Roadmap of Neuro-IT Development: PHASE 1 - Collection of ideas

Edited by Alois Knoll & Marc de Kamps

Version 1.2: November 26, 2003

With contributions from:
John Hallam
Marc de Kamps
Alois Knoll
Riccardo Manzotti
Rolf Müller
Guy Orban
Herbert Peremans
Guilio Sandini
Erik de Schutter
Vincenzo Tagliasco

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 111 researchers from 82 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. The EU has expressed great interest in the creation of a Roadmap: future calls within FP6 may be based on the content of the Roadmap, and also may help to decide if, and which, new fields will be addressed in FP7. Currently, one Roadmap already exists: the Roadmap for Nanoelectronics (NE), which can be found on the web at:
`http://www.cordis.lu/ist/fetnidqf.htm`

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

[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.

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 1.2, is a draft. It contains a few new challenges, which will most likely be merged into one in a future version. This publication of this version coincides with the closing of a 'webconsultation' that was held by theEU in early October 2003. The results of this webconsultation will help determine the next steps that will be taken. Comments, proposals and criticism can be sent to `roadmap@neuro-it.net` at all times **At present we are still in the phase of collecting ideas !** At a later stage the ideas will be finalized, a time line wil be produced an interdependencies between the proposed projects will be established.

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 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.

1.6.3 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.

1.6.4 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.

1.6.5 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 ’brain interface’ project

2.1 Introduction

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.

2.2 Objectives

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.

2.3 Examples

- 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.
- 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

Despite the ongoing debate about the nature of the movement parameters represented by neuronal activity (Todorov, 2000), several recent studies investigated the possibility of predicting limb movements from the activity of multiple single-neurons in the motor cortex. After initial studies in rats (Chapin, Moxon, Markowitz, & Nicolelis, 1999), this was applied successfully to monkeys (Wessberg et al., 2000; Serruya, Hatsopoulos, Paninski, Fellows, & Donoghue, 2002; Taylor, 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. Such techniques could potentially be the basis of neuronal motor prostheses (Chapin et al., 1999; Fetz, 1999; Laubach, Wessberg, & Nicolelis, 2000; Wessberg et al., 2000; Nicolelis, 2001; Craelius, 2002; Serruya et al., 2002; Taylor, 2002; Kvnig & F., 2002; Wickelgren, 2003). A major conceptual breakthrough in this direction was achieved by two recent experiments demonstrating the feasibility of real-time brain-control of a computer cursor in two (Serruya et al. 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 brain signals recorded by the implanted electrodes, due to modulation or real plasticity.

While these animal studies are important proofs of principle, few have attempted realistic neuroprosthetic applications. The monkey study where movement of a robot arm was controlled succesfully (Wessberg et al., 2000) did not (yet) attempt to use this arm for a real task. In fact, the only published study where the robot arm was doing something is the original rat experiment

(Chapin et al., 1999). The problem in controlling robot arms or similar devices in realistic contexts is the absence of somatosensory feedback in current experimental designs. The only form of sensory feedback the animals receive is visual and this is not sufficient to grasp objects or to perform complex manipulations (Johansson, 1996; Todorov & Jordan, 2002).

Another limitation of the current approaches are the simple algorithms for spike train decoding: Taylor, Helms-Tillery, and Schwartz (2002) used cosine tuning functions to represent neuronal firing frequency as a function of movement direction, and the population vector approach (Georgopoulos A. P., 1986) to extract the population-average of the directional information contained in the activity of single neurons. Chapin et al. (1999), Wessberg et al. (2000), Serruya et al. (2002)used a linear filter for movement reconstruction, which did not make any specific assumptions about the neuronal code. Based on realistic models for the neuronal encoding of movement, however, more accurate and robust decoding algorithms can be constructed. Those will certainly be needed to control effectors with many degrees of freedom.

2.5 Problem areas

Neuroscience:

- Identifying optimal brain regions for electrode implantation to use BCI in different tasks, try to get highest information rate for low number of recording points.
- Identifying the coding strategy used in these brain areas necessary to develop corresponding representations in software systems.
- Understanding how the brain integrates sensory and motor systems both in fast motor control and in decision.
- Understanding the limitations to cortical plasticity necessary to get optimal signal transfer through the BCI.

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 in humans. Study long-term effects.
- Research on alternatives to implanted electrodes.
- 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.

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. 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.

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.

2.6 Future research

The program will focus on fully establishing BBCI technology in animal models and on developing ethical guidelines for future use in humans. Technology and materials development geared toward human application should be covered by other programs.

The goal of this challenge is to develop bidirectional bionic animals, defined as animals which use BCI both for sensory input and to interact with their environment. Interaction can be by controlling either a robot or a prosthetic device which allows the animal to move itself. In these models the brain closes the loop between computer controlled sensory input and computer driven action or motor activity. BBCI technology will first be developed in rodents and subsequently ported to application in monkeys. Application in rodents will include BBCI driven navigation and exploration while the monkey model is more appropriate to develop BBCI control of prosthetic limbs.

2.7 Immediate goals

- Identifying optimal brain regions for BCI for limb control in monkeys.
- Studying integration of somatosensory input with fast motor control in rodents or monkeys.
- 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.

- Development of real-time encoding/decoding software for brain input/output signals which is robust to noise, changes in signal quality and brain plasticity.
- Study methods for shared control versus partial autonomy in real-time brain robot interaction.
- Create interfaces between scientists, clinicians and patient organizations to define ethical standards for use of BBCI in human patients.

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. Discussion

Chapter 3

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

3.1 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.

3.2 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.

3.3 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, e.g. (?,)

- **Nanoscale self-asembling structures** were proposed that can build up aggregates of macroscopic size, e.g. for "muscle tissue" (?, ?, FHG-TEG) 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 by König ().

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.

3.4 Expected Results: What will it be good for?

In Table 3.1 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 3.1: Application areas and users of direct and spin-off results from research under the umbrella of Factor-10.

Chapter 4

Successful Thinking and Acting in the Physical World

4.1 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.

4.2 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.

4.3 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.

4.4 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.

4.5 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.

4.6 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.

4.7 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 5

Conscious machines I

5.1 Motivation and objective

The quality of the mechanical components of robots, the available sensors, and in particular computing power have increased over the past few years to an extent that today a large number of new tasks can be processed in the field of ”service robotics”: e.g. materials and tools transport in factories, delivery services in hospitals, household cleaning chores, or underwater inspections. However, the hopes placed in this new class of robots have not been fulfilled. Their acceptance falls far short of the original euphoric expectations. There are two main reasons for this:

- The adaptation of robots of this kind to tasks that deviate only slightly from their original functions is very difficult and in many cases has to be carried out by the manufacturer. This also applies to changes in the environment in which the robots work.
- Communication and cooperation with human beings on a given task is largely an unsolved problem.

Both of these things result in people perceiving these robots as being thoroughly ”dumb” and, worse yet, their services are in many cases not yet seen as being a help. This will only change if and when the fundamental deficit shown by today’s robots, i.e. their lack of *autonomy* and *adaptivity*, is eliminated.

Flexibility of an *artificial being* (artifact) with regard to structurally changing tasks presupposes cognitive functions, i.e. recognition of objects and environment, planning and control of movements and actions, learning of object characteristics and long action sequences (with sensorimotor parameters), transfer of generalizations to new situations, evaluation of situational contexts, generalization and transfer of knowledge learned under specific contextual or environmental conditions to new situations, independent development of autonomous behavior based on experience and background knowledge (analogous to the transfer of learned actions to the human cerebellum), short-term and long-term memory.

On the other hand, a key prerequisite is the adaptation of body and effector mechanisms to the environment. Their behavioral equipment should make it possible for artifacts to work in semistructured environments of the kind in which human beings move quite naturally, e.g. homes, public institutions, factories, power plants. In addition, physical adaptability is the prerequisite for being able to assume tasks which human beings cannot handle or only with considerable technical difficulty, such as in environments hostile to life, based on independent development of appropriate survival strategies. There are two possibilities for adaptation: either the use of self-organization and growth processes (in the sense of an independent adaptive strategy only ”genetically predetermined” by the designer which is based on the constraints dictated by the environment and appropriately assesses the success of adaptation) or the adaptation of mechanics using additional devices or tools. The first possibility corresponds to biological evolution over thousands of years; the second requires the passing on of knowledge based on experience regarding the use of devices from one ”generation” to the next. After all, it is conceivable that artifacts could be equipped with redundant effector mechanisms whose use would improve through practice (through additional availability of resources; this would correspond to the enlargement of areas of the brain used to control frequently used limbs). The development of adequate social and interactive communication between artifacts and humans requires (i) the bidirectional use of all modalities available to human beings (optical, visual, linguistic, auditive, gesticular, mimical) and (ii) the ability of the artifact to predict movement, action, and communication sequences. The first point requires, in addition to the recognition and production of expressions in all modalities, also the recognition and pursuit of human dialogue patterns – an extremely difficult problem. The second point is the prerequisite of insightful behavior. On presentation of a task the various options for completing the task are gone through mentally and the best option is implemented in action. Only a few living beings (humans and humanoid apes) are able to do this, but it is an indispensable demand to be placed on artifacts if they are to be taken seriously by humans as interactive partners.

5.2 Questions

In working to create artificial beings with fundamentally new abilities for afferent adaptation to environmental conditions in the broadest sense of the term and also with a new quality of cognitive ability with regard to environmental perception and learning, cooperation and interaction with humans and their own kind, two intertwined problem areas emerge: development of the perceptual and cognitive apparatus and development of a sense of body.

The following fundamental questions, among others, are connected with the problem of controlling the sensorial system and its independent further development:

- What knowledge should be initially ”planted” in the artifact; what knowledge, what behavioral patterns, and what structures will it learn on the basis of sensorial input; first and foremost, however, *in what way* (what is context-dependent, what is dependent on prior knowledge, what instincts exist)?
- What does the artificial being know about itself (form, condition, behavior, desires, intentions, assumptions, abilities) and its environment; what does the designer need to know?
- What strategies does it have for exploring the environment (its own interaction with it) and how can knowledge be constructed through active behavior (looking for food, curiosity, search for social and communicative contacts)?

- What information can be derived from internal knowledge, for what information is (additional) sensorial input required, and how will this be evaluated in the light of background experience and behavioral results; what will be transferred to memory?

- How can representations be built up which will be appropriate to the extendable depiction of internal knowledge of the external world and of one's own competence and in this context include factors such as space, time, and internal state?

- How are ideas and concepts as those developed by humans learned, how do meanings arise, and how are they connected with sensorial and behavioral patterns. Based on the human use of concepts can independent conclusions be drawn for their use or for one's own behavior?

- How can humans gain access to the experience of the artifact, which is accumulated in the same world as human experience, but is doubtless different? What experience will both sides gather in the process of interacting with one another?

- In the end will there be the possibility that operations on these (or higher) representations will create a separate (ego-)awareness?

Further complex problems are associated with the problem of the development of corporeality, the adaptation of behavioral capabilities to suit the environment, and changes made to the environment by the artifact:

- To what extent is physical structure responsible for the sequence of cognitive processes or for their development?

- Is it possible to learn from the evolution of natural examples of sensors and their excellent adaptation to task niches (what factors are adaptable, what can be changed and what not)?

- How are representations created and how do they interact with the robot structure?

- How can these representations be used to support behavioral planning through trial actions, taking into account one's own dynamic behavior, through anticipatory inclusion of sensory patterns (behavior of other systems), and through the use of experience derived from other contexts?

- How can devices and tools used by humans also be used in a purposeful way by artifacts (e.g. through observation of human interaction with them and corresponding abstraction of these observations)?

- To what extent can actors be adapted to new tasks through simple mechanical modifications? How will these modified actors be controlled, how are their sensorimotor systems to be adjusted, how flexible must/can this adjustment be?

- As a result of new (biological) substrates, will the possibility ultimately exist for ”organs” and ”effectors” to develop through growth (in harmony with cognitive abilities) during the lifetime of the being?

Expressed more generally, does the question arise as to what a *generic architecture* may look like for an autonomous system that develops its own cognitive structures as far as possible, builds up a knowledge base (intelligible only to itself) through interaction with its environment (humans) as well as through modification/design of its effectors and, as such, is capable of machine cognition (in contrast to pure learning) and of developing new and non-preprogrammed *behavior patterns*.

5.3 Method

An operational system is indispensable as a basis on which to study these phenomena, i.e. there is a need to develop beyond software at least a rudimentary platform made up of existing mechanical subsystems and/or to develop appropriate components (a simulation will not get us anywhere here). These platforms made up of sensorial and effector mechanisms must be able to register the environment in its complexity (leading to very complex sensorimotor patterns) and they must make possible far-reaching actions or behavioral sequences to influence the environment. Directly connected with this is the challenge of purposeful organization and control of available resources (this will not involve a central control unit but rather several finely or roughly attuned parallel systems) and their programming, a task that it will hardly be possible to confront with the programming techniques used in today’s robotics.

What will be involved, thus, is carrying out *systematic studies* to analyze the principles of action of biological systems at the signal-processing and conceptualization level with a view to their transferability to the cognitive and behavioral activities of artifacts. This regards in particular:

1. The co-evolution of sensory systems, cognitive activities, and effector capabilities during the lifetime of the artificial being and/or an inheritance component (corresponding to chemo-evolution, bio-evolution, and psycho-evolution in living beings).

2. The control and/or use of growth dynamics of the sensorial system and external structures (morphology). Instead of full coding (specification) of the development and/or growth process, only a certain **disposition** is to be given and, as such, principles of self-organization exploited (in a function-optimizing sense, not with a view to bringing forth entirely new forms).

3. The creation of a *cognitive basis* for the learning of structures in different contexts: formation of categories, concept learning, object naming (e.g., through imitation), and transformation of knowledge in a form intelligible to humans.

4. The development of suitable *substrates*/materials and implementation technologies (hardware: analogue, FPGA, hybrid, biological).

5. The structural coupling of artifacts with the environment as well as their coupling with humans and/or other machines through social and communicative interaction and through suitable dynamic ontologies.

For this reason we no longer speak of robots, but rather consider it more appropriate to refer to these artificial beings as ”biomachines”, thus emphasizing the aspects of cognitive complexity, adaptability, and employed principles of action.

5.4 Applications and transfer: market objective

Although the studies we aim to carry out are for the most part very fundamental in character (involving a very high percentage of theory), it can be assumed that there will be a rapid transfer to practical applications. The developmental state of hardware platforms is already quite high (wheeled mobile robots, insectoids, humanoids, special designs), and most of them will be able to profit from the technologies to be made available here.

Nonetheless, *transfer* is to be strongly pursued in the framework of this program. The practical implications of the use of technologies is to be determined early on together with potential users and studied with a view to their marketability. In cooperation with companies who are to carry out the marketing activities (and which can of course be founded in the framework of the program) it will be determined what industrial conditions need to be complied with (safety, reliability, costs, etc.) and in what (software) form the results are to be presented so that they will have commercial value.

After an initial phase in which scenarios are defined (experience shows that this is one of the most difficult phases in a research program of this kind) deliverables are systematically defined for time intervals of 3 to 4 years on the basis of which method-related progress can be documented and which hold out promise of immediate benefit for the further development of existing product lines. They will be built directly on existing platforms and/or components and will follow the principles of modular design with a view to achieving effective reusability. From the current viewpoint there are at least three market segments that would be of relevance here:

1. Improvement of classical applications. Involved here are the areas of industrial robotics (improvement of programming 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 usefulness here is obvious. The primary problem in marketing will be the attractive implementation of laboratory prototypes in components which manufacturers will actually accept.

2. *Complex new tasks in the field of adaptive service robotics.* Here we will seek to provide practical demonstrations of the potentials held out by these technologies by applying them in new areas without an immediate view to marketing. We intend first of all to implement the *airport scenario* we proposed. A navigation-capable artificial porter at a large airport will register a passenger’s desired destination through a multimode interface (e.g. ”to the

PanAm flight to San Francisco”) and then follow the passenger to the designated destination in constant adaptation of its behavior. After that the requirements will be expanded for our *supermarket scenario*: a traveling ”sales guide” in a large supermarket will register what a customer wants to buy and *lead* the customer (i.e. observing the customer at all times, walking speed, desire to stop and look) to the point where the customer can find the product he or she wants and also, if needed, provide further information on the product, special offers, etc. In a second phase it will also be able to take the product from the shelf (e.g. take it from the bottom shelf for elderly customers), something which will necessarily require independent adaptation and constant further development of behavioral capabilities given the differences in product categories and their constantly changing locations. In this scenario almost all the above-named questions regarding the development of cognitive and behavioral systems can be dealt with in a manageable and, most importantly, in a manner that lends itself to practical demonstration and with thoroughly realistic chances of attracting public attention and a large market.

3. Edutainment. Over the long term we see the greatest need (although over the short term not the biggest market) in the field of ”edutainment”. In recent years there has been an increasingly strong trend towards an interlacing of the areas of education, continuing education, applications research, and technology development in the sense of communication of research and development results through ”hands-on” displays or products. This is perhaps most visible in the toy market. Even big-name institutions such as MIT do not shy away from supplying product ideas in this area. **BioMachines** could assume three new tasks in this connection. First of all they could be part of a larger scenario which humans are to be made to understand (e.g. as animated displays in science centers and theme parks). They could be interactive partners for humans (e.g. as toys, in the film industry, or in theater). Finally, they could explain and demonstrate their own developmental principles, i.e. be both subject and object to themselves and, in doing so, attain a degree of interactivity which software on a computer could never achieve (e.g. virtual beings to illustrate teaching material).

Chapter 6

Conscious machines II

Riccardo Manzotti, Vincenzo Tagliasco e Giulio Sandini LIRA-Lab - University of Genova - Italy

6.1 Introduction

In the last ten years several studies have approached the topic of consciousness from a scientific point of view (Tani, 1998; Jennings, 2000; Buttazzo, 2001; O' Regan & Noe, 2001; Ziemke, 2001; Zlatev, 2001; Perruchet & Vinter, 2002; Rees & G. Kreiman, 2002; Taylor, 2002; Crick & Koch, 2003; Gallese & Metzinger, 2003; Harnad, 2003; Zeki, 2003). The Tucson Conferences 'Towards a Science of Consciousness' (1996-2002) have helped to create the scientific environment for the study of consciousness, playing a role similar to that played by the Macy Conferences on Cybernetics (1946-1953), that prepared the ground for cybernetics and artificial intelligence in the following years.

Numerous approaches are emerging to unravel the study of consciousness. Among them there are a neuroscience approach and a constructivist approach. The first was summarized 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). The constructivist approach was sketched by Edelmann and Tononi in their book: 'To understand the mental we may have to invent further ways of looking at brains. We may even have to synthesize 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). The realization that the study of consciousness may require a physical realization is deeper than just being the 'existence proof' of a theory because the subjective experience that we intuitively define as consciousness requires, by definition, the existence of a subject and, unless we refer to spiritual entities, a subject has to have a physical instantiation. 'Pure consciousness' similarly to 'pure intelligence' (in the sense of not being contaminated by physical matter such as flesh or silicon) is just a philosophical paradox. Therefore a scientific approach to the study of consciousness must be based on physical systems and have as a final goal the construction of conscious machines.

6.2 Objective

A conscious machine is an intelligent machine (in the intuitive sense of the word) that is aware of its existence in the world. A conscious machine has emotions, survival drive, motivations, can decide to be selfish or altruistic, has a drive to learn and 'grow'. A new generation of autonomous machine will be the result of our deepening understanding of what gives rise to conscious experience in mammals. For example self-awareness is a crucial tool to filter-out irrelevant sensory information. The decision of what is relevant pertains mainly to the goals and motivations of the perceiving subject and only marginally on the contents of visual information. Furthermore assigning a 'subjective meaning' to a physical event is a very powerful way of reducing the complexity of the sensory and motor information representing such event or of defining the appropriate level of detail required.
Another case in which consciousness seems to be relevant is during learning. It is well known that each activity passes through a series of phases and that, at the beginning learning needs a lot of conscious activity. A dancer needs to train himself/herself consciously in order to move according to certain rules before being able to acquire the necessary and mostly unconscious automatic sensory/motor coordination. A conscious machine must possess the same capability of self teaching.
Finally, the issue of epigenesist and/or development must be addressed. Current machines have to be carefully programmed in order to accomplish their tasks. From what we know brains of mammals possess the capability of self organizing themselves in response to external sensory stimulations. Their structure is far too complex to be completely genetically coded. Furthermore, the behavioral plasticity of mammals is increased immensely by their being able to develop individual unique cognitive capacities. It can be argued that this behavioral plasticity and consciousness emerged together as a result of a common architectural foundation. If this were to be demonstrated, a conscious machine would develop uniquely as a result of its own epigenetic development in a given environment.
Consciousness refers explicitly to the notion of being there or the sense of the self which, in turn, refers to the understanding of phenomenal experience (Aleksander, 2001). Scientific hypotheses about the nature of subjectivity and the physical condition necessary to its emergence can be advanced (Chalmers, 1996; Damasio, 1999; Edelman & Tononi, 2000; Crick & Koch, 2003; Gallese & Metzinger, 2003). This is unavoidable if we want to deal scientifically and practically with consciousness. An attempt to deal with consciousness must challenge these concepts by providing a unitary framework in which these terms can be used. A theory of consciousness has to be formulated as a working background at the beginning of the projects (a tentative theory of meaning and intentionality). A theory for consciousness must provide a series of hypotheses in order to bridge the gap between cognitive systems and physical systems. These hypotheses must propose a candidate structure as the structure responsible for the occurrence of phenomenal consciousness, access consciousness and self consciousness. They must provide necessary and sufficient criteria to design and build conscious machines.

This theory must propose a structural difference that will be used to test whether an artificial system actually hosts the emergence of consciousness. Therefore, the success of such a theory will be evaluated by checking its capability of expressing cognitive and mental jargon in terms of objective structural conditions (like sensory-motor loops, causal relation between experiential events and subsequent system development).

On the other hand, a scientific theory of consciousness must be translatable into a series of experiments and suggestions about how to replicate the physical conditions for the emergence of phenomenal experience.

6.3 Examples

It is difficult to underestimate the implementation of a conscious machine. The following are a number of future implementations:

- 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);
- truly intelligent ’situated artificial communicators’, e.g. for situation-dependent human machine interfaces;
- truly ’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’).

6.4 Current state of technology

Although the filed of artificial consciousness is new, a number of projects have sprung up over the last decades. In 2002-2003, two calls for projects (FET, Future Emergent Technology), ’Beyond Robotics’ and ’Presence’ from the European Union explicitly encouraged projects in this field by addressing the issue of ’machine consciousness’, ’phenomenal experience in machines and robots’, and ’machine awareness’. At the same time well-established research projects have continued to develop in this area. Gerard Edelman’s attempts to build an intentional robot capable of mimicking the cortical neural structure of the human brain is a noteworthy example (Edelman & Tononi, 2000). Igor Aleksander had presented many results in this area (Aleksander, 2000, 2001). At the same time many other laboratories have been trying to establish an common engineering background wherein consciousness can be tackled (McCarthy, 1995; Grossberg, 1999; Manzotti & Tagliasco, 2001; Perruchet & Vinter, 2002; Taylor, 2002; Haikonen, 2003; Harnad, 2003; Manzotti & Glatzeder, 2003). Equally interesting for their potential cross fertilization are the neuro-scientific approaches to the problem of consciousness (Baars, 1988; Atkinson, Thomas, & et al., 2000; Cotterill, 2001; O’ Regan & Noe, 2001; Dehaene, 2002; Rees & G. Kreiman, 2002;

Taylor, 2002; Crick & Koch, 2003; Gallese & Metzinger, 2003; Zeki, 2003). Whether these results will endorse a robust implementation of a conscious being is still hard to say. However, the increase in the number of scientific researchers, explicitly dealing with the design of conscious machines, bears witness to the emergence of a new field of investigation within the scientific community.

6.5 Problem areas

A few problem areas of main interest can be envisaged:

- attention control,
- complexity reduction,
- access consciousness,
- phenomenal consciousness,
- self-awareness,
- self generate motivations.

Attention control and complexity reduction are strictly related since an efficient attention system dramatically decreases the complexity for incoming data by introducing helpful expectations and selection criteria. On the other hand, complexity reduction is needed in order to grasp those meaningful unities which can be used to govern attention. Consciousness can be achieved at different levels. It is customary to distinguish between three different levels of consciousness: access consciousness, phenomenal consciousness and self consciousness (or self awareness). The distinction between the first two is due to Ned Block (Block, 1997). This distinction arises from the thought that the phenomenal properties of consciousness are of a different character to the cognitive, intentional or functional properties of consciousness. For Block, the phenomenal properties of consciousness are experiential properties. On the other side of the coin, we have what Block refers to as access consciousness to encapsulate the tasks involved in cognition, representation and the control of behavior. Finally the last kind of consciousness (self consciousness) is the most complex one involving the capability of self monitoring mental states and recognizing them as part of the history and identity of a unique agent. The implementation of conscious machine can deal separately with each of this kind of consciousness, achieving their separate realization. Finally it is important to notice that a conscious machine will also be able to self generate new motivations, new goals to be achieved. In this way a conscious machine should be able to self teach itself what has to be achieved and also to achieve a much higher degree of behavioral plasticity. Classic neural network learn how to achieve a task which is given to them by their designers, conscious machine will learn by themselves what has to be achieved. This capability should be of great importance for long distance mission (like space exploration) as well as for long term assignment (as robots left on their own for long period of time in unpredictable environment as internet, jungles, the bottom of the sea).

6.6 Future Activities

As it has been mentioned in the previous paragraph the field of conscious machines is at the very beginning of its development; the future activities coincide with the birth of the discipline itself.

6.7 Ethical considerations

Every technical achievement unavoidably entails ethical responsibility and inevitable social pitfalls: hence the implementation of conscious machines carries a heavy load of ethical responsibilities. Two main concerns are evident:

- ethical problems and responsibilities deriving from the interaction between human beings and conscious machines;
- ethical problems and responsibilities towards conscious machines as such.

The former issue is relevant and belongs to the broader problem of the introduction of robots and automatic machines in a social context. Although the existence of autonomous machines is aimed at achieving a better quality of life for workers and citizens, it has often been seen as potentially dangerous since it modifies the structure of human work. Besides, conscious machines are seen as potentially capable of developing their own goals and ends (as has been anticipated by many science-fiction movies) and thus capable, in this respect, of becoming potential competitors for human beings. The issue of responsibility towards a conscious machine as such is still by and large theoretical, since no real 'candidate' is available. However, it is conceivable that such a candidate would possess at least a phenomenal consciousness of the world in which it is living. Regards these machines - similar in their degree of subjective development to mammals like rats, cats, dogs or similar - some kind of ethical rule must be applied. Our understanding of the ethics of animal experimentation has significantly evolved in the last century and by the same token our attitude towards machines will change as soon as we are able to discern in animals the existence of some form of consciousness. A final issue concerns the fact that the construction of conscious machines does not necessarily require the use of living biological materials. It follows that dealing with conscious machines without biological parts many problems could do away with a number of serious bioethical issues.

6.8 Discussion

The theoretical and philosophical issues concerning conscious machines have been extensively debated over the past forty years (Dennett, 1969, 1991; Minsky, 1991; Searle, 1992; Chalmers, 1996). However, the technical implementation of conscious machines is still far from being a certainty; many people still have doubts about its feasibility. Only the attempt to design and build conscious machines jointly with neuroscience findings can add new stimuli to an already mature conceptual framework.

Chapter 7

The ’constructed’ brain

7.1 Introduction

In the field of NeuroIT, as it was defined in the introduction, there are at least three major challenges, which, when taken on, could lead to signifcant advances in technology:

- The interface between the Central Nerve System (CNS) and machine (explained in chapter 2).
- The creation of ’intelligent’ machines: machines that demonstrate flexible behaviour and that are able to adapt to unforeseen circumstance. And also the creation of a brand of machines which can be customized easily to tasks outside the original design specifications, without requiring a re-design of the machine (see chapter 3 and 6).
- New and better ways to study the brain. If one were able to do ’in silico’ simulation of drug tests the pharmaceutical industry would be able to design new drugs in a more systematic way, and, hopefully, in the long run would be able to do without animal (and human) testing.
- An understanding of the brain itself. A better understanding of brain function will beyond a doubt have a profound impact on the treatment of psychological disorders, which are a source of distress to many and also a source of substantial economic damage.

We are stating the obvious when we say that the progress in all these challenges is hampered by the complexity of the brain. Nevertheless, it is useful to look at the problem of complexity a bit closer. The way we propose to address this complexity is the basis for the challenge described in this chapter, and, in our opinion, is essential for progress in the three areas mentioned above. This complexity has at least two different aspects, a methodological and a sociological aspect:
Methodological aspect:

- The brain is complex in a way that is different from the way that a computer programme or machine is complex. Unlike machines, or programmes, the brain seems hard to divide into modules with a well-defined function or structure. Although it is clear that there is some relation between brain area and brain function, it becomes increasingly clear that many aspects of the brain are hard to study in isolation.

- The brain is simply huge, in terms of the number of components and in terms of the physical and chemical processes that are involved in its functioning. Even if the brain were simple from an engineering point of view, its size creates a complexity of its own. This is reflected in the number of publications that is produced at the moment and also in size of data sets of modern experimental techniques.

Sociological aspect:

- There is a large number of disciplines involved in the study of the brain, each with their own methodology, terminology and traditions. Neuroscience and psychology, for instance, have long standing research traditions with relatively little interaction. Although the two are slowly merging into the field 'cognitive neuroscience', there is still a long way to go before a common terminology and methodology will be developed. In some cases this has lead to replications of results in one field, which have been long known in the other. Often it also leads to lengthy discussions, which find their cause in semantic difference, rather than fundamental issues.

- Similarly, the application of results from brain sciences are finding their way into engineering only slowly. Although there are examples to the contrary, 'bio-inspired' applications in engineering or IT often refer to the application of artificial neural networks. As such, they are part of a larger set of statistical learning techniques, where a biological and cognitive perspective is absent. Many introductory texts in the field of machine learning, or computer vision ignore recent advances in cognition completely. A better dissemination of recent findings in cognitive neuroscience and better ways to access them, would undoubtedly be beneficial for the development of truly bio-inspired applications.

The central issue in the difficulties in creating 'intelligent' applications, then, is a lack of insight in why the brain performs cognitive functions so well and so fast. Is it because the brain is massively parallel, on a scale that we still have not realized in hardware or software ? Or do we understand the computational architectures of the brain insufficiently, as argued in chapter 8 and below ? Is the fact that the brain is massively parallel sufficient for an efficient performance of cognitive functions, or is the fact the brain codes information by means of *spiking* neurons somehow a crucial factor ?

We need a comprehensive view of the processes that take place in the brain, during the performance of cognitive tasks and interaction with the outside world, to answer these questions. And we need to answer these questions in order to be able to develop engineering principles for bio-inspired hardware in a systematic way. Using this view, we can decide what cognitive function we can implement in existing hardware, or what kind of hardware we would need to implement given cognitive functions. Using this view we can decide if and how we can extend our sensory capabilities, as described in chapter 2. Moreover, we need an *infrastructure* that allows people from various disciplines to work together, to develop a common methodology and that integrates the vast body of knowledge, which is now scattered over a large number of individual disciplines.

7.2 Objectives

We are reaching the point, where we have sufficient computing power to simulate a complete brain in considerable detail. This may be a human brain, or a smaller one (monkey, cat, insect) depending on what current computer power allows us. This could be called a ’virtual’ brain, in analogy to the ’virtual cell’. It is essential to include possibilities to integrate hardware into the framework (see examples), and the resulting structure might be called the ’incorporated brain’. Or it might be possible to provide an artefact with the brain simulation, thus creating an ’embodied brain’. The total of possibilities we call the ’constructed brain’.

Therefore we propose to:

- Create a framework that allows a large-scale, coarse simulation of the brain, with sufficient flexibility to create more detailed simulations locally, where needed, or to increase overall sophistication when computer power increases.
- Allow for the integration of structural and functional data and simulation algorithms into the framework, thus creating a repository for data and simulation methods.

7.3 Examples

It is not hard to find interesting examples for application of the framework:

- The testing of computational strategies, to see if they can realized within the brain, with the aim of applying them to hardware later.
- ’In silico’ experiments that would be hard or unethical in living creatures. E.g. effects of drugs and hormones on cognitive performance come to mind.
- Virtual lesions, to study the consquences of trauma, or the function of a certain brain region are another possibility.
- Impact of diseases (e.g. Alzheimer’s) and aging on cognitive performance.
- To study where and how artificial implants should be used. To quote a researcher describing the implantation of an artificial retina: ”*The problem is not to transmit an image with high resolution, but to send useful information to the right locations in the CNS.*” (Fernandez, 2003)

7.4 Current state of technology

In this section we will try to summarize the state-of-the-art of the technology which is presently available to initiate this project. We can not give an exhaustive overview of all tools, databases etc., which are avialable at present, but we will provide examples which exemplify the state-of-the-art.

7.4.1 Simulation tools

At the neuronal level, there are several simulation packages, e.g, GENESIS (Bower & Beeman, 1998), NEURON (Hines & Carnevale, 1997), CATACOMB `http://www.compneuro.org/catacomb`, to name but a few. (A good overview of tools that are available at present 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 at the neuronal level. With the exception of CATACOMB, they are very much 'stand-alone' however: few provisions are made to help these package interface to other programmes and databases.

A number of publications describe sophisticate algorithms, which are able to simulate large groups of neurons (e.g., Hansel et al., 1998; Mattia & Giudice, 2000). Unfortunately, many of the programmes that are used to obtain theoretical results, are not made publicly available with some exceptions (e.g., SPIKENET (Delorme et al., 1999)) On a higher level, simulation tools seem to be dominated by neural network simulation packages. For example PDP++ (OReilly & Munakata, 2000) and SNNS (Zell, 1995) come to mind. In general, these packages seem oriented towards artificial intelligence and machine learning, rather than high-level cognitive modelling. An exception is NSL the Neuron Simulation Language which aims to provide a platform for both the simulation of ANNs, as well as neuronal simulations

On the whole, it seems that almost all software that is available was written with a very specific problem domain in mind. For a large number of specific problem domains very sophisticated simulation tools are available. There are some examples of tools which aim for a larger scope, such as NSL, but if they can deal with the various conceptual levels that this project entails, as well as provide the necessary computational performance is still undecided. In general, almost no efforts are made to standardize data formats, coding practices and dissemination of software. Little or no attention is given to the design of interfaces, which allow different simulation tools to be used together.

7.4.2 Databases

The number of databases, as well as the variety of data that they contain, which can be found on the web is astounding. An overview can be found on `http://www.neuroinf.org` (links). There are a few databases, which are defined very professionaly, with good use of modern database techniques, such as COCOMAC (Stephan et al., 2001), which provides extensive information on macaque brain connectivity. A website, created by van Essen and coworkers, contains extensive information on surface based atlasses of human and macaque cortex. An interesting attempt to create a database that is useful for reanalysis of fMRI data is the fMRIDC (van Horn et al., 2001) initiative, announced in the Journal of Cognitive Neuroscience. Here authors are invited submit the datasets, which were used to support their publications. Futhermore there are databases on topics as varied as hippocampus neuron data, ion channels, cortical connections in cat areas. In general, the quality of databases that are publicly available on the web is poor. The number of broken links on web pages that refer to these databases is high and the design of most databases does not conform to the high standards set by e.g. COCOMAC. There is relatively little effort going on to standardize data format and database design. There are, however, some efforts to address this problem: NeuroML (Goddard et al., 2001), is an XML extension which aims to enhance the interoperability of databases, simulations and computational models. NeuroML is a relatively recent development and it still must establish itself firmly. Also it must be extended to incorporate higher cognitive concepts to become useful in a ’virtual brain’, but it and its American counterpart BrainML (`http://brainml.org`) seem to be one of the few initiatives to look beyond one single problem domain.

7.4.3 Theory

There are numerous papers on the behaviour of individual neurons. One of the best known, of course is the seminal paper by Hodgkin and Huxley (1952), but by now there are literally thousands of papers, addresing aspects of morphology, ion channels, receptors, cable theory and so on. This is an important line of research, which has a large number of participants and which is very active.

On a somewhat larger scale, important topics in NS modelling and theory are: how does the cortical code work (e.g. various authors, 2001) (rate coding, precise inter spike times (e.g., Maass, 1997). How are Long Term Potentiation and Depression (LTP and LTD) realized and what role do they play in learning, etc. How do cortical and subcortical structures apply these mechanisms to produce behaviour ?

Somehow one must find ways to incorporate this information in higher-level descriptions of the brain. It is impossible, now and in the foreseeable future to simulate the billions of brain cells, and their interconnections, even at a very superficial level. And even if this were possible, it would be of limited use. One would still have to extract higher-level cognitive information from the spike trains of these billions of neurons. Similar problems in physics have lead to the rise of statistical physics, where, for instance, a realistic description of the behaviour of gasses sometimes requires the inclusion of quantum effects. The challenge is to apply a microscopic theory which is developped to describe individual particles (or very small systems of particles) to the vast number of molecules present in macroscopic volumes. The use of statistical techniques was essential to achieve this.

Techniques of statistical physics are used increasingly often to model the behaviour of large groups of neurons. The behaviour of large groups of neurons under input is very complex, but in the last decade significant progress has been made on this subject. It was 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), even with the inclusion of some neuronal details (Casti et al., 2002), can be described by powerful sets of equations. Although the solution of such equations is not trivial, it is computationally much more efficient than a straightforward simulation of a large group of neurons. These techniques have also been applied to cortical circuits, which are belived to underly working memory (e.g., Amit & Brunel, 1997), attention, the formation of orientation columns in visual cortex (Nykamp & Tranchina, 2000), etc.

This is an extremely interesting development, because, first, it is possibly an important step to the description of the large-scale cortical networks. It is interesting to note that recent developments in fMRI allow the identification of large-scale functional cortical networks, using techniques like Structured Equation Modelling , or more recently, Dynamic Causal Modelling. If good descriptions of neural activity for higher-level cognitive processes can be found, it may be possible to simulate fMRI and EEG signals. Thus, it would be possible to confront models in a very direct way with experimental data. Secondly, a large number of higher-level cognitive models is still implemented in terms of 'connectionist' models. Explanations for coordinate transformations between various frames of reference (head-centered, eye-centered etc.) (e.g., Zipser & Andersen, 1988; Pouget & Snyder, 2000) for instance use perceptron networks trained with 'back-propagation', as are models for attention (e.g.; van der Velde & de Kamps, 2001), models for long-term memory formation in the hippocampus complex (e.g., Rolls & Treves, 1998). It is extremely interesting to see how 'connectionist' concepts can be rooted in neuroscience, because this is likely to provide constraints for using 'connectionist' techniques in a biological (and also cognitive) context (see e.g. Gerstner, 1995; Maass, 1997; de Kamps & van der Velde, 2001).

A final important role for theory is the determination of the computational architecture of the brain. The human cortex is remarkably uniform, and this gives rise to the idea that there is a relatively small number of cortical configurations which underly the computational performance of the human brain. The idea is that, despite the fact that the human cortex performs an astounding variety of complex computational tasks, from language processing to visual object recognition, a relatively small number of computational principles are applied everywhere in the human cortex. The notion of 'cortical circuits', for instance, has been around for a while (Douglas, Martin, & Whitteridge, 1989), and theoretical and experimental evidence is mounting that there are a number of basic cortical circuits which underly cognitive performance. However, also on a larger scale there appear to be computational architectures in the cortex. It has been argued, for instance, that the visual cortex is a so-called 'blackboard architecture' (van der Velde, 1997; Bullier, 2001): different high-level apects of visual stimuli, such as color, form, motion, etc., are processed by high-level visual area. Feedback information from higher to lower visual areas could lead to a re-evaluation of information in lower visual areas and, for instance, solve binding questions. The evidence for a 'blackboard architecture' in the visual cortex is quite strong and the notion that similar principles could also be involved in the processing and production of language has been voiced by various researchers.

The investigation of 'computational architectures' like these is important, becasue it relates to possible hardware implementations: if there is a small number of 'computational architectures' in the cortex, and if they are understood, we problably understand how a massively parallel structure of relatively slow elements is able to perform complex forms of computation. We would be able to evaluate if we could emulate such a structure in existing or projected hardware, and if this would lead to an application with the desired performance.

7.4.4 Towards large-scale cortical structures

A few years ago, techniques like PET and fMRI were mainly used to compare activity in brain areas between two conditions. A cognitive task was performed by the subject, during which (correlates of) brain activity were measured. Brain activity was also measured during the 'base line' condition, where the subject was disengaged from the cognitive task as much as possible.

Using eloborate statistical analysis a difference between activity in brain areas between the two conditions was established, which lead to conclusions of the kind 'area X play a role in cognitive task Y'. Although important information can be obtained from such analyses, there are obvious limitations. First of all, it is hard to create 'base line' conditions, which truly distinguish the state of mind of the subject performing a task from a 'rest' state ('try to think of nothing'). Secondly, it is well-known that some cortical functions (e.g. long-term memory) are distributed over a large part of the cortex and a localization of function is probably not very meaningful. Thirdly, even if two brain areas display about the same level of activity between the two conditions (which means that they are 'subtracted out' in the comparison between the two conditions), this does not mean that they do not perform a function in the cognitive task under investigation. Although the overall level of activity may be the same in the two conditions, the activity may be of a different kind, between the two conditions, as for instance SEM modelling (see below) has shown. Finally, to see anything at all, sometimes analyses must include several test persons and such analyses are notoriously difficult, due to person-to-person differences in individual brains. These disadvantages have made some people sceptical about the use of fMRI and PET research.

There are, however, several recent developments which are extremely interesting from the point of view of the large-scale structure of the brain. First of all higher magnetic fields in fMRI scanners and a better understanding of the BOLD effect (which is a correlate of neural activity) have lead to so-called event related fMRI: responses in individual subjects, caused by single changes in conditions in the cognitive task that is studied can be observed. Secondly, more sophisticated analysis methods are used which allow a more subtle use of data than a simple comparison between two conditions. One of those technique is Structural Equation Modelling (SEM) (see (van Hulle, 2003) and references therein), which does not simply look at the level of activity of a brain area, but which also takes correlations between brain areas in a single condition in account. In this way, effective connectivity between brain areas can be established, and correlation between brain areas can be studied. This has lead to the identification of large-scale functional networks for cognitive tasks, which can supplement structural networks, which can be identified in anatomical studies, which are e.g. represented in the work of van Essen of the COCOMAC database. The implications of the discovery of these networks may be profound. Already it has been demonstrated that there is difference in these networks between young subjects and old ones, which is not reflected in behavioral measures. These discoveries by themselves are important enough. In section 7.6, where we discuss future developments, we would like to point out the great possibilities that recent developments in this field, combined with recent developments in theory (see section 7.4.3) would offer: for the first time 'top-down' modelling is possible, where the existence of large-scale functional andeffective cortical networks is given and where models on a smaller level (cortical circuits, small neural networks) are used to reproduce the large-scale networks.

Although the techniques discussed here are relatively new and new techniques like Dynamic Causal Modelling (DCM) (van Hulle, 2003) are still under investigation, already new approaches to analyze large-scale cortical structures are being considered, such as Bayesian network approaches (e.g. Neal, 2000). Such approaches, in combination with a good model for the underlying neural activity, may prove an important step forwards in constructing an overall model of the brain.

7.4.5 Hardware

There are two possible implementations for this kind of project. A simulation on a large-scale super computer, including highly networked high perfromance clusters. Or specily designed chips using Very Large Scale Integration (VLSI) techniques. To our knowledge VLSI techniques have been used for relatively small systems (see (IJspeert & Mange, 2003) for an overview). There are, however, initiatives to develop VLSI architectures, which have a size and level of connectivity comparable to the brain (Hammerstrom, 1999).

Large supercomputers are available in national centra throughout Europe. The computational power of these computer is impressive and still growing. (Some numbers on current supercomputers to be inserted here.).

7.5 Problem areas

Many of the problem areas already have been discussed above. We will review them here.

- A true, multidisciplinary overview of 'brain science' is lacking. 'Brain science' is still composed of several traditional disciplines, which have relatively little knowledge of each other's work, methodology and terminology.
- Information which is essential for good modelling and theory making is scattered over thousands of sources, many of them still paper.
- There is little structure in databases which are publicly available. The quality of databses ranges from unstructured lists of facts to well-designed, professional datastructres. There is also little consistency: some databases vanish from public view after a while.
- Little use if made of existing techniques to share information between databases or to construct data models in such a way, that databases can readily interface to each other.
- Similar problems affect simulation tools.
- Simulation tools are not used enough. Substantial progress in theory, for instance, is not available because it is not cast into software which is publicly available. There is a large of amount of data, which is well-described and well-established, but which is relatively hard to obtain. A good comprehensive model that represents everything which is known about a 'cortical column' or 'cortical circuits' which is publicly available, in a well-defined format would be of considerable value.
- Large-scale cortical structures are in the process of being discovered. There is no methodology to capture these large-scale strutures, and to structure the data from which they are inferred. Such a methodology is essential for a future large-scale model of the brain.
- The complexity of issues concerning software is underestimated. Too often programming is viewed as something that every scientist can do on the side. Software design and implementation should be regarded as important, even key issues.

- New techniques are necessary to visualize the complex data sets, delivered by current experimental methods. These techniques will probably essential in visualizing a large-scale model of the brain itself.

Some of the problems above are being addressed by current initiatives. In particular the transmigration of important information from paper to electronic databases is clearly in progress, due to initiatives like the Human Brain Project. Also, some disciplines slowly start to merge (one might consider 'cognitive neuroscience' as the marriage of cognitive psychology and neuroscience). Overall, howver, the research community still seems very fragmented. Many of the tools that are being created at the moment are very useful for other researchers in the field in which they are created, but are relatively difficult to use for researchers from another field. A modeller, for instance, will certainly appreciate a good brain atlas. If he wants to use the underlying information, however, it will not do to make this available over a Graphical User Interface (GUI) based web application. The modeller may want to use the underlying information, that is used to generate these atlasses, and may want to access this information in hiw own programmes through direct interaction with a local database. Although, as mentioned above, there are some initiatives that try to improve the interoperability of various databases and simulation tools, it remains to be seen if they will find wide acceptance. The success of these initiatives will be determined by the number of supportes and users they will gain. For many research groups there is no strong incentive to seek multi-disciplinary collaboration, because substantial scientific progress in their own field is still possible, and as long as this is the prevailing situation, there will be no strong drive to support initiatives for interoperability.

7.6 Future Activities

7.6.1 A ground plan for the brain

A first step would be to make an inventory of hardware which is suitable for a first implementation of a project of this kind. A supercomputer would be the first logical choice. A second step would be to make an inventory of existing techniques and projects which are going in the same direction as this initiative and to invite them into an organisational structure.

An essential concrete step would be to implement a coarse large-scale brain structure on the chosen hardware right from the beginning. To initiate such a step, probably a conference or workshop is necessary, to provide the minimum of coordination, which is required for starting such a project.

7.6.2 A 'start up' programme

But clearly, such a step takes time and it is important to start-off immediately. So let us define a **start up** programme.It is possible to start with small-scale projects. These projects must model relatively limited aspects of human cognition, or they must emulate parts of human cognition or motor tasks in hardware (artificial retina, cochlear implants, robot arms/hands, etc.). **To qualify for the 'constructed brain' project, the most important deliverables should be**

software libraries. The quality of these libraries should be verified by independent experts and these libraries should conform to high standards with respect to quality of code, quality of interfaces, documentation, and maintenance. Moreover, they should work on various, pre-assigned platforms (which could include high-performace computers).

The requirement that models be published, but be cast in a well-designed software library as well, is a significant step forward towards models that are re-usable, either in other (more complicated !) models, or in hardware implementations. The requirement that hardware have a well-defined software interface opens interesting possibilities: expensive hardware could be kept locally in a labarotory set-up, with remote access available via the software interface.

Particular stimulation could be considered for projects which promise synergy: this could be 'model-model',or 'model-hardware' combinations that are i) related and ii) use or design a common sofware interface. Examples: a retinal implant, interfaced with model of early visual processing, an artificial hand, whose sensors are interfaced to a model of sensory-motor cortex, a model of visual processing interfaced with a model of auditory processing interfaced with a model for multi-modal representations, etc.

Again, it is emphasized that the most important deliverables of these project should be software libraries, which would distinguish this 'start up' project from other funding projects in this area of research. These software libaries themselves can be an important object of study in a later stage: how can they be maintained, how can they be set up, in such a way that they can easily be extended, and cooperate with other projects which have not been conceived yet. This is a central issue in software engineering in general, but in the context of the 'constructed brain' an extra quality enters: as we learn more on how the brain works, some of this insight might make its way back into our techniques for constructing software which is reliable, fault-tolerant and 'degrades gracefully'.

In parallel the following activities must be undertaken, in order to refine the initial structure in an iterative way.

- The collection of databases which allow direct access to: neuronal data, small scale brain structures (e.g, details on the structure of a cortical column), structural and functional connectivity databases, and databases for large-scale cortical structures. We say collection, because these databases already exist or are being created. The specific goal of the collection of databases is to provide modellers with the possibility of accessing data at all levels, directly in one software development. This entails the creation of interoperability techniques, methods to load parts of large databases into core in a flexible way.

- The creation of of an external environment for the brain to interact with. The environment in first instance would most likely consiste of simulated sensory input and output, which would corresponds to very simple abstractions of 'real-world' simulations. In later stages these abstractions can be extended, and also the environment could be extended to include 'real' sensory input and motor output. In a later stage, sequences of sensory input and the consequences of motor actions on sensory input can be considered ('closing the sensori-motor loop').

- The development of theoretical methods which bridge the gap between two orders of magnitude or complexity. Statistical mechnical methods which describe the collective behaviour of large groups of neurons are one example. The description of cortical circuits with dynamical models, which capture the essence of more detailed and realistic neuronal models, for instance the phase space portrait, but which require less computing power to evaluate. A succesful programme which describes cortical circuits could probably be the input for a similar programme for higher-level cortical structures.

- Investigation of 'computational architectures', both from theoretical and from experimental point of view is important. If there exists a 'cortical principle' of parallel computation, of at least a few relatively simple ones, as we have argued in section 7.4.3, this could well lead to an understanding of how to implement 'cortical principles' in hardware, and thus offer a realistic estimate as to what cortical function can be succesfully emulated into which type of hardware.

- The creation of visualization tools which allow an overview of the project, at every possible level.

7.7 Discussion

In the first steps of the project, that we propose in the section above, there are many activities which are now done by various grous around the world. We hope, that by creating a single project which can serve as a framework for these activities, and also by creating a central hardware platform, a 'condensation point' for these activities will emerge. This will provide a strong incentive for various groups to work together in a natural way.

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.

Chapter 8

The Brainprobe project-Tools for Neuroscience

8.1 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 8.1

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

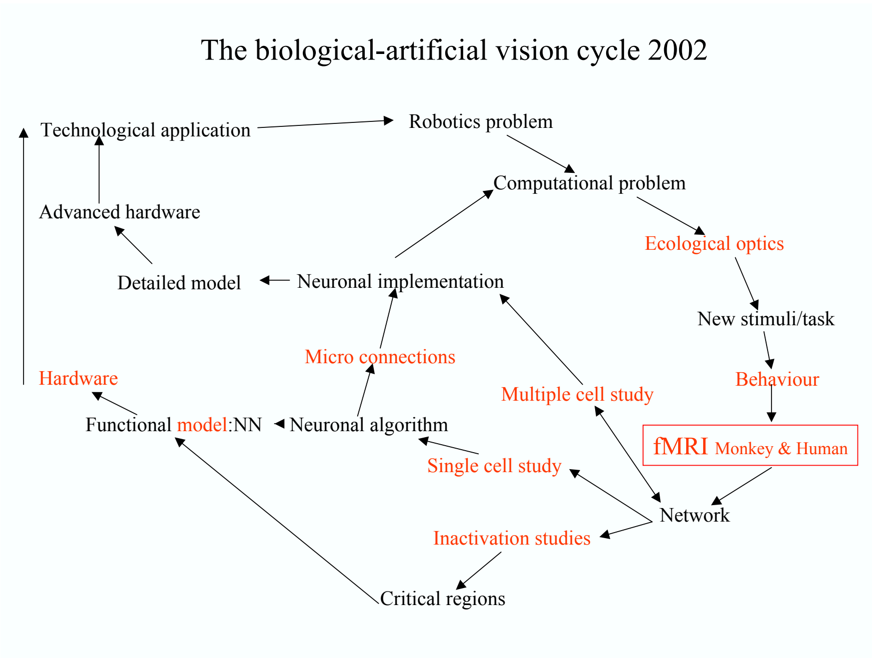

Figure 8.1: The different levels of possible interactions between neuroscience and robotics.

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.

8.2 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

8.3 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

8.4 Current state of technology

8.4.1 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.

8.4.2 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).

8.4.3 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.

8.4.4 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.

8.4.5 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 rostroventral 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.

8.4.6 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).

8.5 Problem areas

8.5.1 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.

8.5.2 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.

8.5.3 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.

8.5.4 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.

8.5.5 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.'

8.5.6 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.

8.6 Future research

8.6.1 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.

8.6.2 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

8.6.3 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 sub-part (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.

8.6.4 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 preextant 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.

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

See other chapter.

8.6.6 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)).

8.6.7 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.

8.7 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.

8.8 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.

8.9 Acknowledgements

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

References

Aleksander, I. (2000). *How to build a mind.* London: Weidenfeld & Nicolson.

Aleksander, I. (2001). The self 'out there'. *Nature*, *413*, 23-23.

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.

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

Atkinson, A. P., Thomas, M. S. C., & et al. (2000). Consciousness: mapping the theoretical landscape. *Trends in Cognitive Sciences*, *4(10)*, 372-382.

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. J. (1988). *A cognitive theory of consciousness.* Cambridge: Cambridge University Press.

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. (1997). *On a confusion about a function of consciousness. the nature of consciousness. o. flanagan and g. guzeldere.* Cambridge (Mass): MIT Press.

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.

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.

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.

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.

Cotterill, R. M. J. (2001). Cooperation of the basal ganglia, cerebellum, sensory cerebrum and hippocampus: possible implications for cognition, consciousness, intelligence and creativity. *Progress in Neurobiology*, *64*, 1-33.

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.

Craelius, W. (2002). The bionic man: restoring mobility. *Science*, *295*, 1018-1021.

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. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness.* New York: Harcourt Brace.

Dehaene, S. (2002). *The cognitive neuroscience of consciousness.* Cambridge (Mass): MIT Press.

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. C. (1969). *Content and consciousness.* London: Routledge & Kegan Paul.

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.

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.

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.

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. 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.

Fernandez, M. E. (2003). Designing a brain-machine interface for direct communication with visual neurons. *Presentation on NeuroIT.net workshop on July 8th, 2003, Alicante*.

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.

Fetz, E. E. (1999). Real-time control of a robotic arm by neuronal ensembles. *Nature Neuroscience*, *2*, 583-584.

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.

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.

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

Georgopoulos A. P., K. R. E., Schwartz A. B. (1986). Neuronal population coding of movement direction. *Science*, *233*, 1416-1419.

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.

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.

Hammerstrom, D. (1999). *The OGI white paper.* http://www.ece.ogi.edu/ strom/research/papers_talks/wp.pdf.

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.

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.

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.

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.

van Hulle, M. (2003). Contribution to roadmap. *www.neuro-it.net*.

IJspeert, A., & Mange, D. (2003). Contribution to roadmap. *www.neuro-it.net*.

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

Johansson, R. S. (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.* Academic 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.

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.

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.

Kvnig, P., & F., V. P. (2002). Neurons in action. *Science*, *296*, 1829-1832.

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.

Laubach, M., Wessberg, J., & Nicolelis, M. A. (2000). Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature*, *405*, 567-571.

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.

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).

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.

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.

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.

Minsky, M. (1991). *Conscious machines. machinery of consciousness.* National Research Council of Canada.

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

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

Neal, R. M. (2000). On deducing conditional independence from d-separation in causal graphs with feedback. *Journal of Artificial Intelligence*, *12*, 87-91.

Nicolelis, M. A. (2001). Actions from thoughts. *Nature*, *409*, 403-407.

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.

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.

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.

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.

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). http: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.

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

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.

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.

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.

Searle, J. R. (1992). *The rediscovery of the mind.* Cambridge (Mass): MIT Press.

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.

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.

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

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.

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

Todorov, E. (2000). Direct cortical control of muscle activation in voluntary arm movements: a model. *Nature Neuroscience*, *3*, 391-398.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*, 1226-1235.

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.

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.

Wickelgren, I. (2003). Power to the paralyzed. *Science, 299*, 497-497.

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.

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.

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.