# THE ARTIFICIAL BRAIN

**A SEMINAR REPORT**

***Submitted By***

**RON A. MATHEW**

***In partial fulfillment for the award of the Degree***

***of***

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**SCHOOL OF ENGINEERING**

**COCHIN UNIVERSITY OF SCIENCE & TECHNOLOGY,**

**KOCHI-682022**

**AUGUST-2008**

**DIVISION OF COMPUTER SCIENCE & ENGINEERING**

**SCHOOL OF ENGINEERING**

**COCHIN UNIVERSITY OF SCIENCE &**

**TECHNOLOGY,KOCHI-682022**

# Certificate

Certified that this seminar report ***The Artificial Brain*** is the bonafide work of ***Ron A. Mathew***, who has carried out the seminar under my supervision.

SEMINAR GUIDE
**Mr. Pramod Pavithran**
*Lecturer*

HEAD OF THE DEPARTMENT
**Mr. David Peter**
*Lecturer*

# ACKNOWLEDGEMENT

First and foremost I would like to acknowledge the fact that this seminar topic is the result of the inspiration gotten from the introductory class taken by Mr.Pramod Pavithran, who is also my Seminar Guide for one of my subjects. His class took a brief journey through the infinitesimal possibilities of Artificial Neural Networks and artificially created systems that could think for themselves. That very class provided me with the motivation to work on this particular topic.

Apart from profoundly thanking God for his graciousness in providing me with much help through this project, I would also like to extend my heartfelt gratitude to my Head of Department, Mr. David Peter, who provided a lot of valuable support, mostly being behind the veils of college bureaucracy. Also, I would like to lend my gratitude towards Mrs. Sheena Mathew, our Seminar Instructor who was responsible in guiding us through the task of giving good seminars.

Lastly, but not at all the least, I would like to thank my Seminar Guide, Mr. Pramod Pavithran who proved to be more than just a seminar guide and being more of a friend than just a teacher at college.

The role played by my friends in the successful completion of this seminar cannot be underestimated at all, with their advises proving to be inevitable words of help.

Ron A. Mathew

## ABSTRACT

We have always been interested in the notion of consciousness fact, which is, for us, the fact that an individual endowed with a brain can think of something related to his position in the world right here right now. It is not about the continuity, or the performance, nor the profoundness of the thought, but it is about thinking of something in a knowable manner and which can be specified from a linguistic or mathematical angle, without it being an automatic and predefined response to a given situation.

By analogy to the notion lengthily investigated by philosophers, psychologists, neurobiologists, we will pose the question of artificial consciousness: how can one transpose the fact of "thinking of something" into the computable field, so that an artificial system, founded on computer processes, would be able to generate consciousness facts, in a viewable manner. The system will have intentions, emotions and ideas about things and events related to it-self. The system would have to have a body that it could direct and which would constrain the system. It would also have to have a history, and intentions to act and, most of all, to think. It would have to have knowledge, notably language knowledge. It would have to have emotions, intentions and finally a certain consciousness about itself.

We can name this system, by sheer semantic analogy, an artificial brain. However we will see that its architecture is quite different from living brains. The concern is transposing the effects, the movements; certainly not reproducing the components like neurons and glial cells.

We should keep in mind principally one characteristic of the process of thinking unfolding in a brain: there is a complex neural, biochemical, electrical activation movement happening. This movement is coupled to a similar but of a different mode in the nervous system deployed in the whole body. This complex movement generates, by selective emergence and by reaching a particular configuration, what we call a thought about something. This thought rapidly leads to actuators or language activity and descends then in the following thought which can be similar or different. This is the very complex phenomenon that has to be transposed into the computable domain.

Hence, we should approach the sudden appearance of thoughts in brains at the level of the complex dynamics of a system building and reconfiguring recurrent and temporized flow. We can transpose this into computer processes architectures containing symbolic meaning and we should make it geometrically self-controlled. Two reasonable hypotheses are made for this transposition:

- analogy between the geometrical dynamics of the real brain and of the artificial brain. For one, flows are complex images, almost continuous; for the other, these are dynamical graphs which deformations are evaluated topologically.
- combinatory complexity reduction of the real brain in the computable domain by using symbolic and pre-language level for this approach. The basic elements are completely different; they are not of the same scale.

However, once these hypotheses made, one should not start to develop an architecture that will operate its own control from the aspects of its changing geometry. One needs to ask the proper question about consciousness fact generation. A philosopher, a couple of decades ago, M. Heidegger, asked the proper question: *what brings us to think about this thing right here right now?* The answer, quite elaborate, to this question will conduct to a system architecture choice that will take us away from reactive or deductive systems. The system will generate intentionally its consciousness facts, intention as P. Ricoeur understood it.

There are no consciousness facts without intention to think. This settles the question, considered as a formidable, of freedom to think. One thinks of everything according to his memory and his intuition on the moment, but only if it is expressible as a thought by the system producing thoughts. Some might see something infinite in this process; however it is not our case. A finite set of component which movements occur in a finite space has only a finite number of states in which it can be. Also, as the permanence of the physical real apprehensible by the sense is very strong, the preoccupation to think by man is quite limited, in his civilizations. Let us point out that artificial systems that will think artificially will be able to communicate directly at the level of forms of the ideas, without using a language mediator, and hence, would be co-active as well as being numerous in space.

For different reasons, numerous people think that the path of artificial consciousness' investigation should not be taken at all. I feel differently, because, discoveries have been the very root of our existence, from fire to the mighty F-16s. The mind is a work of art moulded in mystery, and any effort to unlock its doors should be encouraged because, I am sure, that its discovery is only going to help us respect the great architect more.

## TABLE OF CONTENTS

# PROLOGUE

As human beings, what makes us special is our brain. We don't have legs that are fast enough or arms strong enough to survive in this world. But, we have a brain, a gross looking, collection of tissues with no mind blowing gadgetry of any sort, the world's most understated advanced piece of technology that makes us what we are.

This seminar is only an insight, a small preview to the larger answers to be found and I also do not intend to probe into the inner details because it could prove to be an extensive task.

The Artificial Brain, as fictional as it may sound, proved to be an eye opener for me. Opinions that research into the creation of a brain-alike is discouraged citing religious reasons, but it should be understood that, any understanding so far received in this field of research has only proved to be a revelation of the mighty power of God and I am sure it will continue to be so.

Ron A. Mathew

# I. The Brain and Consciousness:

# A Biological Approach

# Chapter 1

## The Biology behind our Consciousness and Emotions

### 1.1 The Brain

The *cerebral cortex*, the convoluted "grey matter" that makes up 80% of the human brain, is responsible for our ability to remember, think, reflect, empathize, communicate, adapt to new situations and plan for the future. The cortex first appeared in mammals, and it has a fundamentally simple repetitive structure that is the same across all mammalian species.

The brain is populated with billions of *neurons*, each connected to thousands of its neighbors by dendrites and axons, a kind of biological "wiring". The brain processes information by sending electrical signals from neuron to neuron along these wires. In the cortex, neurons are organized into basic functional units, cylindrical volumes 0.5 mm wide by 2 mm high, each containing about 10,000 neurons that are connected in an intricate but consistent way. These units operate much like microcircuits in a computer. This microcircuit, known as the *neocortical column* (NCC), is repeated millions of times across the cortex. The difference between the brain of a mouse and the brain of a human is basically just volume - humans have many more neocortical columns and thus neurons than mice.

### 1.2 The generation of emotions in neurobiology

Nowadays, the generation of emotions is a well-studied problem in neurobiology. It is not merely about psychological knowledge, where we observe some behavioral effects, but we have the knowledge of the internal functioning of the brain in relation with the nervous system driving to emotional states. We can lean on the available results in neurobiology and transpose them to construct a similar system, having the same principles of architecture and functionality, but using multiagent paradigm with the notion of self-reconfirmation.

In biology, an emotion is a vital function of the central nervous system that triggers to typical states. It is a psychic and physiological behavioural state produced by a neural activity from an inductive, driving to some bodily behavior. Biologists found the "center of the emotions" and can precisely describe the emotional process of the production of pleasure. The center of pleasure in the brain has the following architecture:

- The first part of the central nervous system generating emotional behaviours is the *hypothalamus*. Nervous cells of the hypothalamus control the hormonal secretions of the hypophysis and so control the hunger, the thirst, the circulating electrolyte rate…
- The second component intervening in the emotional process will be limbic system. It is composed of the tegmental ventral area, situated to the basis of the brain, and of the core accumbens, situated deeply under the frontal cortex of amygdalins lobe. Limbic system is connected to the hypothalamus by its median part. It also communicates with the neo-cortex by its lateral part. The center of the emotion is going to communicate therefore with the whole of the brain.
- The other components of the system are the epencephala and the pituitary gland that generate tendencies respectively toward a goal by the production of dopamine, and that achieve the pleasure by the production of morphine molecules. Considering more particularly the pleasure, the called "rewards – punishments" system was localized and well clarified by neurobiologists. The steps permitting to have the generation of a pleasure emotion, are:
    i. At the beginning the system is in a state of neutral normal functioning,
    ii. A signal, in the biologic meaning of this term, trigger the process of generation of a specific type of biochemical components,
    iii. In Epencephala system, at the reception of the components freed by the previous process, an incentive expressed by flux of dopamine hires the system to enter in activation toward an expressed goal,
    iv. A center of pleasure - displeasure, the pituitary gland, values the success or the failure of the reach of the goal stated by the incentive, generating morphine molecules, and hires the pursuit of the process or makes stop. In any case, the action is a real process that tampers limbic system by memory effects.

# II. The Artificial Consciousness and Self-Generating Emotions

# Chapter 3

## The Artificial Consciousness

### 3.1 System generating consciousness facts

A system capable enough to generate facts related to consciousness is going to be very complicated to recreate. But if we try to classify the different components or the levels that guide the brain in coming to decisions, we can come up with the following components.

This system can have five components:

- ❖ An organizational memory that is a large memory of facts, knowledge, rules and events where anything coming from this memory is systematically adapted to the current context. Such a memory in not knowledge base but is a continuous dynamic interpretation of knowledge.
- ❖ A subsystem building in a strictly constructivist way the current artificial idea here and now: that is the construction of the current idea like the well-controlled activity of some large agent organization.
- ❖ A subsystem generating emotions as the alteration of the activity of the subsystem expressing the current idea, that is some specific field altering the focus of this generation, according to the specificity of the emotion.
- ❖ An input–output subsystem linking the system producing artificial consciousness facts with the body of a robot or any software data flow.
- ❖ An interface subsystem expressing the mental map that is the representation of the current artificial consciousness fact with the reasons of this generation.

These five levels to the thought process can help simplify the task involved in developing a system with an artificial consciousness.

### 3.2 A Pictorial Representation of the System

The actual system will control the behavior of an autonomous robot.


Fig.C.3.1

#### 3.2.1 General architecture of the system

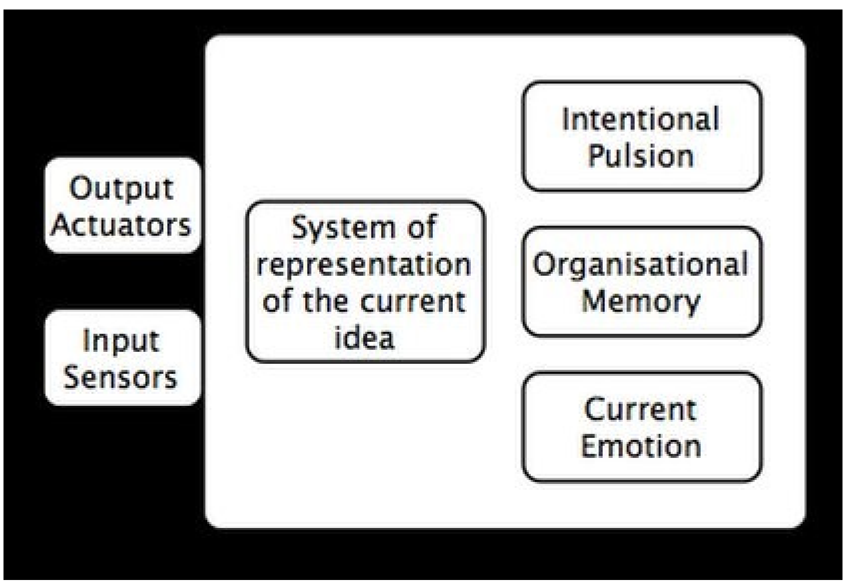



Fig.C.3.2

The system builds inline a representation of the state of the robot from the data of the sensors (its environment, its internal state according to the its organizational memory and its current emotion...). From this representation, the system drives the behavior of the robot via the actuators.

### 3.2.2 The Pathway of Data Transfer and monitoring

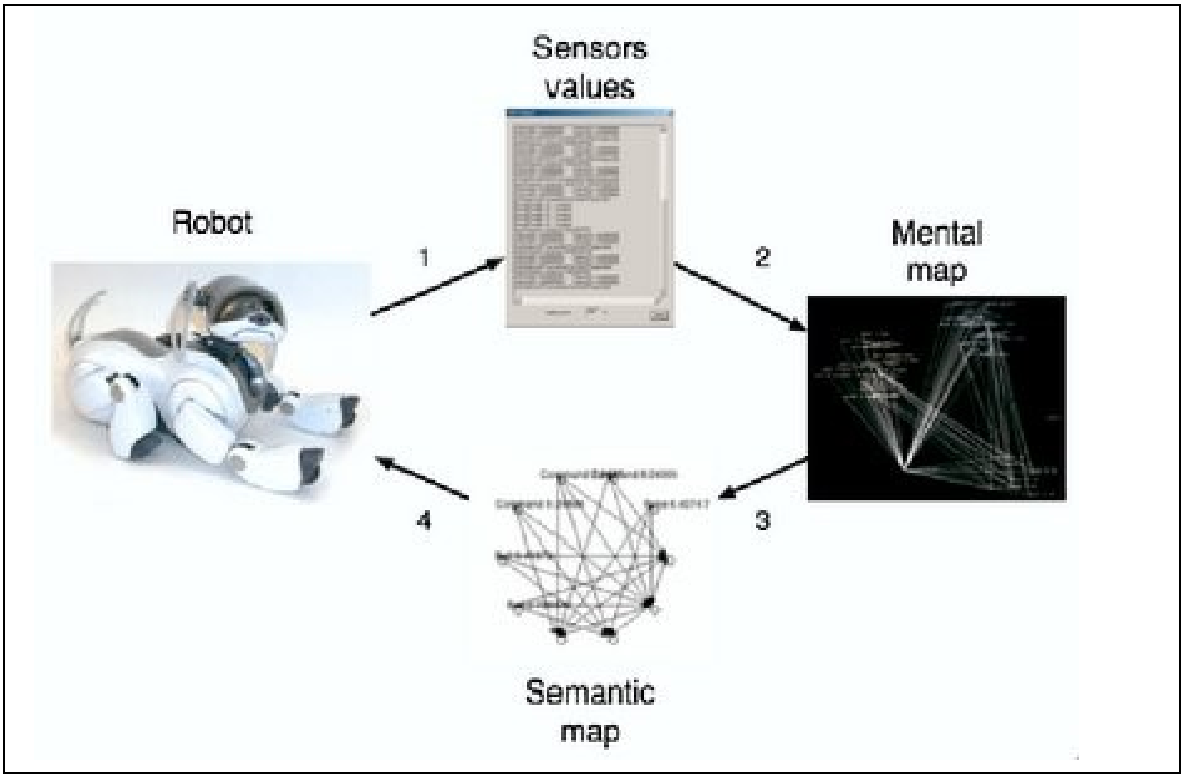



Fig.C.3.3

The image above describes the system loop:

i. Input of the system (sensor data: video, sound...)
ii. Mental card building, agent co-activity
iii. Analysis, semantic interpretation of the mental card
iv. Decision making, orders sent to robot.

This is achieved under the multi-agent paradigm.

The representation is not a static picture, it is like a movie which is generated by an number of active elements, which we can also call agents: this is the transposition of a mental map:

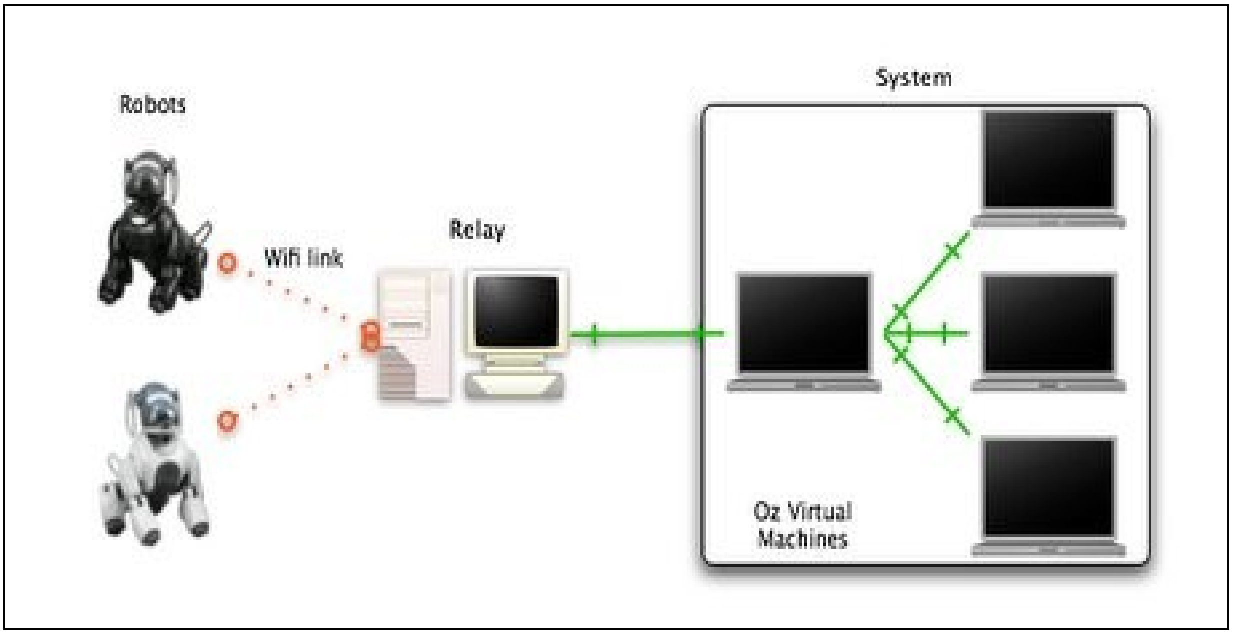



Fig.C.3.4

The system prototype is developed in Oz/Mozart. The system is distributed on several Oz/Mozart virtual machines.

# Chapter 4

## Self-Generating Emotions

### 4.1 General architecture of a self-adaptive system generating emotions

The data processing always drives to the construction of a system running in a computer, but the behaviors of systems are varied. One can classified the systems, according to their behavior, in three categories.

#### 4.1.1 The reactive systems and the automatic action

A reactive system is a system for which each event of the environment is feared like a stimulus that instantaneously puts in action in a way that strictly corresponds to the stimulus characters. There is a causal and sufficient link between outside events and the reactions of the system that are elements of the same nature. Link stimulus – action is a procedural call, a kind of reflex method. The system is constructed and pro grammed to answer some restricted class of events by automatic activations of procedures.

The central problem for such systems is the control and the efficiency of the action event - corresponding subroutine. This type of system doesn't attach any meaning to its actions and only an anthropomorphic observer's can attach some emotion to its behavior. These systems are, with regard to their construction, very decomposable in clearly identified parts with, for each part, a very precise functional role.

#### 4.1.2 The perceptive systems and the selective action

In this case, each event coming from the environment is considered as a complex fact. There is symbolization of the event feared, representation in an internal entity composed of many data and using knowledge basis. This entity, that symbolizes some aspects of the things of the environment, is clearly defined and its structure can be complicated. The problem is then to well recognize the whole characters of the event for the precise identification and have a reaction in an appropriate manner. Some recognition characters permit to identify the event and to drag the system reaction suitable and precise. The system distinguishes elements of the environment while symbolizing them by many predefined characters specified by ontology. It is constructed with a mediating module between environment and subsystem of action, but it remains a solver of problems. The central problem is the fusion of data, achieving the loop "event, recognized symbols, structure of such symbols, interpretation, action, event again". These systems are, as for their construction, totally decomposable [*Mataric 1995*].

The two types of systems, reactive and perceptive, solve the well-stated problems and they are constructed typically for that. Their possible property of training is of type backing, in view to improve their reactive performance, this one tackling the gap between real event and recognized one. But it exists a third type of system, of a very different nature. For these new systems, inspirited of the living ones, it is not anymore the main question to solve well-known problems in formal ways, but rather to experiment possibilities of self-organization of some of their components, expressing problems they formulate for themselves, and treat in their own [*Brooks 1991*], [*Dautenhahn 1997*]. Such a system, when it copies the behavior of a living organism, must have the same reason to adopt some behavior that the living organism it simulates. These systems, called the *self-adaptive systems*, have radical differences with the two previous ones [*Cardon 2004*].

#### 4.1.3 The self-adaptive systems and the motivated action

A self-adaptive system is a system composed of two different strongly linked parts [C.f. Fig 2]:

- a substratum part, rational in way, managing the inputs and the reactive effects as well as the logical and rational automatic actions,

- a specific part deliberately representing the current situation of the system in its environment, according to some subjective characters and controlling the substratum part.

Such system is active for itself with the means of its structure indeed. The system part representing the current situation will be constituted of many entities in permanent reorganization action, adapting this internal organization at a time to the actions towards the environment and also to its own organizational state, like in the brain.

**4.2 The Self-adaptive system**

*A self-adaptive system is a system composed of a rational substratum and of a sub-system of representation of the current situation, controlling the substratum and formed of entities having the capacity of adaptive reorganization. This subsystem of representation adapts its organization at a time with the state of the environment, and so with its organizational state, following its own tendencies.*

Such a system can continuously construct representations of events, according to reasons that will be its own, according to its specific situation in its environment, according to the possibilities of its structure. The notion of representation, the internal object that represents a scene or a thing of the environment, is first. The notion of adaptativity is considered in the strong sense therefore. Determinants of this representation are constituted the first, according to the capacity of the system, according to its organizational memory, under some stimuli, while putting an aim toward something it takes into consideration. The system notices something that it aims. It conceives a situation, elaborates some plans, and after acts. The construction of such a system is not merely decomposable in very precise functional parts, because its active parts can hold roles that evolve during the generation of plans. Its fundamental property is the re-organization of its active parts [Dretske 1988]. A self-adaptive system can evidently degrade its structure in the one of a perceptive system, while tampering its organization and while having an automatic behavior.

**4.2.1 The notion of representation in a self-adaptive system**

*An artificial self-adaptive system, before to act, builds a purposeful representation of its environment, according to some own points of view. It acts while using this representation systematically. It set up itself in adequacy with its environment, in the sense that its representations have the tendency to consolidate its global internal state for adequacy with the evolutionary characters of the environment.*

Let's note that the three systems we have presented form an architectural hierarchy:

A self-adaptive system contains effectively subsystems that are only perceptive, every perceptive system containing some strictly reactive components. This hierarchical character corresponds to the evolution of the living organisms and also corresponds to the evolution of data processing systems towards the artificial life.

The part of the self-adaptive system that will produce representations and will use emotions will be composed of two distinct parts [Fig.C.4.1], as the brain in the body: a substrate sub-system and a specific component that represents the situation and emotions at a time. These two parts are:

- a reactive sub-system, the substrate that will seize information coming from the environment and will execute standards actions. It will regroup the Man Machine Interface (MMI), sensors and effectors, bases of knowledge and functional modules.
- a sub-system for the global control. This second sub-system has two parts:
  - a sub-system generating the current representation of the situation, that will be charged to build an effective representation of the environment and to construct the current plan of action, using composition of local plans.
  - an emotion generator sub-system, strongly coupled to the previous one that will explicitly express the fundamental emotional tendencies, while tampering the production of the representation system, altering the plan of action of the system in a subjective way.

### 4.3 Emotional tendencies and adaptability

The representational component of self-adaptive systems manages continuous structural modifications, driven by its tendencies that one must consider as unavoidable.
These systems must solve adequate reaction problems indeed, but while systematically taking account the satisfaction of their tendencies.

### 4.3.1 Emotional tendencies

*The self-adaptive systems are conceived so that they must satisfy the imperative tendencies that drive their behavior in a decisive manner. These tendencies are in various numbers and are contradictory. They will be called the emotional tendencies. They are the global light driving the modifications of the organization of the system of representation to give a subjective tendency.*

The emotional tendencies, for the natural organisms, are emotions that advise them to survive, to feed, to reproduce, to maintain themselves in satisfactory situation in their environment... That is these reasons that drive their behaviors, first while generating some appreciable representations of their world and also allowing to act for reasons on their environment. And with the necessity to behave and to act imposed by these emotional tendencies, they are brought to solve some varied problems with efficiently. But the resolution of problems is therefore, for these systems, a mean and not a goal.

One set up that the emotional tendencies present in the self-adaptive system must be numerous and must permit opposition and choice always. In the case where the system would have only one tendency or that all of these would be strongly in agreement and constituting, for example, a hierarchy with a permanent dominant need, the system would amount to a reactive one, without tendencies. So, the system will be complex.

### 4.3.2 Self-adaptive system and emotional tendencies

*A self-adaptive system is a system where the reason to function is driven by the satisfaction of emotional tendency, modifying strongly its representation of the environment. To satisfy these tendencies, the system must adapt its sub-system of representation of the situation and so modify its behavior. The structure of its representation system will be, for that reason, very plastic. The system will modify its representations while maintaining them in a domain compliant to its emotional tendencies.*

This type of adaptativity sound in an organizational way therefore, while specifying that the structure of the system of representation is fundamental to constantly drive the action. This sub-system re-organizes itself so that the global system stands in structural concordance situation with solicitations of the environment, while following pressures of its emotional tendencies. The fundamental emotional tendencies, to be able to finely modify the current plan of action of the system, will express themselves as the characters of some organizations of software agents constituting the system of representation. Indeed, a software agent has two characters, active and cognitive; it represents knowledge and action, and is neither a particle, nor variable or process only. The fundamental emotional tendencies will be the global characters driving reorganizations of agent's organizations. It is a strong organizational hypothesis, since we will represent tendencies by types of shapes expressing the movement of agent organizations [Cardon 2004]. It is the geometric hypothesis about emotional tendency expression in large software agent Organizations.

### 4.4 The fundamental geometrical hypothesis

*The geometrical hypothesis, with regard to the fundamental emotional tendency expression, consists to set up that the tendencies can be represented by some kinds of movements of geometric shapes in some dynamic space where speed up the organizations of software agents.*

# Chapter 5

## The multiagent approach of a system with an emotional behaviour

The architecture of a system generating emotions is radically different of an input - output one stepping levels of computation according to some predefined steps. We have to define specific inner-control of a system producing fuzzy states as the emotional ones, composed of proactive entities (entities that run for themselves), generating inner cycles of activities with specific rhythms, according to the real process of emotions in the brain and corresponding to different types of effective actions or movements. This architecture will essentially be founded on aggregation and breaking of software agents groups rather than formal neuron systems.

### 5.1 Expressing of an agent group

*Because they are proactive, organization of software agents can represent systems in a totally organizational way, where the form of the activities and the links between agents directly lead to an effective activity of the system, in a continuously adaptive reaction.*

#### 5.1.1 Basic computable component producing the emotion: the computable oscillator

The biologic presentation of the emotional activity describes the existence of neuron domains speeding up in loops, acting each other's for the propagation of flux of activations. We have to describe a system where the activity form is made of emergent feedback loops, positive, negative and additive. We precise the basic architectural element of the system with the following component, the computable oscillator:

**Computable oscillator**

*A computable oscillator is a software organization of agents whose activation forms quickly and by its own functioning many cycles of activity having specific intensity and speed.*

This notion spreads into the computable the one of systemic feedback loop. Such an oscillator is an organization of software agents that coordinates them, modifies the link of theirs activities, synchronizes some of them. The oscillator is formed by emergence of a self-kept structure distinct of the other agents of the organization. Such a group must emerge then to control itself, to pass from a uniform state to another where a looped process transforms a group into an oscillator. Mathematically, it is about an emergent sub-graph in the strongly coupled activation of an agent organization. Such an emergent oscillator leads to the adaptive activity the outputs of the system and must control other attempts of emergent loops.

The running system will be formed of a structured set of such elementary oscillators, permitting a global and local backing and the inhibition and the stop of the process. There is no central controller in the system but self-control distributed into the emerging components and their synchronization using negotiations. The system will function by self-control and self-regulation of its oscillators, with local limits cycles and a general faculty more or less conservative.

### 5.1.2 Typical element of the architecture: the adaptive component

One considers a system with a sub-system composed of interfaces and material components, and with a specific part generating representations and emotions, the two being very strongly linked. The generating system is composed, at the minimal level, of a set of agent organizations called the *aspectual agents*, able to easily produce by emergence groups of some computable oscillators. This architecture is typically evolutionary. From the inputs of the system, regrouped into classes according to data (the sensors for a robot), the system should produce lot of loops during its activation.

There is not initial state driving automatically to a specific state of reaction, but unceasing transformations driving to progressive actions. The very basic element of conception of the system is the aspectual agent [Cardon - Lesage 1998]. It is a software agent, typically proactive, whose role is at a time factual and symbolic. These agents serve to form some subsystems similar to those of the limbic one into the human brain. An oscillator is a group of synchronized agents that emerges and exert some power on its context. For the emergent and the control of this oscillator, we represent the behavior of every agent group by a mechanism of

self-observation founded on the geometric shape of the agents' activities. We are going to associate to the notion of behavior of all agent groups the one of geometric shape. We hear shape into the classic geometric sense of the term, like hyper-graphs.
Let's notice that agents being some rational entities, it is possible to associate them a precise notion of state characterized by values of specific vector.

### 5.2 Geometric state of an aspectual agent

*The state of an aspectual agent is the meaningful characters that permit to describe its current situation at a time and to predict its future behavior. This will be a specific vector.*

It is clear that one will always bring back each of these characters to an element of R. So, an agent's state will be a point of Rm if there are m characters defining the behavioral agent's state [Cardon 1999, 2004].

#### 5.2.1 Map of activity of an aspectual organization of agents [Lesage 2000]

*A map of activity of an organization of agents is a temporal representation of the geometrical set of the significant characters of the agent behaviors. This is a dynamic geometrical object.*

To use the notion of shape, that is to represent a map of agent activity by geometric forms, it will be necessary to first represent each agent by a vector of activity. The map of activity of an organization of agents will be then a set of points in the $R^m$ corresponding space, according to the m typical characters of each agent's behavior. The constitution of such a map is possible because agents are only rational entities. Their typical characters are expressed from their structure according to their actions only.
Using a classic mathematical transformation, we express the form of the agents activities as a graph, where each node is a group of similar vectors of agents and each link is the valuation of the qualified communications between groups of agents [Fig.C.5.1].

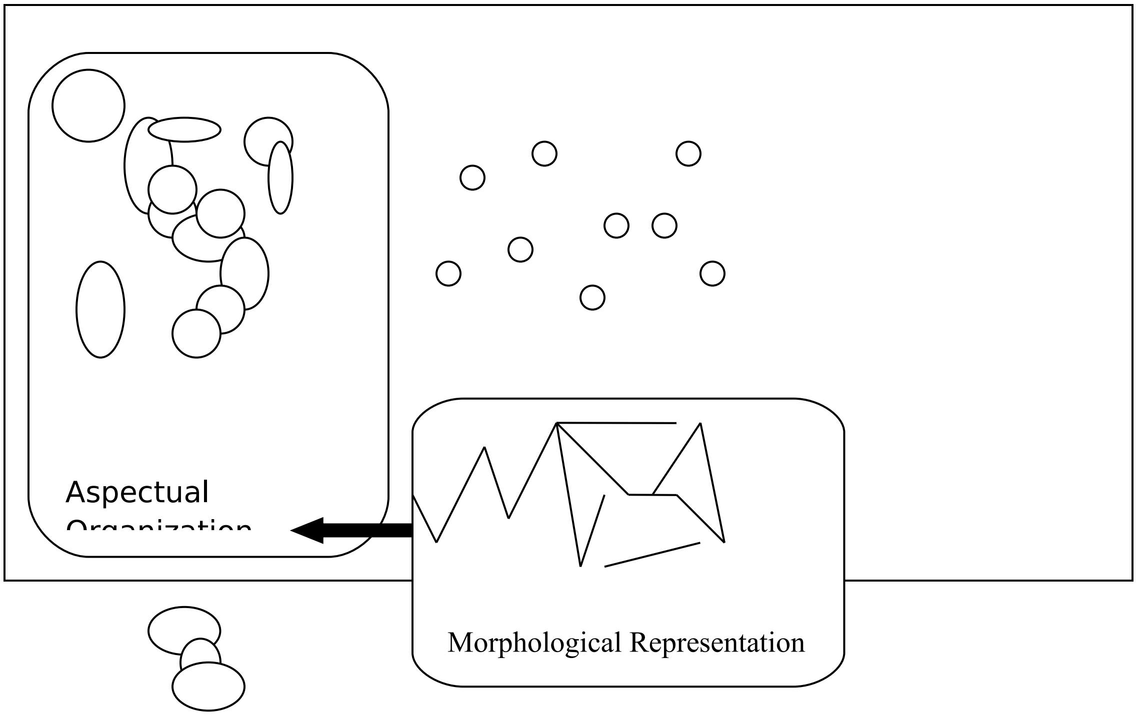



Fig.C.5.1.The Morphology expressing the aspectual agents of organization

To represent the behavior of a self-adaptive component organization with its geometric aspects, one must define a specific dynamic space, the *morphological space* [Cardon 1999]. The so-called *agents of morphology*, distinguishing groups and extracting forms will achieve the assessment of the active shape of agent groups [Fig.C.5.2].
Finally, another organization of agents, after the aspectual and the one of morphology, is going to take in consideration the state of the aspectual organization to achieve an analysis of the aspectual agents behavior. It is about representing the sense of the activation of the aspectual agent organization with its geometrical characters and produced by the agents of morphology. The *agents of analysis* are going to provide a cognitive view of that has been expressed by the geometric and semantics information coming from the morphology agents, above the aspectual agent landscape, an interpretation of graphs indeed [Fig.C.5.2].

**5.2.2 Rational functioning of the system**

*The system, reduced to the aspectual agents doing the foreseen rational tasks, with the agents of morphology expressing the shape of the aspectual activation and with the agents of analysis achieving the synthesis of the aspectual agents functioning to plan the reactions, will be qualified of rational functioning, that to be-to-say under inner-control but without any emotion indeed.*

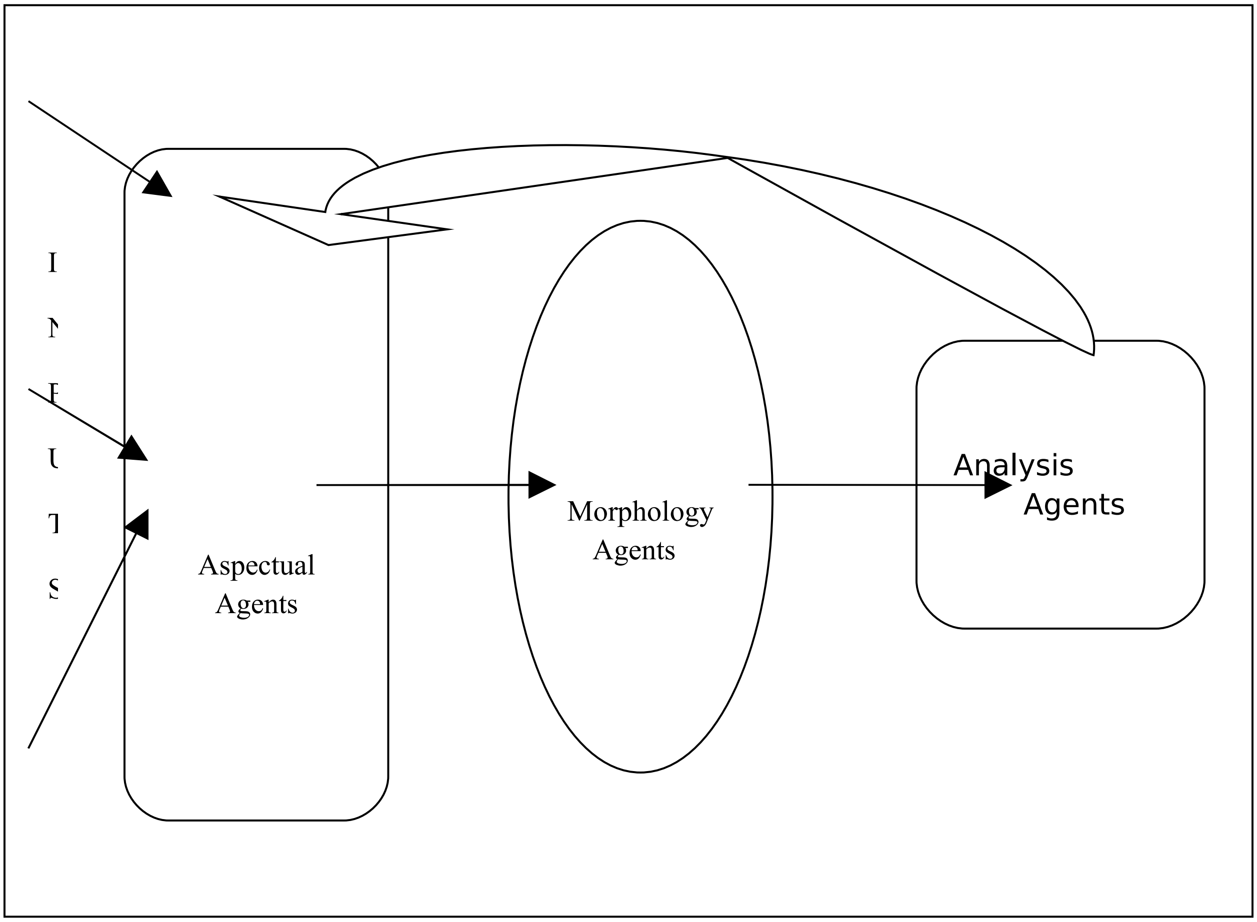



Fig.C.5.2.The three rational organizations of a self adaptive component

We can precise the notion of self-controlled representation with a specific coarse grain component: the adaptive component [Fig.C.5.2]. This notion is going to permit to exceed the strictly rational behavior of the system and to define the emotions.

**5.3 Self-adaptive component**

*A self-adaptive component is an aspectual agent group linked with a morphological representation and an analysis agent group allowing to describe and to interpret its own specific* *activity, while controlling it in order to make emerge a computable oscillator. Such a component is emergent and does not exist at the conception level therefore.*

A linked set of self-adaptive components could express, by its functioning and in a very dynamic way, the emotional states.

### 5.3.1 Basic organization of self-adaptive components: the aspectual agents

According to the inputs of the system (robot sensors and effectors for example), the agent organizations generating behavioral decisions and emotions will be composed of self-adaptive components, every one being an organization of aspectual agents centered on a symbolic and geometric role bound to a morphological characterization and provoking an organizational answer taking into account the information coming from its context. The behavior of this set of components will be appreciated in two ways:

- the action of agents on the output of the system (flux, intensity…), what will leads to an emotive effect.
- the shape of the organization, witch will produce an observation of emotion like a mental photography. We base the architecture of the system on a representation of situations theory, where the reactions of the system necessarily pass by the continuous erecting of an adapted representation, the current one, that is a computable construct about the sense of the action of numerous groups of agents.

  We now specify the different categories of aspectual agents. The self-adaptive components are constructed using light software agents. One kind of aspectual agents is bound to the outputs of the system (for example the embodiment parts of the robot). While preserving the robotic terminology, these agents are of two types:

    - the *sensor aspectual agents*, SAA: they interpret the information coming from the environment,
    - the *effectors aspectual agents*, EAA: they propose an effective action of the output parts. Their structure is the one of a classical light agent [Krogh 1995], [Campagne 2004].

### 5.4 The SAA, the sensors aspectual agents

They are aspectual agents that achieve an automatic interpretation "cognitive and variable in amplitude" of this one, for each new information coming from the environment in the input buffers. The knowledge module of these agents achieves the interpretation. These two modules, module of knowledge and the behavioural one, are composed of rules.

### 5.5 The EAA, aspectual effectors

The EAA operate on the environment of the system by efficient actions (on the MMI or with the body of the robot). These agents are solicited by the aspectual and analysis ones and propose actions. For achieved these actions, the EAA must coordinate them in a social way (prey - predatory principle). Rules of action of the EAA are multilevel and are, for example, the followings:

- ❖ if the state of several solicited aspectual agents is of such category, then to make the proposition of such immediate action (with a notion of hierarchy and priority)…
- ❖ if the proposition of action is reinforced by some agents of analysis, then to act of such manner…
- ❖ if an alarm is triggered then to propose to hire such action with such priority and such speed of realization then…

These aspectual agents must define very local plans: that is the numeric synthesis of their past actions, proposed current action, possible values of future actions. Let's notice that this notion of plan is strictly local and only defined by the values of characteristic variables, permitting to memorize the details of the local actions.

# Chapter 6

## The Artificial Consciousness and Emotions generated

### 6.1 The Artificial Emotions

The principle of generation of an emotion is therefore the next one. The input sensors, via agents of interface, make to speed up structuring agents that make speed up the corresponding morphology agents. The incentive organization generates a specific form in the morphology that is to reach: the incentive morphology. The current morphology describing the aspectual activity has to transform itself into the incentive morphology. For that, the aspectual agents have to make some specific activity. The analysis agents control them in that way. The characters of the incentive and of the current morphologies are the determinants for the emotion. If the current morphology has a complicated shape, far away from the incentive morphology, the system expresses a state of tension that it is going to try to systematically reduce. This tendency

to the reduction corresponds precisely to the discharge of energy that evoked S. Freud in the antagonistic action of impulses toward states of quietude [Freud 1966].

The pleasure is the case where the current morphology reaches slowly but steadily the incentive one. After the activity of the interface agents, the aspectual agent organization generates some global activity, a global process geometrically represented by the organization of morphology. The incentive morphology generates a singular pick that activated some corresponding aspectual agents. The current descriptive morphology of the aspectual agents, considered by projection in

$R^2$, leads to the constitution of a lot of successive picks, from progressive features, and this shape converges toward the pick of the incentive morphology. In its looped functioning, agents of analysis are going to reduce the morphology to only one prominent shape. There will be reduction of tension and sensation of pleasure, as it describes S. Freud.

### 6.2 Conclusion

The emergence of emotions in an data processing system, or in a robot, has been presented as the stabilization in a while of activities in a complex multiagent system, in an organizational and constructivist way. This emergence, represented by functioning with periodic loops of activities, is the global state of a set of agent organizations.

To achieve this type of functioning, some agents, the agents of morphology, represent behaviors of aspectual agents that themselves represents the minimal elements of significance. The fact that the activity of a system is endowed of emotions founds finally on a strong coupling process between the computations that organize it and the representation of computations that permits the own-control of group of agents.

The importance of such a process of coupling, binding the parts to the whole, binding groups of agents to their significance represented by agents of morphology, is strong. It is the fundamental principle of the functioning of systems that we called self-adaptive, and are to day the alone able to self-control complex systems. It generalizes the notion of feedback and systemic loop and go to the realizations of autonomous system essentially producing states by emergence.

A system generating emotions while using a bodily substratum, while proceeding to an organizational emergence thus has, according to us, a complex structure at the level of organization. But this organization can also produce a representation of itself, of its own morphology leading to the notion of "*its own body*". System can use this morphology as an engagement to act since at a time. Using geometric and cognitive aspects, that is the sign of the organizational semiotics summarizing at the same time the process of reorganization and its result. And the result of this constructive self-observation can be delivered by the system to all the observers. In this sense, such a system can express itself rather than only display values merely. The difference then, brings in such a system that expresses itself subjectively according to its intentions and another one that would proceed to displays complicated information very well adapted for sound the user, is large and makes a kind of rupture in the field, very vast, of the present computer science.

# III. The Artificial Brain in Reality:

# The Blue Brain Project of 2005

# Chapter 7

## The Blue Brain Project

### An Insight

A network of artificial nerves is evolving right now in a Swiss supercomputer. This bizarre creation is capable of simulating a natural brain, cell-for-cell. The Swiss scientists, who created what they have dubbed "Blue Brain", believe it will soon offer a better understanding of human consciousness. This is no sci-fi flick; it's an actual 'computer brain' that may eventually have the ability to think for itself. Exciting? Scary? It could be a little of both.

The designers say that "Blue Brain" was willful and unpredictable from day one. When it was first fed electrical impulses, strange patterns began to appear with lightning-like flashes produced by 'cells' that the scientists recognized from living human and animal processes. Neurons started interacting with one another until they were firing in rhythm. "It happened entirely on its own," says biologist Henry Markram, the project's director. "Spontaneously."

The project essentially has its own factory to produce artificial brains. Their computers can clone nerve cells quickly. The system allows for the production of whole series of neurons of all different types. Because in natural brains, no two cells are exactly identical, the scientists make sure the artificial cells used for the project are also random and unique.

Does this 'Brain' have a soul? If it does, it is likely to be the shadowy remnants of thousands of sacrificed rats whose brains were almost literally fed into the computer. After opening the rat skulls and slicing their brains into thin sections, the scientists kept the slices alive. Tiny sensors picked up individual neurons, recorded how the cells fired off neurons and the adjacent cells' responses. In this way the scientists were able to collect entire repertoires of actual rat behavior- basically how a rat would respond in different situations throughout a rat's life.

The researchers say it wouldn't present much of a technological challenge to bring the brain to life. "We could simply connect a robot to the brain model," says Markram. "Then we could see how it reacts to real environments."

Although over ten thousand artificial nerve cells have already been woven in, the researchers plan to increase the number to one million by next year. The researchers are already working with IBM experts on plans for a computer that would operate at inconceivable speeds – something fast enough to simulate the human brain. The project is scheduled to last beyond 2015, at which point the team hopes to be ready for their primary goal: a computer model of an entire human brain.

### 7.1 What is the Blue Brain Project all about?

The Blue Brain project is the first comprehensive attempt to reverse-engineer the mammalian brain, in order to understand brain function and dysfunction through detailed simulations.

### 7.2 How it came to be?

In July 2005, EPFL and IBM announced an exciting new research initiative - a project to create a biologically accurate, functional model of the brain using IBM's Blue Gene supercomputer. Analogous in scope to the Genome Project, the Blue Brain will provide a huge leap in our understanding of brain function and dysfunction and help us explore solutions to intractable problems in mental health and neurological disease.

At the end of 2006, the Blue Brain project had created a model of the basic functional unit of the brain, the neocortical column. At the push of a button, the model could reconstruct biologically accurate neurons based on detailed experimental data, and automatically connect them in a biological manner, a task that involves positioning around 30 million synapses in precise 3D locations.

In November, 2007, the Blue Brain project reached an important milestone and the conclusion of its first Phase, with the announcement of an entirely new data-driven process for creating, validating, and researching the neocortical column.

## 7.3 What the Blue Brain Project is not

The Blue Brain Project is an attempt to reverse engineer the brain, to explore how it functions and to serve as a tool for neuroscientists and medical researchers. It is not an attempt to create a brain. It is not an artificial intelligence project. Although we may one day achieve insights into the basic nature of intelligence and consciousness using this tool, the Blue Brain Project is focused on creating a physiological simulation for biomedical applications.

## 7.4 Building the Microcircuit

### 7.4.1 Modeling Neurons

Neurons are not all alike - they come in a variety of complex shapes. The precise shape and structure of a neuron influences its electrical properties and connectivity with other neurons. A neuron's electrical properties are determined to a large extent by a variety of ion channels distributed in varying densities throughout the cell's membrane. Scientists have been collecting data on neuron morphology and electrical behavior of the juvenile rat in the laboratory for many years, and this data is used as the basis for a model that is run on the Blue Gene to recreate each of the 10,000 neurons in the NCC.

### 7.4.2Modeling connections

To model the neocortical column, it is essential to understand the composition, density and distribution of the numerous cortical cell types. Each class of cells is present in specific layers of the column. The precise density of each cell type and the volume of the space it occupies provides essential information for cell positioning and constructing the foundation of the cortical circuit. Each neuron is connected to thousands of its neighbors at points where their dendrites or axons touch, known as synapses. In a column with 10,000 neurons, this translates into trillions of possible connections. The Blue Gene is used in this extremely computationally intensive calculation to fix the synapse locations, "jiggling" individual neurons in 3D space to find the optimal connection scenario.

### 7.4.3 Modeling the column

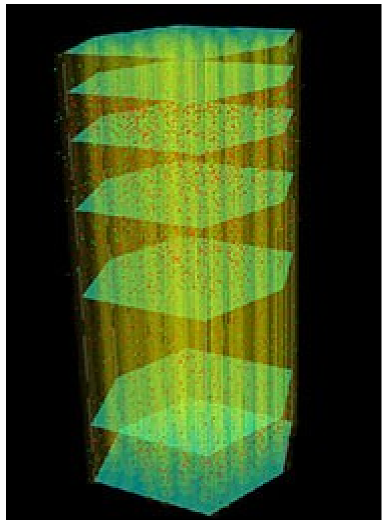
The result of all these calculations is a re-creation, at the cellular level, of the neocortical column, the basic microcircuit of the brain. In this case, it's the cortical column of a juvenile rat. This is the ***only*** biologically accurate replica to date of the NCC - the neurons are biologically realistic and their connectivity is optimized. This would be impossible without the huge computational capacity of the Blue Gene. A model of the NCC was completed at the end of 2006.

In November, 2007, The Blue Brain Project officially announced the conclusion of Phase I of the project, with three specific acheivements:

1. A new **modeling framework** for automatic, on-demand construction of neural circuits built from biological data

2. A new **simulation and calibration process** that automatically and systematically analyzes the biological accuracy and consistency of each revision of the model

3. The first **cellular-level neocortical column** model built entirely from biological data that can now serve as a key tool for simulation-based research



C.7.1.The Neurocortical Column (NCC)

**7.5 Simulating the Miocrocircuit**

Once the microcircuit is built, the exciting work of making the circuit function can begin. All the 192 processors of the Blue Gene are pressed into service, in a massively parallel computation involving the complex mathematical equations that govern the electrical activity in each neuron when a stimulus is applied. As the electrical impulse travels from neuron to neuron, the results are communicated via inter-processor communication (MPI). Currently, the time required to simulate the circuit is about two orders of magnitude larger than the actual biological time simulated. The Blue Brain team is working to streamline the computation so that the circuit can function in real time - meaning that 1 second of activity can be modeled in one second.

**Interpreting the results**

Running the Blue Brain simulation generates huge amounts of data. Analyses of individual neurons must be repeated thousands of times. And analyses dealing with the network activity must deal with data that easily reaches hundreds of gigabytes per second of simulation. Using massively parallel computers the data can be analyzed where it is created (server-side analysis for experimental data, online analysis during simulation).

Given the geometric complexity of the column, a visual exploration of the circuit is an important part of the analysis. Mapping the simulation data onto the morphology is invaluable for an immediate verification of single cell activity as well as network phenomena. Architects at EPFL have worked with the Blue Brain developers to design a visualization interface that translates the Blue Gene data into a 3D visual representation of the column. A different supercomputer is used for this computationally intensive task. The visualization of the neurons' shapes is a challenging task given the fact that a column of 10,000 neurons rendered in high quality mesh (see picture) accounts for essentially 1 billion triangles for which about 100GB of management data is required. Simulation data with a resolution of electrical compartments for each neuron accounts for another 150GB. As the electrical impulse travels through the column, neurons light up and change color as they become electrically active.

A visual interface makes it possible to quickly identify areas of interest that can then be studied more extensively using further simulations. A visual representation can also be used to compare the simulation results with experiments that show electrical activity in the brain. This calibration - comparing the functioning of the Blue Brain circuit with experiment, improving and fine-tuning it - is the second stage of the Blue Brain project, expected to be complete by the end of 2007.

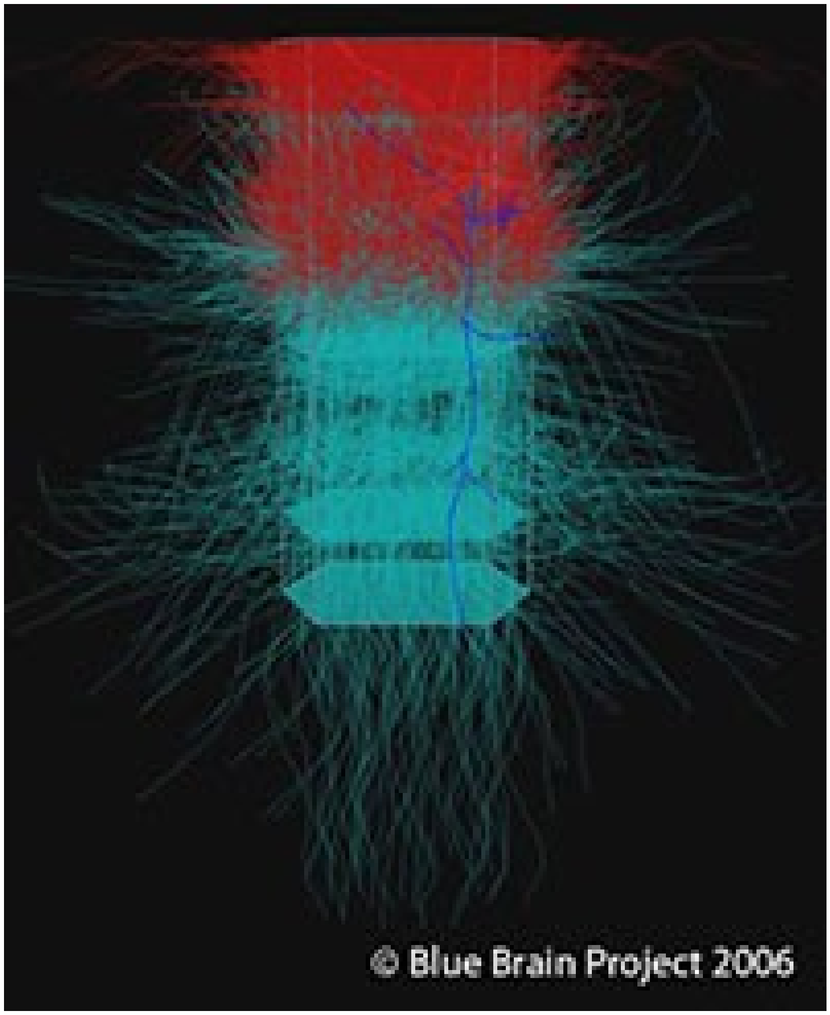



Fig.C.7.2.Visual Interface showing fluctuations in the NCC

**7.6 What comes next in the Blue Brain Project?**

Phase I marks the completion of a proof-of-principle simulation-based research process that has resulted in a cellular-level model of the neocortical column. We have achieved biological fidelity such that the model itself now serves as a primary tool for evaluating the consistency and relevance of neurobiological data, while providing guidance for new experimental efforts. These new data will serve to further refine the neocortical column model. The assembled process allows neuroscientists to investigate scientific questions by integrating the available experimental data and evaluating hypotheses of network dynamics and neural function.

The completion of phase I provides the basis now for increasing the resolution of the models down to the molecular level and expanding the size of the models towards the whole brains of mammals.

In the future, information from the molecular and genetic level will be added to the algorithms that generate the individual neurons and their connections, and thus this level of detail will be reflected in the circuit's construction. The simulations can be used to explore what happens when this molecular or genetic information is altered -- situations such as a genetic variation in particular neurotransmitters, or what happens when the molecular environment is altered via drugs.

The project will continue to expand and will necessarily involve additional scientists and research groups from around the world.

## 7.7 Why the Blue Brain Project is Important

### 7.7.1 Gathering and Testing 100 Years of Data

The most immediate benefit is to provide a working model into which the past 100 years knowledge about the microstructure and workings of the neocortical column can be gathered and tested. The Blue Column will therefore also produce a virtual library to explore in 3D the micro architecture of the neocortex and access all key research relating to its structure and function.

### 7.7.2 Cracking the Neural Code

The Neural Code refers to how the brain builds objects using electrical patterns. In the same way that the neuron is the elementary cell for computing in the brain, the NCC is the elementary network for computing in the neocortex. Creating an accurate replica of the NCC which faithfully reproduces the emergent electrical dynamics of the real microcircuit, is an absolute requirement to revealing how the neocortex processes, stores and retrieves information.

### 7.7.3 Understanding Neocortical Information Processing

The power of an accurate simulation lies in the predictions that can be generated about the neocortex. Indeed, iterations between simulations and experiments are essential to build an accurate copy of the NCC. These iterations are therefore expected to reveal the function of individual elements (neurons, synapses, ion channels, receptors), pathways (mono-synaptic, disynaptic, multisynaptic loops) and physiological processes (functional properties, learning, reward, goal-oriented behavior).

### 7.7.4 A Novel Tool for Drug Discovery for Brain Disorders

Understanding the functions of different elements and pathways of the NCC will provide a concrete foundation to explore the cellular and synaptic bases of a wide spectrum of neurological and psychiatric diseases. The impact of receptor, ion channel, cellular and synaptic deficits could be tested in simulations and the optimal experimental tests can be determined.

### 7.7.5 A Global Facility

A software replica of a NCC will allow researchers to explore hypotheses of brain function and dysfunction accelerating research. Simulation runs could determine which parameters should be used and measured in the experiments. An advanced 2D, 3D and 3D immersive visualization system will allow "imaging" of many aspects of neural dynamics during processing, storage and retrieval of information. Such imaging experiments may be impossible in reality or may be prohibitively expensive to perform.

### 7.7.6 A Foundation for Whole Brain Simulations

With current and envisage able future computer technology it seems unlikely that a mammalian brain can be simulated with full cellular and synaptic complexity (above the molecular level). An accurate replica of an NCC is therefore required in order to generate reduced models that retain critical functions and computational capabilities, which can be duplicated and interconnected to form neocortical brain regions. Knowledge of the NCC architecture can be transferred to facilitate reconstruction of sub cortical brain regions.

### 7.7.7 A Foundation for Molecular Modelling of Brain Function

An accurate cellular replica of the neocortical column will provide the first and essential step to a gradual increase in model complexity moving towards a molecular level description of the neocortex with biochemical pathways being simulated. A molecular level model of the NCC will provide the substrate for interfacing gene expression with the network structure and function. The NCC lies at the interface between the genes and complex cognitive functions. Establishing this link will allow predictions of the cognitive consequences of genetic disorders and allow reverse engineering of cognitive deficits to determine the genetic and molecular causes. This level of simulation will become a reality with the most advanced phase of Blue Gene development.

## EPILOGUE

The Artificial Brain, howsoever crazy or deep-rooted in fiction it may sound is indeed an inspirational theme. Success or failure, whatever may the end results be, I am sure that any research of the brain will only force us into accepting the presence of the one who is able to do so.

` A system that may be able to think for itself, feel emotions that may vary from happiness to sadness, love to anger and maybe even sensuality is indeed a technology that seems light years away, but it wouldn't be wrong if we believe, that if we start today, the lesser the number of light years to its realisation.

Apart from the sheer fantasy of a system that can reflect, understand and respond, this advanced technology may help us in understanding and even provide solutions for people suffering from ailments that impair parts of the brain. Just like the vast skies and oceans have been conquered, in the same way, it could be possible that we may actually unlock the mysteries of the mind.

I, as a student, have been forever inspired after being involved in this particular subject of which very little has been studied or written about. It's exciting to think of all the possibilities that could open up from this discovery. Answers to mental retardation and brain abnormalities can be solved at earlier stages. Diseases like Alzheimer's and Parkinson's that rid a person of his dignity could be eradicated totally. The list is endless, everything from brain disorders to the truths about the mysteries of dreams and dejavu(s) could be found out.

Pragmatism is no longer the issue as far as science is concerned. So, an Artificial Brain is not at all matter for just science fiction magazines, it could be real and the Blue Brain Project is a sure and steady step towards it.

## BIBLIOGRAPHY


- Bechara, H.D., Damasio, A.: Emotion, decision making and the orbito frontal cortex. Cerebral Crotex **10** (2000)

- [Brooks 1991] Brooks R., *Intelligence without reason*, in Proc. of the 1991 International Joint Conference on Artificial Intelligence, p. 569 - 591, 1991.
- [Campagne 04] Campagne J.C., Cardon A.,*Using Morphology to Analyse and Steer Large Multi-Agents Systems at Runtime*, Selmas'04 IEEE, Edinburg, Scotland, 24-25 May 2004.
- [Campagne 04] Campagne J.C., Cardon A., Collomb E., Nishida T., *Massive Multi-Agent System Control,* FAABS III 2004, IEEE Workshop on Formal Approaches on Agents-based Systems, NASA Goddard Space Center, Greenbelt MA, USA, April 2004,
- [Cardon - Lesage 1998] Cardon A., Lesage F., *Toward Adaptive Information Systems : considering concern and intentionality*, KAW'98, Banff, Canada, 17-23 Avril 1998.
- [Dautenhahn 1997] Dautenhahn K*., Biologically inspired robotic experiments on interaction and dynamic agent-environment coupling*, in Proc. Workshop SOAVE'97, Ilmenau, p. 14 - 24, September 1997.
- [Dretske 1988] Dretske F., *Explaining Behavior*, MIT Press, 1988.
- [Freud 1966] Freud S., *The Complete Psychological Works of S. Freud*, J. Strachey, The Hogarth Press, London, 1966.
- [Lesage & al. 1999] Lesage F., Cardon A., Tranouez P., *A multiagent based prediction of the evolution of knowledge with multiple points of view*, KAW' 99, Communication