the repetition paradigm. In this paradigm, trails with two identical or different stimuli are compared with trials in which two stimuli are presented of which it is unknown whether or not the brain treats them 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, one concludes that a brain region treats them either as identical or as different.

The time-related fMRI with long trials allows the experimenter to separate processes that operate at different instants of the trial, such as visual processing, maintenance and response selection in working memory trials (D'Esposito et al., 1995; Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000). Of course this is natural in tasks such as working memory tasks where the delay has to be long, but may be more difficult to apply to other tasks such as performing mathematical operations. Indeed, it requires introduction of long delays which may be unnatural for this type of tasks.

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, e.g. 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 monkey is indeed the only adequate animal model for higher order brain functions. It poses however a severe problem (Orban, 2002) since the comparison entails both a change in species and a change in technique and one needs to understand the effect of both factors. This cannot be solved easily unless one resorts to a new technique, fMRI in the awake monkey (Vanduffel et al., 2001; Nakahara, Hayashi, Konishi, & Miyashita, 2002; Vanduffel et al., 2002). With fMRI in the awake monkey these questions can be uncoupled. On one hand, one compares within the same species single cell recordings, and other local electrical changes, with MR signals (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). On the other hand, one compares with the same technique human and monkey brains, and addresses the homology problem between these two primate brains (Nakahara et al., 2002; Vanduffel et al., 2002).

In monkey the functional MR signals are smaller than in humans and the initial measurements with simple BOLD were heavily contaminated by artifacts (L. et al., 1998; Dubowitz et al., 1998). This can be solved either by resorting to high fields (Logothetis, Guggenberger, Peled, & Pauls, 1999) or by using a contrast agent (Vanduffel et al., 2001). In this latter study monocristalline iron oxyde nanoparticle (MION), developed by Weissleder at MGH, was used as contrast agent. This agent not only produces a 5 fold increase in the contrast to noise ratio but also provides MR signal arising from the capillaries located in the gray matter rather than from small veins above the cortex as BOLD does. Given the lack of problems resulting from long lasting chronic use of this contrast agent there is hope that it might be approved for use in humans, if not for routine use in normal subjects, at least in patients. A gain in signal would be welcome in clinical fMRI, e.g. in pre-operative assessment. 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 as performed with fMRI only provide a static description of a set of cerebral regions active (or more active) in given experimental conditions. 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. Functional connectivity is distinct from the anatomical connections, which are fixed (although modifiable by plasticity). Depending on the task the anatomical connections will be used differently and functional connectivity refers to these adjustable strengths of existing connections. It differs from effective connectivity, which simply refers to the positive or negative correlation of activities in two regions.

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, both in PET (McIntosh et al., 1994) and fMRI (Büchel & Friston, 1997). Task-related changes in connectivity have also been considered with this technique (Büchel & Friston, 1997). Alternatives to SEM, but that allow for non-linear and/or time-variant connection weights, have also been introduced, e.g. based on Kalman filtering (Büchel & Friston, 1998) and Volterra kernels (Friston & Büchel, 1998). Bullmore and co-workers (Bullmore et al., 2000) tested whether or not the suggested network could have been drawn from a distribution of ”optimal” models generated by a heuristic search algorithm. More recently, Dynamic Causal Modelling (DCM) ((Friston, Harrison, & Penny, 2003); see also SPM2 beta release) has been introduced to determine the connection weights in a predefined network. Generally, this predefined network is unknown in humans.

Tracing anatomical connections with MRI

In vivo tract tracing refers to local injections in to a brain region of a tracer that can be visualized in the MR. So far only one study has been performed in the monkey (Saleem et al., 2002) 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. When further developed this technique will need verification in animal models, in which anatomical connections are known, as opposed to inferred in humans.

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 (e.g. Dale et al., 2000) this problem is not completely solved in humans. It is worth noting that EEG and fMRI signals can in principle be acquired simultaneously, MEG and fMRI cannot. One should also note that MEG reflects in principle activity mainly of pyramidal cells in banks of sulci, while the EEG reflects more the pyramidal cells on the gyri.

In that sense EEG and MEG are complementary. So far these fusion of imaging techniques has not been attempted in animal models, although again this is the only way to validate them.

Other imaging technologies with limited use

Optical recording (Grinvald, Lieke, Frostig, & Hildesheim, 1994) has a good spatio-temporal resolution but its applicability to old world monkeys is restricted because it requires flat pieces of cortex that are accessible. For example in the visual system only V1, a 2 mm wide strip of V2 and V4 can be studied. Similarly 2-deoxyglucose technique, which has an excellent sensitivity and spatial resolution (Tootell, Hamilton, Silverman, & Switkes, 1988; Vanduffel, Tootell, & Orban, 2000) also has a limited use because only one or two conditions can be studied (single label and double label 2DG). Also this technique is very invasive and critically depends on the availability of specific films (at least the double label version). For the sake of completeness we mention 2- photon and intracellular synaptic imaging.

9.4.2 Multiple single cell recordings

Obviously more information can be obtained from recording multiple single neurons rather than a single neuron. Equally obvious, one should not confuse multiple single cell recording with multi-unit recordings which can be more or less noisy. The noisier the recording, the less clear it is that one records from local neurons as opposed to fibers of unknown origin. The initial techniques (e.g. tetrodes, Thomas recording electrodes) allowed recording of small number 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 (Hoffman & McNaughton, 2002; Donoghue, 2002). 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, e.g. functional architecture see Diogo, Soares, Koulakov, Albright, and Gattass (2003), in addition to synchronization of signals between areas (Hoffman & McNaughton, 2002). One of the other motivations behind these multiple recordings is to control a robot arm or other artifacts by the brain signals obtained. For this purpose the recording of single neurons on each electrode may not be required, local field potentials may suffice (B., Pezaris, Sahani, Mitra, & Andersen, 2002). Potentially these recording could be chronic, allowing to address important questions such as changes in neural activity with learning or prolonged testing of the same set of neurons under widely varying conditions, which may be required to crack problems such as the code of shape representation in IT. The critical point here is not so much to obtain long-term recording but to prove that one is recording from the same neuron over long periods of time.

Links of fMRI with neuronal activity

In a seminal study Logothetis et al. (2001) compared fMRI signal to electrical neural recorded simultaneously with an electrode from the cortex imaged. This revealed that fMRI signals correlate with local field potentials more than with spike activity. It seems thus likely that MR signals reflect more the afferent input and even more local processing in an area than the output of that regions to other brain regions. A far more complex problem is the relationship between selectivity or tuning observed in single cell recordings with average activity of large groups of neurons underlying the vascular response measured with fMRI. Adaptation experiments are touted here as the solution for fMRI but this has not been proven. Indeed the only attempt of validation so far has reached paradoxical results (Tolias, Smirnakis, Augath, Trinath, & Logothetis, 2001).

9.4.3 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, such as ibotenic acid, are usually combined with behavioral testing (e.g. Orban, Saunders, & Vandenbussche, 1995). Note that the ibotenic acid lesions are more specific than surgical excision, as fibers 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 (Squire, 1986). Electrical stimulation is generally used for manipulations in the opposite sense, i.e. driving a brain area. There is long tradition to use electrical stimulation in motor studies. It's application in sensory regions is more recent (Salzman, Murasugi, Britten, & Newsome, 1992) and seems to depend on some uniformity in the neuronal properties at the stimulation site. This was the case in the Salzman et al. (1992) study in which all the cells in a direction column of MT/V5 share the same preferred direction.

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) or muscimol (Gaba agonist). This has been combined with behavioral measures or single cell recordings in an effort to identify afferent systems (Ferrera, Nealey, & Maunsell, 1994) The problem here is to inactivate large enough regions to obtain reliable effects especially in behavioral studies. An alternative is local cooling, which generally can affect large enough regions and can be more rapidly reversed, but which is difficult to restrict to a given anatomical region (Girard, Lomber, & Bullier, 2002). Pharmacologically it has recently (Jones and Sillito, unpublished) become possible to locally increase neuronal activity, even only stimulus driven activity (and not spontaneous activity).

Finally it is worth mentioning that in humans systemic injection of pharmacological agents is used in pharmacological challenge studies (Rosier et al., 1997) 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.

9.4.4 Visual and sensory systems

Monkey visual system

It is now more than ten years since Felleman and Van Essen (1991) 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 visual cortical areas. Each of these areas is on average connected to 10 other afferent and efferent regions. Thus the primate visual system is an extremely complex system that adapts its configuration to the visual task at hand. In comparison rodents have only a few extrastriate areas. Thus except for matching with molecular studies, the physiological 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. Cortical areas are identified by the four criteria: connection pattern, cyto-and myelo-architecture, topographic organization and functional properties. The evidence is lacking in case of division of the infero-temporal cortex or that of the intraparietal sulcus (IPS) and superior temporal sulcus (STS). 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. But even areas that have been explored abundantly such as area MT/V5, are not well understood: their role in visual perception is far from clear. Initially, MT/V5 was a motion area but it is becoming increasingly an area involved in 3D analysis (Bradley, Chang, & Andersen, 1998; Xiao, Marcar, Rraiguel, & Orban, 1997; Orban, Sunaert, Todd, Van Hecke, & Marchal, 1999). The overall impression of systems or integrative Neuroscience is that of a very conservative field. This is largely due to the labor-intensive character of the single cell studies: it takes one man year to perform a study exploring the response of neurons in a given area to a new set of stimuli. 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. One should note that in Europe the ultra-conservative policy for academic recruitments, related to job security, hampers the recruitment of exactly this sort of innovative researchers.

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. Segmentation, extraction of 3D shape from different cues, building representations of objects and actions are not well understood. The task dependency of the visual system (Dupont et al., 1993; Fias, Dupont, Reynvoet, & Orban, 2002) has been explored to some extent in human imaging, but has hardly been addressed in single cell studies. The role of feedback which is anatomically well documented, has hardly been explored physiologically (Hupi et al., 1998).

Human visual system

Functional imaging has shown that in general terms to visual system of all primates are similar. The early retinotopic regions (V1,V2, V3) are similar in lay-out in humans and monkeys(Sereno et al., 1995; Fize et al., 2003). In the same vein the visual system in both species is divided in dorsal and ventral stream related to where and what processing respectively. These stream process to some degree different attributes (Ungerleider & Mishkin, 1982; Haxby et al., 1994), for different behavioral purposes (Goodale & Milner, 1992), using different cognitive operations (Fias et al., 2002). 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 (Tootell et al., 1997; Vanduffel et al., 2001; Tsao, Conway, & Livingstone, 2003). The IPS of humans processes motion information, and in particular extracts 3D from motion, much more than its monkey counterpart (Vanduffel et al., 2002; Orban et al., 2003). For years there have been heated discussions about homologies, e.g. the debate between Tootell and Zeki (Hadjikhani, Liu, Dale, Cavanagh, & Tootell, 1998; Bartels & Zeki, 2000) related to the color-processing region. This was largely based on an absence of relevant data. Now that both brains can be imaged exactly in parallel, these problems can be rigorously addressed. We have no idea of how many cortical areas the human visual system contains, certainly more than. One should remember that some 80the parts underlying higher functions, is still unknown.

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.

9.4.5 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. Histologically, the caudal sector is characterized by its almost complete lacks of granular cells (agranular frontal cortex).

The agranular frontal cortex (henceforth referred to as motor cortex) is cytoarchitectonically not homogeneous, but constituted of several distinct motor name areas named according to a terminology derived from Von Economo F1- F7 (see (Rizzolatti, Luppino, & Matelli, 1998)). Five areas lie on the lateral cortical surface, two on its mesial surface. Comparing this motor cortex subdivision with the classical map of Brodmann, F1 corresponds to Brodmann area 4 (primary motor cortex) while the other motor areas (F2-F7) lie inside Brodmann area 6. 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 (F2 and F7) and ventral motor areas (F4 and F5) 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. Indeed the primary motor cortex is involved in the execution of movements and the school of Georgopoulos has shown that the parameters direction and distance of movement to a target are encoded independently in M1 and premotor neurons. Recent neurophysiological data showed, however, that motor areas play a broader role in behavior and are involved in functions traditionally considered proper of higher order associative cortical areas.

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 (SMA and pre-SMA) are involved in the control of sequences of movements (Tanji et al., 1996). Recent studies of the connections of the motor areas with cortical areas outside the agranular frontal cortex (”extrinsic connections”) showed that there is a marked difference in connection organization between the posterior motor areas - areas F1, F2, F3, F4 and F5 - and the anterior motor areas - areas F6 and F7 - (Luppino & Rizzolatti, 2000)(Luppino and Rizzolatti, 2000). The posterior motor areas receive their main cortical input form the parietal lobe (”parieto-dependent” motor areas). In contrast, the anterior motor areas receive their main cortical connections from the prefrontal cortex (”prefronto-dependent” motor areas).

This subdivision of motor areas is in accord with their connections with other motor areas (”intrinsic connections”). The prefronto-dependent areas do not send fiber to F1 (the primary motor area), but have diffuse connections with the other motor areas. In contrast, the parieto-dependent areas are connected with F1 and are linked among them in a precise somatotopic manner (Matsumura & Kubota, 1979; Muakkassa & Strick, 1979; Matelli, Camarda, Glickstein, & Rizzolatti, 1986; Luppino, Matelli, Camarda, & Rizzolatti, 1993).

Another anatomical finding that strongly supports the validity of this subdivision is the organization of cortico-spinal projections. The parieto-dependent motor areas send direct projections to the spinal cord, while the prefronto-dependent do not (Keizer & Kuypers, 1989; He, Dum, & Strick, 1993; Galea & Darian-Smith, 1994; He, Dum, & Strick, 1995). Specifically, F1, F2, F3, part of F4, and that part of F5 that is buried in the inferior arcuate sulcus (F5 of the arcuate bank or F5ab) give origin to the cortico-spinal tract, while F6 (pre-SMA) and F7 project to the brainstem.

From these anatomical data, it appears inescapable to conclude that the two sets of areas play different roles in motor control. Parieto-dependent areas receive rich sensory information originating from the parietal lobe and use it for action. This process occurs in parallel in several circuits, each of which is involved in specific sensory-motor transformations, e.g. for reaching or grasping. F1, F3 and that part of F2 that is located around the superior precentral dimple (dimple sector) use, for this process, somatosensory information, while F4, F5 and the rostro-ventral part of F2 use also visual information. The scanty sensory information that reaches the prefronto-dependent motor areas renders very unlikely that they also play a role in sensory-motor transformations. The prefronto-dependent motor areas receive higher order cognitive information, related to long-term motor plans and motivation. On this basis, it appears logical to posit that these areas have a control function. Most likely they determine when and in which circumstances the activity generated in the parieto-dependent areas -potential motor actions- becomes an actual motor action.

9.4.6 Cognitive systems

It is well established (Fuster & Alexander, 1971) 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. Neurons in higher order cortices such as parietal or infero-temporal cortex share this property. Typical for prefrontal neurons is that the delay activity represents the sample whereas delay activity of infero-temporal cortex only represents the last stimulus (Miller, Erickson, & Desimone, 1996). Although initially (Wilson, O'Scalaidhe, & Goldman-Rakic, 1993) it was claimed that delay activity in dorsal and ventral parts of prefrontal cortex, linked with dorsal and ventral visual pathways respectively, were specialized for spatial and figural information respectively, recent studies mainly from Miller's group (Rao, Rainer, & Miller, 1997) suggest this is not the case. Also in human imaging it has been difficult (Smith & Jonides, 1999)to show a segregation in prefrontal cortex between object and spatial working memory. Alternative schemes proposed segregation within prefrontal cortex of maintenance and executive functions in working memory.

In addition to delay activity, the task dependency of prefrontal activity has been recently documented physiologically (Rainer, Rao, & Miller, 1999; Asaad, Rainer, & Miller, 2000), as well as its role in categorization (Nieder, Freedman, & Miller, 2002). 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 (Tremblay, Hollerman, & Schulz, 1998). These responses are probably reflecting input from the basal ganglia and the substantia nigra (reflecting reward incongruence). In addition to prefrontal cortex, parietal cortex has been shown to contribute to cognitive functions. According to Goldberg and coworkers, a saliency map resides in area LIP (Gottlieb, Kusunoki, & Goldberg, 1998). This saliency is dependent both on physical factors (pop-out in the stimulus) and on behavioral factors (aspects in the image that are relevant for the present behavior). LIP works in tandem with prefrontal cortex, in particular with the Frontal eye field (FEF) for controlling explicit and implicit shifts in attention (Schall, 1995; Schall, Morel, King, & Bullier, 1995; Corbetta et al., 1998).

9.5 Problem areas

9.5.1 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 scaling of these problems with size of the electrode array should also be addressed. The stability of the recordings is the other problem. It is one thing to have single cell quality recording for months or years, which some groups claim to have achieved, it is another thing to be confident that each day on a given electrode you record from the same neuron. This is probably the most important problem, since it would extend the use of the technique tremendously, e.g. many training experiments would become possible.

9.5.2 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 for somato-sensory stimulation, visual stimulation is generally restricted to a part of the visual field.

The present day trend is to go for higher fields to increase the field strength to increase S/N ratio and resolution. Perhaps we should envision the opposite possibility: lower field strength in which wide bore magnets can be used and in which some subject movement is tolerable. The use of a contrast agent such as MION would still provide a reasonable S/N ratio and resolution (in fact the effect of MION compared to Bold increases with lower field strength).

The development of new sequences providing new type of information about brain function remains important, as is the development of new coils.

9.5.3 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 that for children could be used.

9.5.4 Few mathematical tools

Just like mathematics were developed for physics then for economics, we need mathematics for biology and in particular for neuroscience. Of course statistics is used, as in many sciences, image processing etc. What we need are new mathematical tools to deal with the multiple electrode signals and/or the MRI signals. Mathematics for complex dynamic systems might prove useful. One should keep in mind that in many cases the data are sparse and non-stationary.

9.5.5 Education of the public

In its majority the public is supportive of medical research even that involving animals, the more so that 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. To quote Thomas Insel, the new director of the national Institute of Mental health: 'Often, the general public assumes that new drugs or new treatments develop almost from whole cloth, without realizing that there's often a decade or more of basic science that feeds in to new discoveries that have clinical significance. It is important for us and a challenge for us to make sure the public understands the importance of investing in basic science as a pathway to improving therapeutics.'

9.5.6 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. We must invent something like the RO1 of NIH, especially for those who have acquired expertise in brain studies of primates. They face the most difficulties to return to Europe.

9.6 Future research

9.6.1 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, or at least quality multiunit activity, over long periods of time, 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 pre-amplifying systems is important, as are wireless connections 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.

9.6.2 Improve and diversify brain imaging techniques

To improve the 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.

A better understanding of the vascular phenomena and neural activity phenomena underlying the different MR signals is critical for interpretation of fMRI signals: 'what are deactivations ' do we see inhibition and excitation in fMRI? Is adaptation or priming, as it is sometimes referred to, really a measure of selectivity. Can neuronal selectivity be revealed by other means in the fMRI.

Devise new sequences or new types of MR measurements to extend the range of brain phenomena that can be visualized non invasively: anatomical connections, transmitter systems and other important neurochemical substances.

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 Single cells and EEG are being measured in the scanner but this is an exception, these techniques should become routine and robust.

All brain imaging techniques, used in humans and even in clinical settings, have yet to be properly validated. It is essential to validate them in realistic animal models. For higher cognitive functions, which are the essence of human functional imaging, 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. The potential savings in money and time are vast (up to 80

9.6.3 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. Probably a mathematical treatment of complex dynamic systems can help, but the stochastic nature and non stationarity of the signals, as well as their sparseness and incompleteness should be considered. If possible these techniques should allow introduction of a priori information such as cell types present, anatomical connections etc. This should allow study of input output relationships between brain regions, of functional architecture of a region, or of the local circuits present in a region of subpart (canonical column e.g.). Fusion of different brain imaging modalities such as fMRI and EEG or MEG should be further improved. The visualization of the results both spatially and in time will be important. Techniques to provide general interpretation and integration schemes, such as new coordinate systems, brain atlases, and warping algorithms to compare brains, are important. fMRI measurements contain a wealth of information that is only feebly used; Development of new signal processing tools to extract relevant signals of activity, connectivity and their dynamics are key. Concepts such as the Granger causality (Granger, 1980) hold some promise to model cerebral networks from neuro-imaging data by testing causality between two timeseries (e.g., Freiwald et al., 1999; Chavez et al., 2003). Furthermore, one should perform an exploratory network analysis, rather than starting from a predefined network, perhaps in combination with Granger causality, but then extended to a conditional definition, i.e., causality from x to y, given additional time series z. 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.

9.6.4 Visual and motor systems

The newly development tools hold promise to unravel the important issues of systems neuroscience in ways directly relevant to functioning of the human brain and to understanding the malfunctioning brain.

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, categorization and visuo-motor 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 it is completely unclear which dimensions of 2/3D shape, which material properties, which action primitives or scene dimensions are encoded at high levels in the visual system. 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. The role of parietal cortex in representation of space (or probably multiple spaces such as near, reachable, or distant space,) in decisions and in attention should be further studied, in addition to its links with the premotor cortex (see below). Most robots or other intelligent artifact use vision as one of their main senses, unless these biological vision functions are better understood at the algorithmic level, construction of intelligent vision machines that share at least some of the performances of the human visual system are illusory.

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. These

areas (and especially F6) represent the main avenue through which the frontal lobe controls the motor cortex. It is through this area that the action representations coded in the parieto-premotor circuits become actions. Understanding this control may be of enormous advantage for constructing robots or other artifacts 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? Recent anatomical data suggest that the interaction between the pragmatic aspects and semantics aspect of objects are mediated by input coming from infero-temporal lobe to the motor representations of the inferior parietal lobule. The knowledge of how semantic and pragmatic description of object interact will be of enormous value for the construction of artifacts able to interact with object in an intelligent way (intelligent artificial arms).

3. The discovery of mirror neurons (see Rizzolatti, Fogassi, & Gallese, 2001) provided the basis for understanding the mechanism of imitations. Yet, there are virtually no data on how mirror neurons are re-organized when an individual learn a new action. Experiments are needed to test hypotheses postulating that the observed action (the model) activates pre-extant motor acts coded by mirror neurons and then through the action of the prefrontal lobe these motor acts are re-combined in new configurations identical to those of the model. It obvious that imitation mechanism now is a fundamental issue because of the enormous economical possibilities that will open for construction of robot or other devices 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. The connections however between these areas, including the backward projections from F1 to F5, may explain the formation of neurons or neuronal assemblies able to integrate higher order sensory information conveyed to premotor areas with knowledge concerning action execution, giving the individual a deep knowledge of action meaning. Understanding how motor knowledge is formed could be a fundamental step in the construction of really intelligent artifacts.

As mentioned above for the visual system, many of the fundamental problems outlined above 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 combine use of monkey and human experiments will solve in addition to problem sketched above also other problems here not discussed such as the role of the dorsal premotor areas in movement organization, the role of premotor cortex in association leaning, and last but not least the neural mechanism of intentionality, that is the understanding of why an action has been preformed (its distant purpose).

Other systems

Given its importance in cognitive functions prefrontal cortex should be explored more vigorously at this side of the Atlantic ocean. 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, auditory regions.

9.6.5 Development of software for modeling at and across different levels of integration

See other chapter.

9.6.6 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. Advances in cognitive and neuroscience theory should lead to richer descriptions of the tasks the system is performing. In addition is should generate hypotheses which can be tested with more refined paradigms. Such paradigms should contain gradually less and less abstraction compared to real life situations (e.g. monkeys using tools, taking elevators (in virtual reality)).

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

The complexity of the brain is such that the amount of data collected and to be collected is enormous. It is generally accepted that modeling will be important to generate understanding. However in addition to modeling theories about the brain and its functions are important. A model is nothing else than a vast piece of software that captures the data as closely as possible. Beyond that we have to understand the data. Just like playing with a mathematical formula that captures physical reality generates understanding of this reality, we will have to run lots of simulations on the model to understand what each neuron, circuit, functional region is contributing to the behavior that we are observing.

9.7 First impetus

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

1. To strengthen the research basis, support of systems neuroscience proposals in primates with as hallmark novel and incisive designs. As mentioned above the most performing European laboratories easily match the competition of the American of Japanese labs. We need more of them, covering all aspects of neuroscience. So we should encourage young researchers, who often have been in the US, to set up their own independent group. We may have to device 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. This should however remain the minority as Europe has been wasting huge amounts of money supporting systems studies in lower animals while the same questions were addressed in primates in the US or Japan. This is one of the main reasons pharmaceutical research is leaving Europe for the US. The main requirement would be for the studies to introduce new more realistic designs, such as those the Japanese have been introducing: use of a tool, monkeys taking the elevator (in virtual reality) etc, or tackling in more incisive ways different cognitive functions.

To encourage such proposals, flexible mechanisms should be devised, perhaps on a continuous basis rather than a single call.

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 favor proposals linking neurophysiological teams with SMEs e.g. Thomas recording in Germany or with engineering groups. Here the EU could play a role of catalyst in bringing these groups together. Also proposals that favor the understanding of the blood supply and other physiological requirements of the brain, should be welcome. Again it is imported to realize that in rodents most of these problems do not occur or have largely be solved, and that we need proposals targeting the primate brain. As these multi-electrodes are introduced we should support the development of software to record, display and analyze this wealth of data.

3. To foster the integration of different non-invasive imaging techniques in the primate, notably fMRI, EEG and MEG. This integration is not yet solved in humans and even if it were there is no obvious way to validate the solution in humans. In non-human primates the verification is easy: compare the generators (or fields of 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. Also progress in these non-invasive techniques allows reducing the number of animals in research.

   The particular problem will be to find mechanism for the European groups that are performing monkey fMRI to acquire expensive equipment required for MEG. The creation of EU sponsored centers of excellence of a particular type, in which the neuroscience team is linked with companies producing the equipment and functions as a testsite, is a possible mechanism. Again we are envisioning projects linking academia with industry but under close EU patronage to guarantee the long-term perspective over the short-term view traditionally adopted by industry. This strategy might prove particularly interesting to foster other developments, which are also technology-driven, e.g. scanning of a walking person.

4. Increase the awareness in mathematical circles for the need of neuro-mathematics. This will take time but the sooner we start the better. The traditional way would be to call workshops and symposia. It is not clear that this will work. Sustained support of brilliant

individuals, who are interested in building bridges and attract mathematicians to at least understand their questions and problems, may be a way forward.

9.8 Ethical problems

Although the 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. While is may true that only 20The need for primate research is not always well understood by the general public. 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. The general public should realize that pressing immediately for applied research is generally a waste of money leading if any to ad hoc, non robust solutions.
2. The distance between basic biological research and clinically relevant medical research is long in general but especially so in brain research due to the immense complexity of the brain.
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. It is clear that even the primate model is not perfect (since the monkey brain is too small compared to that of humans). The adequacy of the primate model can now be handled efficiently since imaging allows addressing the homology between the brains of different species.

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.

9.9 Acknowledgements

The author is indebted to Prof. G. Rizzolatti for contributing to the motor system section.

References

A., N. M. (2003). Brain-machine interfaces to restore motor function and probe neural circuits. *Nat. Rev. Neurosci.*, *4*, 417-422.

Aleksander, I. (2005). *The world in my mind.* Exeter, UK: Imprint Academic.

Aleksander, I., & Dunmall, B. (2003). Axioms and tests for the presence of minimal consciousness in agents i: Preamble. *Journal of Consciousness Studies*, *10(4-5)*, 7-18.

Aleksander, I., & Morton, H. B. (2003). A digital neural model of visual deficits in Parkinson's disease. In *Proceedings of iwann03.* Minorca.

Amit, D. J., & Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex*, *7*, 237-252.

Amit, D. J., & Tsodyks, M. V. (1991). Quantitative study of attractor neural network retrieving at low spike rates: I. substrate - spikes, rates and neuronal gain. *Network*, *2*, 259-273.

Anafocus. (2005). *Anafocus website.* http://www.anafocus.com/.

Anderson, J. A. (1995). *An introduction to neural networks.* MIT Press.

Asaad, W., Rainer, G., & Miller, E. (2000). Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology*, *84*, 451-459.

various authors. (2001). Network: computation in neural systems, special issue.

B., P., Pezaris, J., Sahani, M., Mitra, P., & Andersen, R. (2002). Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nature Neuroscience*, *5*, 805-811.

Baars, B. (1997). *In the theater of consioucness.* New York: Oxford University Press.

Baars, B. (2002). The conscious access hypothesis: origins and recent evidence. *Trends Cognit. Sci.*, *6*, 47-52.

Bartels, A., & Zeki, S. (2000). The architecture of the colour centre in the human visual brain: new results and a review. *Europena Journal of Neuroscience*, *12*, 172-193.

Belliveau, J., Kennedy, J., D.N., McKinstry, R., Buchbinder, B., Weisskoff, R., Cohen, M., Vevea, J., Brady, T.J., & Rosen, B. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging ? *Science*, *254*, 716-719.

Block, N. (1995). On a confusion about a function of consciousness. *Behav. Brain Sci.*, *18*, 227-287.

Boahen, K. (2000). Point-to-point connectivity between neuromorphic chips using address events. *IEEE Trans. Circuits Syst. II, AnalogDigit Signal Process.*, *47(5)*, 416-434.

Boahen, K. (2002). A retinomorphic chip with parallel pathways: Encoding on, off, increasing, and decreasing visual signals. *Journal of Analog Integrated Circuits and Signal Processing*, *30*, 121-135.

Boahen, K. (2004). A burst-mode word-serial address-event channel-iii: Analysis and testing. *IEEE Transactions on Circuits and Systems I*, *51*, 1292-1300.

Bojarczuka, C., Lopesb, H., & Freitasc, A. (2001). *Data mining with constrained-syntax genetic programming: Applications, paper presented to medical data set, intelligent data analysis in medicine and pharmacology.* London, U.K.: Medinfo 2001.

Bower, J. M., & Beeman, D. (1998). *The book of GENEIS: exploring realistic neural models with the GEneral NEural SImulation sytem.* New York: Springer Verlag.

Bradley, D., Chang, G., & Andersen, R. (1998). Encoding of three-dimensional structure-from-motion by primate area mt neurons. *Nature*, *392*, 714-717.

Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, *2*(1), 14-23.

Brooks, R. A. (2002). *Robot: The future of flesh and machines.* London: Penguin Books.

Büchel, C., & Friston, K. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex*, *7*, 768-778.

Büchel, C., & Friston, K. (1998). Dynamic changes in effective connectivity characterized by variable parameter regression and kalman filtering. *Human Brain Mapping*, *6*, 403-408.

Bullier, J. (2001). Feedback connections and conscious vision. *Trends in Cognitive Science*, *5*, 369-370.

Bullmore, E., Horwitz, B., Honey, G., Brammer, M., Williams, S., & Sharma, T. (2000). How good is good enough in path analysis of fMRI data? *NeuroImage*, *11*, 289-301.

Buttazzo, G. (2001). Artificial consciousness: Utopia or real possibility. *Spectrum IEEE Computer*, *18*, 24-30.

Calabretta, R., Nolfi, S., Parisi, D., & Wagner, G. P. (2000). Duplication of modules facilitates the evolution of functional specialization. *Artificial Life*, *6*, 69-84.

Casti, A., Omurtag, A., Sornborger, A., Kaplan, E., Knight, B., Victor, J., & Sirovich, L. (2002). A population study of integrate-and-fire-or-burst neurons. *Neural Computation*, *14*, 957-986.

Chaitin, G. (1975). Randomness and mathematical proof. *Scientific American*, *232(5)*.

Chalmers, D. J. (1996). *The conscious mind: in search of a fundamental theory.* Oxford University Press.

Chapin, J. K., Moxon, K. A., Markowitz, R. S., & Nicolelis, M. (1999). Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature Neuroscience*, *2*, 664-670.

Chicca, E., Badoni, D., Dante, V., D'Andreagiovanni, M., Salino, G., Carota, L., Fusi, S., & del Giudice, P. (2003). A VLSI recurrent network of integrate-and-fire neurons connected by plastic synapses with long-term memory. *IEEE Transactions Neural Networks*, *14*, 1297-1307.

Chicca, E., Lichtsteiner, P., Delbruck, T., Indiveri, G., & Douglas, R. J. (2006). Modeling orientation selectivity using a neuromorphic multi-chip system. In *Proceedings of IEEE international symposium on circuits and systems.*

Corbetta, M., Akbudak, E., Conturo, T., Snyder, A., Ollinger, J., Drury, H., Linenweber, M., Petersen, S., Raichle, M., Van Essen, D., & Shulman, G. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*, 761-773.

Corbetta, M., & Shuman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.*, *3*, 201-215.

Cotterill, R. (1989). *No ghost in the machine.* London: Heinemann.

Courtney, S., Petit, L., Maisog, J., Ungerleider, L., & Haxby, J. (1998). An area specialized for spatial working memory in human frontal cortex. *Science*, *279*, 1347-1351.

Crick, F. (1994). *The astonishing hypothesis.* London: Simon and Schuster.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Sem Neurosci.*, *2*, 263-275.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, *8*, 97-107.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6(2)*, 119-126.

Dale, A., Liu, A., Fischl, B., Buckner, R., Belliveau, J., Lewine, J., & Halgren, E. (2000). Dynamic statistical parametric mapping: combining fmri and meg for high-resolution imaging of cortical activity. *Neuron*, *26*, 55-67.

Damasio, A. (1994). Descartes error.

Delbruck, T. (1993). Silicon retina with correlation-based, velocity-tuned pixels. *IEEE Trans. Neural Net.*, *4*, 529-541.

Delbruck, T., & Liu, S. C. (2004). A silicon visual system as a model animal. *Vision Research*, *44(17)*, 2083-2089.

Delorme, A., Gautrais, J., van Rullen, R., & Thorpe, S. (1999). SpikeNET: A simulator for modeling large networks of integrate and fire networks. *NeuroComputing*, *26*, 989-996.

Dennett, D. (2003). *Freedom evolves.* Viking Books.

Dennett, D. C. (1991). *Consciousness explained.* Boston: Little Brown and Co.

D'Esposito, M., Detre, J., Alsop, D., Shin, R., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature*, *378*, 279-281.

Diáz, J., E., R., Sbatini, S., Solari, F., & Mota, S. (2005). A phase-based stereo system-on-chip. *Biosystems*, in press.

Dickmanns, E. D. (2002). Expectation-based, multi-focal, saccadic (ems) vision for ground vehicle guidance. *Control Engineering Practice*, *10*(8), 907-15.

Diogo, A., Soares, J., Koulakov, A., Albright, T., & Gattass, R. (2003). Electrophysiological imaging of functional architecture in the cortical middle temporal visual area of cebus apella monkey. *Journal of Neuroscience*, *23*, 3881-3898.

Djurfeldt, M., Johansson, C., Ekeberg, Ö., Rehn, M., Lundqvist, M., & Lansner, A. (2006). Maasively parallel simulation of a brain-scale neuronal network. *Royal Institue of Computer Science and Communication,TRITA-NA-P0513*.

Dobelle. (2005). *Dobelle website.* http://www.dobelle.com.

Donoghue, J. (2002). Connecting cortex to machines: recent advances in brain interfaces. *Nature Neuroscience (supplement)*, *5*, 1085-1088.

Douglas, R. J., Martin, K. A. C., & Whitteridge, D. (1989). A canonical mircrocircuit for neo-cortex. *Neural Computation*, *1*, 119-130.

Dreyfuss, H. (1972). *What computers can't do, a critique of artificial reason.* New York: Harper and Row.

Dubowitz, D., Chen, D., Atkinson, D., Grieve, K., Gillikin, B., Bradley, J., W.G., & Andersen, R. (1998). Functional magnetic resonance imaging in macaque cortex. *NeuroReport*, *9*, 2213-2218.

Dupont, P., Orban, G., De Bruyn, B., Verbruggen, A., & Mortelmans, L. (1994). Many areas in the human brain respond to visual motion. *Journal of Neurophysiology*, *72*, 1420-1424.

Dupont, P., Orban, G., Vogels, R., Bormans, G., Nuyts, J., Schiepers, C., De Roo, M., & Mortelmans, L. (1993). Different perceptual tasks performed with the same visual stimulus attribute activate different regions of the human brain: A positron emission tomography study. *Proceedings of the National Academy of Sciences of the USA*, *90*, 10927-10931.

Edelman, G. (1992). *Bright air, brilliant fire: On the matter of the mind.* Basic Books.

Edelman, G., & Tononi, G. (1998). Consciousness and complexity. *Science*, *282*, 1846-1851.

Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness. how matter becomes imagination.* London.

Eggert, J., & van Hemmen, J. L. (2001). Modeling neural assemblies: Theory and implementation. *Neural Computation*, *13*, 1923-1974.

European Commission, R. (1998). *http://europa.eu.int/comm/research/fp6/pdf/faq_en.pdf.* Prepared for the OECD megascience forum.

Felleman, D., & Van Essen, D. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1-47.

Ferrera, V., Nealey, T., & Maunsell, J. (1994). Responses in macaque visual area v4 following inactivation of the parvocellular and magnocellular lgn pathways. *Journal of Neuroscience*, *14*, 2080-2088.

Fias, W., Dupont, P., Reynvoet, B., & Orban, G. A. (2002). The quantitavtive nature of a visual task differentiates between ventral and dorsal stream. *Journal of Cognitive Neuroscience*, *2*, 661-670.

Fize, D., Vanduffel, W., Nelissen, K., Denys, K., Chef DH́otel, C., Faugeras, O., & Orban, G. (2003). The retinotopic organization of primate dorsal v4 and surrounding areas: an fmri study in awake monkeys. *Journal of Neuroscience, in press.*

Fox, P., Mintun, M., Raichle, M., Miezin, F., Allman, J., & Van Essen, D. (1986). Mapping human visual cortex with positron emission tomography. *Nature*, *323*, 806-809.

Franklin, S. (2003). Ida, a conscious artefact? *Journal of Consciousness Studies*, *10(4-5)*, 47-66.

Freeman, W. (1975). *Mass action in the nervous system.* Academic Press.

Freeman, W. (2000). *Neurodynamics.* London: Springer-Verlag.

Friston, K. (1997). *Analysing brain images: principles and overview. in: Human brain function (frackowiak r.s.j., friston, k.j., frith, c.d., dolan, r.j. and mazziotta, j.c. eds.), pp. 25-41.* San Diego: Academic Press.

Friston, K., & Büchel, C. (1998). A nonlinear analysis of functional brain architectures. *NeuroImage*, *7*, S776.

Friston, K., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage, in press*.

Funes, P., & Pollack, J. (1999). Computer evolution of buildable objects. In P. Bentley (Ed.), *Evolutionary design by computers* (p. 387-403). Morgan Kaufmann.

Fuster, J., & Alexander, G. (1971). Neuron activity related to short-term memory. *Science*, *173*, 652-654.

Galea, M., & Darian-Smith, I. (1994). Multiple corticospinal neuron populations in the macaque monkey are specified by their unique cortical origins, spinal terminations, and connections. *Cerebral Cortex*, *4*, 166-194.

Gallagher, S., & (Editors), J. S. (1999). *Models of the self.* Thorverton, UK: Imprint Academic.

Gallese, V., & Metzinger, T. (2003). Motor ontology: The representational reality of goals, actions and selves. *Philosophical Psychology*.

Gerstner, W. (1995). Time structure of the activity in neural network models. *Physical Review E*, *51*, 738-758.

Gerstner, W., & van Hemmen, J. L. (1992). Associative memory in a network of 'spiking' neurons. *Network: Computation in Neural Systems*, *3*, 139-164.

Gibbs, W. W. (2004). A split at the core. *Scientific American Magazine*, *November*.

Girard, P., Lomber, S., & Bullier, J. (2002). Shape discrimination deficits during reversible deactivation of area v4 in the macaque monkey. *Cerebral Cortex*, *12*, 1146-1156.

Goddard, N. H., Hucka, M., Howell, F., Cornelis, H., Shankar, K., & Beeman, D. (2001). Towards neuroml: Model description methods for collaborative modelling in neuroscience. *Philosophic Transactions of the Royal Society B 356*, *1412*, 1209-1228.

Goodale, M., & Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *1*, 20-25.

Gottlieb, J., Kusunoki, M., & Goldberg, M. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, *391*, 481-484.

Grinvald, A., Lieke, E., Frostig, R., & Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *Journal of Neuroscience*, *14*, 2545-2568.

Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Consciousness and Cognition*, *8*, 1-44.

Hadjikhani, N., Liu, A., Dale, A., Cavanagh, P., & Tootell, R. (1998). Retinotopy and color sensitivity in human visual cortical area v8. *Nature Neuroscience*, *1*, 235-241.

Haikonen, P. O. (2003). *The cogntive approach to conscious machine.* London: Imprint Academic.

Hameroff, S. (1997). Quantum computing in microtubules: an intra-neural correlate of consciousness? *Cognit Stud Bull Jpn Cognit Sci Soc*, *4(3)*, 67-92.

Hansel, D., Mato, G., Meunier, C., & Neltner, L. (1998). On numerical simulations of integrate-and-fire neural networks. *Neural Computation*, *10*, 467-483.

Harnad, S. (2003). Can a machine be conscious? how? *Journal of Consciousness Studies*.

Haxby, V., Horwitz, B., Ungerleider, L., Maisog, J., Pietrini, P., & Grady, C. (1994). The functional organization of human extrastriate cortex: a pet-rcbf study of selective attention to faces and locations. *Journal of Neuroscience*, *14*, 6336-6353.

Hayward, J. (1999). A dzogs-chen buddhist interpretation of the sense of self. In S. Gallagher & J. Shear (Eds.), *Models of the self* (p. 379-395). Thorverton, UK: Imprint Academic.

He, S., Dum, R., & Strick, P. (1993). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere. *Journal of Neuroscience*, *13*, 952-980.

He, S., Dum, R., & Strick, P. (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *Journal of Neuroscience*, *15*, 3284-3306.

Heeger, D., Boynton, G., Demb, J., Seidemann, E., & Newsome, W. (1999). Motion opponency in visual cortex. *Journal of Neuroscience*, *19*, 7162-7174.

Hines, M., & Carnevale, N. T. (1997). The neuron simulation environment. *Neural Computation*, *9*, 1179-1209.

Hochberg, L. R., Mukand, J. A., Polykoff, G. I., Friehs, G. M., & Donoghue, J. P. (2005). Braingate neuromotor prosthesis: nature and use of neural control. In *Program no. 520.17.* Washington DC: Society for Neuroscience.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation on nerve. *Journal of Physiology*, *117*, 500-544.

Hoffman, K., & McNaughton, B. (2002). Coordinated reactivation of distributed memory traces in primate neocortex. *Science*, *297*, 2070-2073.

Holland, J. (1975). *Adaptation in natural and artificial systems.* Ann Arbor, Michigan: University of Michigan Press.

van Horn, J. D., Grethe, J. S., Kostelec, P., Woodward, J. B., Aslan, J. A., Rus, D., Rockmore, D., & Gazzaniga, M. S. (2001). The Functional Magnetic Resonance Data Center (fMRIDC): the challenges and reward of large-scale databasing of neuroimaging studies. *Philosophic Transaction of the Royal Society of London B*, *356*, 1323-1337.

IBM. (2006). *http://www.research.ibm.com/bluegene.*

Implants. (2005). *Implants website.* http://digitial-implants.com.

Indiveri, G. (1999). Neuromorphic analog vlsi sensor for visual tracking: Circuits and application examples. *IEEE Transactions on Circuits and Systems II, Analog and Digital Signal Processing*, *46(11)*, 1337-1347.

Indiveri, G., Chicca, E., & Douglas, R. (2006). A vlsi array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity. *IEEE Transactions on Neural Networks*, *17(1)*, 211-221.

Indiveri, G., & Douglas, R. (2000). Robotic vision neuromorphic vision sensor. *Science*, *288*, 1189-1190.

ITRS. (2005). *International technology roadmap for semiconductors.* http://public.itrs.net.

Jennings, C. (2000). In search of consciousness. *Nature Neuroscience*, *3(8)*, 1-1.

J.S., J. (1990). *Neural network design and the complexity of learning.* Cambridge, MA: MIT Press.

Kamimura, A., Murata, S., Yoshida, E., Kurokawa, H., Tomita, K., & Kokaji, S. (2001). Self-reconfigurable modular robot – experiments on reconfiguration and locomotion. In *Proc. ieee/rsj int. conf. on intell. robots and systems* (pp. 606 – 612). IEEE Press.

de Kamps, M., & van der Velde, F. (2001). From artificial neural networks to spiking populations of neurons and back again. *Neural Networks*, *14*, 941-953.

de Kamps, M., & van der Velde, F. (2006). Neural blackboard architectures: the realization of compositionality and systematicity in neural networks. *Journal of Neural Engineering*, *3*, R1-R12.

Kauffman, S. (1993). *The origins of order.* Oxford University Press.

Keizer, K., & Kuypers, H. (1989). Distribution of corticospinal neurons with collaterals to the lower brain stem reticular formation in monkey (macaca fascicularis). *Experimental Brain Research*, *74*, 311-318.

Kitano, H. (2003). Computational systems biology. *Nature*, *420*, 205.

Knight, B. W., Manin, D., & Sirovich, L. (1996). Dynamical models of interacting neuron populations in visual cortex. *Symposium on Robotics and Cybernetics: Computational Engineering in Systems Applications, E. C. Gerf, France: Cite Scientifique*.

Knoll, A. (2002). Cui bono robo sapiens? *Autonomous Robots*, *12*, 5-12.

Koza, J. (1992). *Genetic programming, on the programming of computers by means of natural selection.* Cambradge, MA: MIT Press.

Kramer, R., J.; Sarpeshkar, & Koch, C. (1997). Pulsed-based analog vlsi velocity sensors. *IEEE Trans. Circuits and Systems II*, *44*, 86-101.

Kwong, K., Belliveau, J., Chesler, D., Goldberg, I., Weiskoff, R., Poncelet, B., Kennedy, B., D.N.and Hoppel, Cohen, M., & Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Science of the USA*, *89*, 5675-5679.

L., S., Reber, P., Costanza, J., Wong, E., Buxton, R., Zola, S., Squire, L., & Albright, T. (1998). fmri of monkey visual cortex. *Neuron*, *20*, 1051-1057.

Lebedev, M. A., Carmena, J. M., O'Doherty, J. E., Zackenhouse, M., Henriquez, C. S., Principe, J. C., & Nicolelis, M. A. (2005). Cortical ensemble adaptation to represent velocity of an artificial actuator controlled by a brain-machine interface. *J. Neurosc.*, *25(19)*, 4681-4693.

Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pac. Phil. Quart.*, *64*, 354-361.

Lichtsteiner, P., Posch, C., & Delbruck, T. (2006). A 128×128 120db 30mw asynchronous vision sensor that responds to relative intensity change. *ISSCC Digest of Technical Papers*, 508-509.

Liu, S. C., & Douglas, R. (2004). Temporal coding in a silicon network of integrate-and-fire neurons. *IEEE Transactions on Neural Networks*, *15(5)*, 1305-1314.

Liu, S. C., & van Schaik, A. (2005). An aer ear. In S. B. (Eds.) (Ed.), *The neuromorphic engineering newsletter* (Vol. 2(1), p. 11-12). Neuromorphic Engineering Newsletter.

Logothetis, N., Guggenberger, H., Peled, S., & Pauls, J. (1999). Functional imaging of the monkey brain. *Nature Neuroscience*, *2*, 555-562.

Logothetis, N., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fmri signal. *Nature*, *412*, 150-157.

Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1993). Corticocortical connections of area f3 (sma-proper) and area f6 (pre-sma) in the macaque monkey. *Journal of Computational Neurology*, *338*, 114-140.

Luppino, G., & Rizzolatti, G. (2000). The organization of the frontal motor cortex. *News in Physiological Sciences*, *15*, 219-224.

Maass, W. (1997). Fast sigmoidal networks via spiking neurons. *Neural Computation*, *9*, 279-304.

Maass, W., & Bishop, C. M. (1998). *Pulsed neural networks.* Cambridge (MA): MIT Press.

Mahowald, M., & Delbruck, T. (1989). Cooperative stereo matching using static and dynamic image features. 213-238.

Mandeville, J., & Marota, J. (1999). Vascular filters of functional mri: spatial localization using bold and cbv contrast. *Magnetic Resonance in Medicine*, *42*, 591-598.

Manzotti, R., & Glatzeder, B. (2003). *An ontogenetic-based ontology for consciousness.* Copenhagen: ISI 2003.

Manzotti, R., & Tagliasco, V. (2001). *On building a conscious being. in search of a science of consciousness.* Skovde (Sweden).

Margulis, L. (1998). *Symbiotic planet.*

Mataric, M. (1992). Integration of representation into goal-driven behavior-based robots. *IEEE Transactions on Robotics and Automation*, *8*(3), 304-12.

Matelli, M., Camarda, R., Glickstein, M., & Rizzolatti, G. (1986). Afferent and efferent projections of the inferior area 6 in the macaque monkey. *Journal of Computational Neurology*, *251*, 281-298.

Matsumura, M., & Kubota, K. (1979). Cortical projection of hand-arm motor area from postarcuate area in macaque monkey: a histological study of retrograde transport of horseradish peroxidase. *Neuroscience Letters*, *11*, 241-246.

Mattia, M., & Giudice, P. D. (2000). Efficient event-driven simulation of large networks of spiking neurons and dynamical synapses. *Neural Computation*, *12*, 2305-2329.

McCarthy, J. (1995). *Making robot conscious of their mental states. machine intelligence. s. muggleton.* Oxford: Oxford University Press.

McCullough, W., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, *5*, 115-123.

McIntosh, A., Grady, C., Ungerleider, L., Haxby, J., Rapoport, S., & Horwitz, B. (1994). Network analysis of cortical visual pathways mapped with pet. *Journal of Neuroscience*, *14*, 655-666.

McKibben, B. (1990). *The end of nature.* Penguin.

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of of visual working memory in prefrontal cortex of macaque. *Journal of Neuroscience*, *16*, 5154-5167.

Mitchell, M. (1996). *An introduction to genetic algorithms.* Cambridge, MA: MIT Press.

Mota, S., Ros, E., Origosa, E. M., & Pelayo, F. J. (2004). Bio-inspired motion detection for blind spot overtaking monitor. *International Journal of Robotics and Automation*, *19(4)*, 190-196.

Muakkassa, K., & Strick, P. (1979). Frontal lobe inputs to primate motor cortex: evidence for four somatotopically organized ' premotor ' areas. *Brain Research*, *177*, 176-182.

Nagel, T. (1974). What is it like to be a bat? *Phil. Rev.*, *83*, 434-450.

Nakahara, K., Hayashi, T., Konishi, S., & Miyashita, Y. (2002). Functional mri of macaque monkeys performing a cognitive set-shifting task. *Science*, *295*, 1532-1536.

N. Fragopanagos, S. K., & Taylor, J. (2005). Modeling the attentional blink. *Cog. Brain. Res.*, (in press).

Nieder, A., Freedman, D., & Miller, E. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, *297*, 1708-1711.

Nolfi, S., & Floreano, D. (2000). *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines (intelligent robots and autonomous agents).* Bradford Book.

Nolfi, S., & Parisi, D. (1995). *Evolving artificial neural networks that develop in time.* European Conference on Artificial Life.

Nykamp, D. Q., & Tranchina, D. (2000). A population density approach that facilitates large-scale modeling of neural networks: Analysis and an application to orientation tuning. *Journal of Computational Neuroscience*, *8*, 19-50.

Ohno, S. (1970). *Evolution by gene duplication.* Springer.

Omurtag, A., Knight, B. W., & Sirovich, L. (2000). On the simulation of large populations of neurons. *Journal of Computational Neuroscience*, *8*, 51-63.

Orban, G. (2002). Functional mri in the awake monkey: the missing link. *Journal of Cognitive Neuroscience*, *14*, 965-969.

Orban, G., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., J., T., & Vanduffel, W. (2003). Similarities and differences in motion processing between the human and macaque brain: evidence from fmri. *Neuropsychologia, in press.*

Orban, G., Saunders, R., & Vandenbussche, E. (1995). Lesions of the superior temporal cortical motion areas impair speed discrimination in the macaque monkey. *European Journal of Neuroscience*, *7*, 2261-2276.

Orban, G., Sunaert, S., Todd, J., Van Hecke, P., & Marchal, G. (1999). Human cortical regions involved in extracting depth from motion. *Neuron*, *24*, 929-940.

O' Regan, K., & Noe, A. (2001). A sensorimotor account of visual perception and consciousness. *Behavioral and Brain Sciences*, *24*, 939-939.

OReilly, C., & Munakata, Y. (2000). *Computational exploration in cognitive neuroscience.* MIT Press.

Oya, T., Asai, T., Kagaya, R., & Amemiya, Y. (2005). Noise performance of single-electron depressing synapses for neuronal syncrony detection. *International Joint Conference on Neural Networks*.

Paillard, J. (Ed.). (1991). *Brain and space.* Oxford: Oxford University Press.

Paulesu, E., Frith, C., & Frackowiak, R. (1993). The neural correlates of the verbal component of working memory. *Nature*, *362*, 342-344.

Penrose, R. (1994). *Shadows of the mind.* Oxford: Oxford Univ. Press.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences*.

Petersen, S., Fox, P., Posner, M., Mintun, M., & Raichle, M. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, *331*, 585-589.

Pouget, A., & Snyder, L. (2000). Computational approaches to sensorimotor transformations. *Nature Neuroscience*, *3*, 1192-1198.

Price, C., & Friston, K. (1997). Cognitive conjunction: A new approach to brain activation experiments. *NeuroImage*, *5*, 261-270.

project, T. H. (2003). www.hydra-robot.com.

Rainer, G., Rao, S., & Miller, E. (1999). Prospective coding for objects in the primate prefrontal cortex. *Journal of Neuroscience*, *19*, 5493-5505.

Rao, S., Rainer, G., & Miller, E. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, *276*, 821-824.

Rechenberg, I. (1965). *Cybernetic solution path of an experimental problem.* Ministry of Aviation, Royal Aircraft Establishment, U.K.

Rees, G., & G. Kreiman, e. a. (2002). Neural correlates of consciousness in humans. *Nature Reviews*, *3*, 261-270.

Reeve, R., Webb, B., Horchler, A., Indiveri, G., & Quinn, R. (2005). New technologies for testing a model of cricket phonotaxis on an outdoor robot platform. *Robotics and Autonomous Systems*, *51(1)*, 41-54.

Reichel, L., Liechti, D., Presser, K., & Liu, S.-C. (2005). Range estimation on a robot using neuromorphic motion sensors. *Robotics and Autonomous Systems*, *51 (2-3)*, 167-174.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews of Neuroscience*, *2*, 661-670.

Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the cortical motor system: new concepts. *Electroencephalography and Clininical Neurophysioly*, *106*, 283-296.

Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function.* Oxford: Oxford University Press.

Ros, E., Ortigosa, E., Agis, R., Carrillo, R., & Arnold, M. (2005). Real-time computing platform for spiking neurons (rt-spike). *IEEE Transactions on Neural Networks*, (in press).

Rosier, A., Cornette, L., Dupont, P., Bormans, G., Michiels, J., Mortelmans, L., & Orban, G. (1997). Positron-emission tomography imaging of long-term shape recognition challenges. *Proceedings of the National Academy of Sciences of the USA*, *94*, 7627-7632.

Rowe, J., Toni, I., Josephs, O., Frackowiak, R., & Passingham, R. (2000). The prefrontal cortex: Response selection or maintenance within working memory ? *Science*, *288*, 1656-1660.

Ryle, G. (1952). *The concept of mind.* London: Hutchinson.

Saleem, K., Pauls, J., Augath, M., Trinath, T., Prause, B., Hashikawa, T., & Logothetis, N. (2002). Magnetic resonance imaging of neuronal connections in the macaque monkey. *Neuron*, *34*, 685-700.

Salzman, C., Murasugi, C., Britten, K., & Newsome, W. (1992). Microstimulation in visual area mt: effects on direction discrimination performance. *Journal of Neuroscience*, *12*, 2331-2355.

Schall, J. (1995). Neural basis of saccade target selection. *Review of Neuroscience*, *6*, 63-85.

Schall, J., Morel, A., King, D., & Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams. *Journal of Neuroscience*, *15*, 4464-4487.

Schoenauer, T., Atasoy, S., Mehrtash, N., & Klar, N. (2002). Neuropipe-chip: A digital neuroprocessor for spiking neural networks. *IEEE Transactions on Neural Networks*, *13*, 205-213.

Schwartz, A. B. (2004). Cortical neural prosthetics. *Annu. Rev. Neurosc.*, *27*, 417-422.

Schwefel, H.-P. (1981). *Numerical optimization of computer models.* Chichester: Wiley & Sons.

Science as a vocation. (1919). In H. H. Gerth & C. . Wright Mills (Eds.), *Essays in sociology.*

Seiingwithsound. (2005). *Seeingwithsound website.* http://www.seeingwithsound.com.

Sereno, M., Dale, A., Reppas, J., Kwong, K., Belliveau, J., Brady, T., Rosen, B., & Tootell, R. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889-893.

Serrano-Gotarredona, R. e. a. (2005). Aer building blocks for multi-layer multi-chip neuromorphic vision systems. In Y. W. Dec, B. Schölkopf, & J. P. (Eds.) (Eds.), *Advances in neural information processing systems (nips)* (Vol. 18, p. 1217-1224). MIT Press.

Serruya, M. D., Hatsopoulos, N. G., Paninski, L., Fellows, M. R., & Donoghue, J. P. (2002). Instant neural control of a movement signal. *Nature*, *416*, 141-142.

Shipp, S., Watson, J., Frackowiak, R., & Zeki, S. (1995). Retinotopic maps in human prestriate visual cortex: the demarcation of areas v2 and v3. *NeuroImage*, *2*, 125-132.

Sipper, M., Mange, D., & Sanchez, E. (1999). Quo vadis evolvable hardware? *Commun. ACM*, *42*, 50-56.

S.J., G. (2002). *The origins of evolutionary theory.* Harvard University Press.

Sloman, A., & Chrisley, R. (2003). Virtual machines and consciousness. *Journal of Consciousness Studies*, *10(4-5)*, 133-172.

Smith, E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657-1661.

Sober, & Wilson. (1998). *Unto others, the evolution and psychology of unselfish behavior.* Harvard University Press.

Squire, L. (1986). Mechanisms of memory. *Science*, *232*, 1612-1619.

Stephan, K. E., Kamper, L., Bozkurt, A., Burns, G. A. P. C., Young, M. P., & Koetter, R. (2001). Advanced methodology for the collation of connectivity data on the macaque brain (CoCoMac). *Philosophic Transactions of the Royal Society B 356*, *1412*, 1209-1228.

Stocker, A. (2002). An improved 2d optical flow sensor for motion segmentation. In *International circuits and systems conference (ISCAS2002)* (Vol. 2, p. 332-335).

Strawson, G. (1999). The self and the sesmet. *Journal of Consciousness Studies*, *6*, 99-135.

Száthmary, E., & Maynard Smith, J. (1995). *The major transitions in evolution.* Oxford: Freeman.

Tani, J. (1998). An interpretation of the 'self' from the dynamical systems perspective: A constructivistic approach. *Journal of Consciousness Studies*, *5-6*.

Taylor, D. M., Helms-Tillery, S. I., & Schwartz, A. B. (2002). Direct cortical control of 3d neuroprosthetic devices. *Science*, *296*, 1829-1832.

Taylor, J. G. (1999). *The race for consciousness.* Cambridge MA: MIT Press.

Taylor, J. G. (2002a). From matter to mind. *Journal of Consciousness Studies*, *9(4)*, 3-22.

Taylor, J. G. (2002b). Paying attention to consciousness. *Trends in Cognitive Sciences*, *6(5)*, 206-210.

Taylor, J. G. (2003). Paying attention to consciousness. *Progress of Neurobiology*, *71*, 305-335.

Taylor, J. G. (2004). CHIMERA: Creating a new generation chip by brain guidance. In *Ieee internation conference on neural networks - conference proceedings* (p. 1989-1994).

Taylor, J. G. (2005). Mind and consciousness: Towards a final answer? *Physics of Life Reviews*, *2*, 1-45.

Taylor, J. G., & Fragopanagos, N. (2005). The interaction of attention and emotion. *Neural Networks*, *18(4)*, (in press).

Thompson, A. (1996). Silicon evolution. In J. K. et al. (Ed.), *Genetic programming 1996: Proceedings of the first annual conference (gp96)* (p. 640-656). Springer Verlag.

Thompson, A. (1997). An evolved circuit, intrinsic in silicon, entwined with physics. In *1st int. conf. on evolvable systems.* Springer Verlag.

Thompson, A. (2004). *Adrian thompson's hardware evolution page.* www.cogs.susx.ac.ukusersadrianthade.html.

Tolias, A., Smirnakis, S., Augath, M., Trinath, T., & Logothetis, N. (2001). Motion processing in the macaque: Revisited with functional magnetic resonance imaging. *Journal of Neuroscience*, *21*, 8594-8601.

Tootell, R., Hamilton, S., Silverman, M., & Switkes, E. (1988). Functional anatomy of macaque striate cortex. i. ocular dominance, binocular interactions, and baseline conditions. *Journal of Neuroscience*, *8*, 1500-1530.

Tootell, R., Mendola, J., Hadjikhani, N., Ledden, P., Liu, A., Reppas, J., Sereno, M., & Dale, A. (1997). Functional analysis of v3a and related areas in human visual cortex. *Journal of Neuroscience*, *17*, 7060-7078.

Tremblay, L., Hollerman, J., & Schulz, W. (1998). Modifications of reward expectation-related neuronal activity during learning inprimate striatum. *Journal of Neurophysiology*, *80*, 964-977.

Tsao, D., Conway, B., & Livingstone, M. (2003). Receptive fields of disparity-tuned simple cells in macaque v1. *Neuron*, *38*, 103-114.

Ungerleider, L., & Mishkin, M. (1982). *Two cortical visual systems. in: The analysis of visual behavior (d.j. ingle, r.j.w. mansfield and m.s. goodale, eds.).* MIT Press.

Vanduffel, W., Fize, D., Mandeville, J., Nelissen, K., Van Hecke, P., Rosen, B., Tootell, R., & Orban, G. (2001). Visual motion processing investigated using contrast-agent enhanced fmri in awake behaving monkeys. *Neuron*, *32*, 565-577.

Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J., & Orban, G. (2002). Extracting 3d from motion: Differences in human and monkey intraparietal cortex. *Science*, *298*, 413-415.

Vanduffel, W., Tootell, R., & Orban, G. (2000). Attention-dependent suppression of metabolic activity in the early stages of the macaque visual system. *Cerebral Cortex*, *10*, 109-126.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience.* Cambridge, MA: MIT Press.

van der Velde, F. (1997). On the use of computation in modelling behaviour. *Network: Computationof Neural systems*, *8*, R1-R32.

van der Velde, F., & de Kamps, M. (2001). From knowing what to knowing where: Modelling object-based attention with feedback disinhibition of activation. *Journal of Cognitive Neuroscience*, *13(4)*, 479-491.

van der Velde, F., & de Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavior and Brain Sciences*, *29(1)*, 1-72.

Walker, R., & Miglino, O. (2002). Genetic redundancy in evolving populations of simulated robots. *Artificial Life*, *8(3)*.

Watts, L., Kerns, D., Lyon, R., & Mead, C. (1992). Improved implementation of the silicon cochlea. *IEEE J. Solid-State Circuits*, *SC-27*, 692-700.

Wessberg, J., Stambaugh, C. R., Kralik, J., Beck, P. D., Laubach, M., Chapin, J. K., Kim, J., Biggs, S. J., Srinivasan, M. A., & Nicolelis, M. A. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, *408*, 361-365.

Wheeler, M., & Buckner, R. (2003). Functional dissociation among components of remembering: control, perceived oldness, and content. *Journal of Neuroscience*, *23*, 3869-3880.

Wilson, F. A. W., O'Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, *260*, 1955-1958.

Wolpert, D., & Macready, W. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, *1(1)*, 67-82.

Xiao, D., Marcar, V., Rraiguel, S., & Orban, G. (1997). Selectivity of macaque mt/v5 neurons for surface orientation in depth specified by motion. *European Journal of Neuroscience*, *9*, 956-964.

Xilinx. (2005). *Xilinx web site.* http://www.xilinx.com.

Zeki, S. (2003). The disunity of consciousness. *Trends in Cognitive Sciences (in press).*

Zell, A. (1995). Snns. *http:www-ra.informatik.uni-tuebingen.deSNNS*.

Ziemke, T. (2001). The construction of 'reality' in the robot: Constructivist perpectives on situated artificial intelligence and adaptive robotics. *Foundations of Science*, *6*, 163-233.

Zipser, D., & Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, *331*, 679-684.

Zlatev, J. (2001). The epigenesis of meaning in human beings, and possibly in robots. *Minds and Machines*, *11*, 155-195.