Doc. No. 03-602-H8
Rev. 0 - JUNE 2005

ESA ESTEC
Noordwijk, The Netherlands

Bionics and Space System Design

Adaptability versus Stability

D'APPOLONIA

ESA ESTEC
Noordwijk, The Netherlands

Bionics and Space System Design

Adaptability versus Stability

Prepared by	Signature	Date
Raimondo De Laurentiis		
Donato Zangani		
Verified by	**Signature**	**Date**
Stefano Carosio		
Approved by	**Signature**	**Date**
Andrea Barbagelata		

Rev.	Description	Prepared by	Verified by	Approved by	Date
0	First Issue	RDL/DMZ	SMC	AB	June 2005

TABLE OF CONTENTS

Page

TABLE OF CONTENTS
(Continuation)

LIST OF TABLES

LIST OF FIGURES

REPORT
FINAL REPORT

1 BACKGROUND

Biomimicry is a multi-disciplinary science involving a wide diversity of other domains like electronics, informatics, medicine, biology, chemistry, physics, mathematics, and many others. However, it is quite unusual to find key people or expertise centres that have cognition and expertise in all these disciplines as a whole. Therefore, there is a need for the establishment of a capillary network of contacts through Europe and elsewhere that will enable to reach also those academic centres, which are not much visible due to their reduced dimension or recent origin. Additionally, although some peculiar conditions characterizing space environments can be similarly encountered on earth (e.g. desert zones) and specific solutions found within these terrestrial contexts can be adapted to space conditions, there is a majority of cases, which are subject to conditions which are broadly different from those encountered on earth (e.g. gravity absence). Therefore, the biomimetic approach in the space sector results more complex and has to be considered in a multidisciplinary and cross-sectorial framework to overcome barriers. The problems to be addressed to exploit the potential of the biomimicry approach in the space domain can be summarized as follows:

- biomimicry has become a real science only in recent years and therefore there is no consolidated co-operation environment with space engineers;
- research in biomimicry across Europe and Canada and more generally at world wide level is scattered and fragmented, it is not easy to locate the proper academic experts for a given space application;
- biomimicry is a multi-disciplinary science and it requires several expertise which is difficult to locate in the same organization;

- o some databases with information about possible natural phenomena, biomimetic products, ongoing biomimetic research, biomimetic researchers, published articles exist, but they lack a systematic and a large-scale exploration of the potential of nature in view of applications in engineering, especially as far as the space domain is concerned;
- o in current knowledge-basis the abstraction of the biological functionality is missing, therefore solutions inspired by nature are sporadic and random-governed;
- o space conditions are completely different from life forms habitats and space engineers are so far not fully aware of applications of biomimetics.

Therefore, the overall objectives of the study consists in the development of a co-operation platform between space and biomimicry experts in order to bridge current gaps that exist for an effective application of natural mechanisms and phenomena in space system design and to foster the development of a new generation of space systems. This has been achieved by:

- o performing a comprehensive collection and review of information concerning attempts made since today in Europe and elsewhere in finding solutions through a biomimic approach, including an insight into planned research activities and trends;
- o developing a detailed biomimicry knowledge map that allows to identify expertise and competencies in ESA member states and elsewhere;
- o providing an overview of the unique characteristics and properties of various life forms found in nature (e.g. animals, plants, etc) and to ascertain whether these characteristics could be an inspiration to create innovative space systems;
- o conceptualising several innovative space systems and components which incorporate the design, features and mechanisms of nature's life forms.

All the gathered information have been implemented into a database which is available online at www.bionics2space.org. The added value of the database is in the deep analysis made on each biological system described, supported by literature and reference articles, patents, etc.

The project group has then been focusing on the analysis of the information collected and on whether any of these biological principles might hold potential for application to the design of space systems or provide solutions to space-related technical challenges.

Therefore, the project group has identified twelve different cases in which the application of biological principles could bring a real added value to the solution of technical constraints within the space field. The identified case studies are reported below:

- deployable digging mechanism for sampling below planetary surfaces;
- energy storage structures for deployable systems;
- rigidisation of deployable structures;
- smart swarm on mars;
- robust biologically inspired navigation techniques;
- planetary exploration with free energy (based on sun flowers);
- adaptive and versatile biologically inspired locomotion control;
- balance between adaptability and stability;
- automatic self-assembly in space;
- landing and planetary exploration;
- energy storage structures for deployable systems;
- planetary exploration with free energy (based on dandelion seeds).

Such work has set the base from which a more detailed analysis has been performed: for each of the topics, a responsible among the Bionics Expert Team has been identified; such expert has been in charge of providing to the partners the assessment of the idea of application.

The results of such detailed analysis have been presented in the framework of the Bionics workshop held in ESTEC on November 2004. The output of such event has been the selection of four case studies which have been further assessed by proposing first attempts of engineering solutions inspired by nature. Such case studies are the following:

- energy storage structures for deployable systems;
- case study on adaptability versus stability;
- deployable digging mechanism for sampling below planetary surfaces;
- landing and planetary exploration.

In this report the work undertaken for “Adaptability versus Stability” case study is described.

2 ADAPTABILITY VERSUS STABILITY

2.1 INTRODUCTION AND PROBLEM STATEMENT

For their survival, animals need to maintain a fine balance between adaptability and stability, and this at multiple levels and multiple time scales. Clearly, stability is important in order to maintain the animal's general structure and metabolism, and to keep it as a living entity despite changes in its environment. Similarly, adaptability is important when significant changes occur in the environment in order to adjust the structure and metabolism of the animal accordingly and to keep it viable. In a dynamic environment, an excess of either adaptability or stability can be damaging or even fatal. To take a simple example, an animal whose feeding habits and digestive system does not adapt quickly enough to seasonal changes from winter to summer might die during a season change. Conversely, an animal whose feeding habits and digestive system changes too quickly with the smallest variation in the environment (e.g. changes between day and night, or small variations of location) might waste resources while adapting continuously, and might risk a total disruption (chaos) due to the strong reaction to fluctuating inputs.

Many space related missions, such as the maintenance of a space station and a planetary exploration mission, face the same dilemma. One would like to have systems that self-organize, self-repair, and self-regulate using adaptation mechanisms, but at the same time maintain some notion of stability for carrying out the programmed tasks.

In this document, we first review how the right balance between adaptability and stability is maintained in different biological systems, and what types of techniques have been developed inspired from those systems (Paragraph 2.2). In particular we make an overview of the principles related to homeostasis, gene regulatory networks, bone generation, and reinforcement learning. We then present three specific case studies in the domains of locomotion control (Paragraph 2.3), artificial immune systems (Paragraph 2.4) and robot swarms (Paragraph 2.5). We conclude the document with a short general discussion and suggestions for future projects.

Such approach, enabled us to cover all possible levels of hierarchy in biological systems: sub-organisms (immune systems), organism (locomotion) and super-organism level (swarm). This enabled us to describe all possible set of techniques, achievements and possible future development as complete as possible.

2.2 PRINCIPLES, GENERAL TECHNIQUES, AND SYSTEMS

The definitions of stability and adaptability vary from one scientific field to the other. In this document, we will use the following meanings. We see *stability* as the ability to maintain, and to return to, a steady-state behavior despite changes in the environment. The steady-state behavior can be static or dynamic. We see *adaptability* as the ability to change the structure and/or the functioning of a system according to changes in the environment. By *system*, we mean a particular biological entity (e.g. a cell or an organism) or an artifact (e.g. a robot or a space station), and we assume that there is a (clear) boundary between the system and its environment, e.g. the membrane of a cell. By *environment* we describe the medium which is outside the system under study, and which can possibly contain other systems (e.g. other cells). We assume that there are *interactions* between the system and the environment, with the system changing states of the environment, and the environment changing states in the system.

Using these definitions, it is clear that the processes of stability and adaptability have conflicting effects on a system. In short, stability "cancels out" changes in the environment such as to maintain the structure and functioning of a system, while adaptability "makes use of" changes in the environment to modify the structure and functioning of the system. How to correctly balance the two is intrinsically linked to the *viability* (or *performance*) of the system. Ideally the balance between stability and adaptability should be continuously adjusted during the life time of the system such as to optimize its viability/performance. But the notions of *viability* and *performance* are difficult to define generally since they closely depend on the type of system that is studied. One could say that *viability* corresponds to maintaining the system in good working conditions; e.g. for a cell, viability means maintaining variables such as temperature, acidity level, concentrations of different chemicals, etc, within some metabolic ranges and avoiding lethal values. *Performance* is a notion that is related with the realization of a task (e.g. by an artifact) and is a measure of how well the task is realized; for example, the performance of a mining robot could be the amount of ore that it is able to extract and bring back to a base station. In general, the optimization of the viability/performance of a system corresponds to a multi-objective optimization problem.

The balance between adaptability and stability is maintained at multiple levels in biological systems, from gene networks, to cells, to organs, to living organisms, and even societies of organisms. Examples of systems that exhibit this balance include the metabolisms of cells, the generation of bone structure, the immune system, learning by trial and error, the combination of memory and forgetting, the organization of a bee hive, to name a few. In the sections below we will review some of these systems, before presenting in more detail three specific case studies in the domains of locomotion control, artificial immune systems and robot swarms.

The balance between adaptability and stability in biological systems is explicitly explored in the domains of homeostasis and autopoiesis (see below). It is otherwise implicitly addressed in multiple fields in biology (e.g. cell metabolism) and medicine (e.g. physiology). In fact, the balance is most often studied in cases where the mechanisms maintaining it fail: examples include cancer (i.e. a repair mechanism which goes uncontrolled), HIV (a defense mechanism which becomes deficient), and osteoporosis (a repair mechanism which becomes too weak).

In engineering, the balance between adaptability and stability is mainly studied in fields such as control theory and machine learning. Reinforcement learning, a sub-area of machine learning, for instance specifically addresses the issue of balancing exploration (i.e. trying out new things) versus exploitation (i.e. using what you know), which is directly related to balancing adaptability and stability.

2.2.1 Brief overview of biological examples

This section makes an overview of the principles related to homeostasis, gene regulatory networks, bone generation, and reinforcement learning. Two specific case studies in the domains of locomotion control and artificial immune systems will be presented in more detail in the next sections.

Principles of homeostasis

It might be argued that to break new ground in terms of generating complex, adaptive, autonomous and crucially: *self-organizing* computational behavior all these properties are required for the implementation of systems capable of generating the type of behavior sought by researchers in fields such as robotics, artificial intelligence and operating system design. With this in mind we wish to focus on one of the most impressive abilities of living organisms: their ability to ensure a reasonably stable internal state despite wildly changing external environmental factors. This property, often termed homeostasis, is a major contributor to an organism's autonomy, and is the biological embodiment of the type of behavior described above.

The concept of homeostasis was first advanced by the physiologist Walter Cannon (Cannon, 1932). Drawing on earlier work of Claude Bernard, Cannon advanced the concept of homeostasis as a set of complex physiological systems that act to maintain the internal state of an organism. In particular Cannon suggested that negative feedback may play an important role in the regulation of homeostatic mechanisms. In recent times, the study of homeostasis is often linked to the study of the autonomic nervous system, which controls heart muscle, smooth muscle, and exocrine glands in the body (Iversen et al., 2001). Artificial homeostasis takes inspiration from this ability of all living organisms to maintain a stable internal state in response to environmental factors. In addition, this approach provides flexibility through the specification of an appropriate level of granularity with which to tackle the problem of stability. (Neal & Timmis, 2005). A related and potentially relevant concept is that of autopoiesis (Maturana and Varela, 1980; Mingers 1994). Autopoiesis is concerned, in part, with how dynamic systems interact in such a way as to continually maintain the system. These concepts are particularly relevant to autonomous, adaptive or intelligent agents (Quick, 2003; Di Paolo, 2005).

Gene regulatory networks

Gene regulatory networks (GRN) describe the interaction of genes and gene products. Underlying these networks is the observation that activated genes in organisms cause the production of various molecules (the gene is translated into a protein which may cause the synthesis of other products) and some of the molecules produced are capable of binding to DNA and affecting the transcription of the genes, either facilitating or inhibiting their action.

Bone generation

Bone is an interesting example of a biological system that presents a fine balance between adaptability and stability. Bones have two structurally different parts, *cortical bone* on the outside and *trabecular* bone on the inside. The former is dense and compact, while the latter has a honeycomb of vertical and horizontal bracing spars. This structure is a beautiful example of Nature's way of maximizing strength while minimizing weight, with the honeycomb structure providing amazing strength.

Despite its image of an inert and solid material, bone is living, growing tissue. Throughout an organism's lifetime, old bone is removed (resorption) and new bone is added to the skeleton (formation). During childhood and teenage years, new bone is added faster than old bone is removed. As a result, bones become larger, heavier, and denser. Bone formation continues at a pace faster than resorption until peak bone mass (maximum bone density and strength) is reached around age 30 (Figure 1). After that age, formation tends to slow down which decreases the overall bone mass.

The ratio between formation and resorption is dynamic and depends on the workload. For instance the bone mass of astronauts staying in low gravity for a prolonged period will decrease, while the bone mass of a person training for heavy lifting will increase. Similar mechanisms allow the bone to rapidly repair itself after breaks. This mechanism of continuously forming and removing bone is therefore a powerful mechanism to adapt the bone's strength to its payload and to repair it, while maintaining functionality. When the mechanism is disturbed (e.g. in the case of osteoporosis which affects bone formation), serious problems can arise such as deformation of the spinal column (shortening of the height) and bone collapsing or breaking under normal load[1].

Reinforcement learning

The field of reinforcement learning takes inspiration from the collection of natural phenomena called conditioning. Conditioning is generally held to be the process of learning associations based on experience of interacting with one's environment to get reward (Kimble 1961). It is what makes cats turn up when plates are being scraped, or dogs become excited if they hear their lead being moved. These learning phenomena occur in many situations, in an enormous variety of different animal species.

Conditioning is ultimately driven by reward: the animals learn associations that promote encounters with rewarding situations, such as finding food, or avoid encounters with undesirable situations such as predators or poison.

In the laboratory, conditioning is divided into several different kinds, of which the best known are classical and instrumental conditioning. In classical conditioning, animals learn to associate new sensory stimuli with existing ones. The canonical example is Pavlov's dogs, who learned to associate Pavlov's presence, or the sound of a bell, with the smell of food and responded to his presence or the sound by salivating even when no food was present (Pavlov 1960). Classical conditioning is induced by presenting sensory stimuli in order with a suitable interval between them, with one of the stimuli (the unconditioned stimulus) generating a response from the animal. Thus a dog that hears its lead being fetched shortly before being

---

[1] See http://www.abpi.org.uk/publications/publication_details/targetOsteoporosis/ for more information about bone formation and osteoporosis.

taken for a walk (which it enjoys) will quickly come to associate the sounds of the lead with the walk.

In instrumental conditioning, the animal learns an association between its own actions and their consequences (Mowrer 1956). Unlike classical conditioning, in this case reward is given only when the animal does the desired action. The canonical example of this is when pigeons or rats learn to request food by pressing on a lever. Initially, the pigeon or rat will press the lever in its box by chance, as part of its exploration of the environment it is in. However, having pressed the lever, some food is delivered. The animal learns over time that pressing the lever results in the arrival of food, and will tend to press the lever readily and deliberately when put into the experimental box. Animals can be trained to elicit quite complex sequences of behavior by this technique, building up the sequence a step at a time: the method is widely used to train animals for performance purposes.

The most interesting facets of animal conditioning are apparent universality of the phenomenon – the variety of animal species that exhibit conditioning – and the individuality of it. For instance, rats will readily learn to associate a sound with a painful stimulus (an electric shock, for example) but cannot associate a sound with poison; conversely they readily associate smells with poison but cannot associate smells with electric shocks. Furthermore, certain animals can learn certain associations – particularly fear-driven ones – very quickly: sometimes in a single encounter. Other associations can take much longer to learn. Thus animals are not in general universal learners, able to associate any stimulus with any other stimulus or response; rather they have a finely tuned balance between what is learnable (adaptivity) and what is not (stability) that depends on their ecological niche.

2.2.2 Brief overview of theories and techniques

Control theory

Control theory is a field in engineering which covers all aspects related to the design and analysis of control algorithms for processes and artifacts. In particular control theory aims at designing control laws for determining the behavior of a plant (e.g. a process or a machine) and at analyzing the conditions under which these control laws are “well-behaved”. It is directly relevant to the current discussion, since it provides formal definitions of stability.

First of all, control theory makes a distinction between robustness and stability. When a system is described as a dynamical system, i.e. a set of differential equations, it will have both state variables (quantities which vary) and parameters (quantities which remain constant). A system is then said to be robust when its behavior does not change significantly when parameters of the systems are changed. Stability describes how the state variables evolve, e.g. after a perturbation. Several levels of stability can be formally defined: e.g. Lyapunov stability, asymptotic stability, and exponential stability (Slotine and Li 1991). Roughly speaking, Lyapunov stability describes a system that will remain in the vicinity of the equilibrium state after a perturbation, asymptotic stability describes a system that will eventually return to the equilibrium state after a perturbation, and exponential stability describes a system that will return to the equilibrium state faster than an exponential function. Furthermore these notions of stability can be local, i.e. only valid in a bounded region around the equilibrium state, or global, i.e. valid for the whole state space.

These definitions of stability are in line with the loose one we previously defined. The notion of adaptability is however much less present in traditional control theory, except for the subfield called adaptive control. Astrom & Wittenmark present adaptive control in the following way: “In everyday language, "to adapt" means to change a behavior to conform to new circumstances. Intuitively, an adaptive controller is thus a controller that can modify its behavior in response to changes in the dynamics of the process and the character of the disturbances” (Astrom and Wittenmark 1995). In practice this implies that an adaptive controller is a controller with adjustable parameters, which is tuned on-line according to some mechanism in order to cope with time-variations in process dynamics and changes in the environment. First examples of adaptive control were applied to automatic control of planes. Due to the complexity of the aerodynamics interactions, the control loops require different gains for different flying regimes (different speeds). Different approaches have been developed to tackle this kind of problems, including gain scheduling, model adaptive systems, self-tuning regulators, and dual control (see Sastry and Bodson 1994). While a detailed review is out of the scope of this document, the framework of adaptive control is certainly relevant for future ESA-funded studies on the balance between adaptability and stability.

Homeostasis in neural networks

Biological inspiration in the design of Engineering systems has been a key concept over the past 50 years (Wiener, 1948). One of the earliest electrical systems designed using the concept of homeostasis was the homeostat (Ashby, 1953), an electrical device which, when it was perturbed, searched for the configuration of variables that would return it to its initial condition. More recently homeostatic principles have been used in the design of robot controllers (Di Paolo, 2002). Artificial neural networks (ANN) have traditionally been used to map inputs to outputs via a non linear transformation. This is often done in the context of pattern recognition, where networks are trained to detect the presence of patterns (correlations) in data sets. One commonly used approach to the design and training of neural networks is the use of a multi-layer perceptron (MLP) networks trained through the use of back propagation techniques (e.g. Rumelhart et al., 1986; Hertz et al., 1991). Sample applications which have used this approach are recognition of hand written digits (LeCun et al., 1989) and discrimination of sonar echoes (Gorman & Sejnowski, 1988a, b). In this respect ANNs can be regarded as a relatively mature technology. Recently Neal and Timmis (2005) have proposed a framework for artificial homeostasis that involves the interaction between three distinct components: artificial neural networks (ANN), artificial endocrine system (AES) and artificial immune system (AIS). Each of these component processes is allocated a particular task, and artificial homeostasis is in charge of the interaction between these components to achieve homeostasis within the artifact. It achieves this using positive and negative environmental cues linked to an input-output model of the system. These can be viewed as analogous to a neural reward system. The success of this approach depends on the successful management of the reward system; this is one aim of artificial homeostasis. In this framework an ANN can be developed and trained using well-established techniques (e.g. Back Propagation), and homeostasis is introduced through the interaction of the ANN with other systems. This interaction has the effect of altering the learning algorithm used in the ANN. The biological inspiration for this is the regulatory effects of different hormone concentrations on levels of excitability in neural tissue. Hormone generation and concentration is controlled by separate “gland neurones”, information regarding the concentration of individual hormones is then passed to each neurone in the ANN in turn. Each neurone in the ANN maintains a list of hormone receptors, matching of a hormone by the receptors results in cell specific action to modify the neurone’s behaviour. In the example of Neal and Timmis, the action is to modify individual weights, through a multiplicative action dependent on the product of hormone concentration, sensitivity of connection and match between receptor and hormone. This allows homeostasis to readily interact with the ANN learning rule. An important issue with ANN design is that of generalization, i.e. the ability of a specific neural network to adapt to changes in the input data with training. This is essentially an adaptability-stability issue, and a range of techniques are available to analyse this problem in the context of ANNs (Lippmann, 1987; Hertz et al., 1991). Neal and Timmis (2005) have applied the idea of artificial homeostasis to the control of a mobile robot platform equipped with ultrasound range sensors.

Gene regulatory networks in evolutionary algorithms

Gene regulatory networks GRN are computational models of the interaction of genes and gene products. Underlying these models is the observation that activated genes in organisms cause the production of various molecules (the gene is translated into a protein which may cause the synthesis of other products) and some of the molecules produced are capable of binding to DNA and affecting the transcription of the genes, either facilitating or inhibiting their action. Because GRNs implement dynamical systems, they can balance stability in the form of attractor states with adaptability as environmental conditions, such as externally produced proteins, switch them between different regions of their phase space.

In the computational models, `genes' when active cause the production of `proteins' whose concentrations affect the activity of the producing, or of other, genes. The proteins typically have a production rate dependent on the activation of their gene, and a decay rate which may also be genetically fixed, while the instantaneous gene activation depends on the current local concentration of various proteins. The result is a dynamical system whose state variables are the protein concentrations and the gene activities. Proteins are usually also coupled to particular observable behaviors, for example actions if the GRN is being used as a control system for an agent, or transformations of a body structure if the GRN is being used to control the development of morphology.

Early work on this kind of system was done by Torsten Reil (Torsten Reil, 1999). The same principles, with somewhat varying implementations, have been used by Eggenberger Hotz (2003) for control of multicellular morphology by varying the local geometry of individual cells in a sheet so as to produce 3D effects such as folding, and by Taylor (2004) for control of a swarm of underwater robots.

A key problem with a GRN, whether as a controller or a developmental system, is how to design it: what arrangement of genes and proteins is appropriate for a given system behavior. A way to finesse this problem is to use a suitable evolutionary algorithm, in which collections of genes are rewarded according to the fitness of the behavior they generate (measured by some suitable, typically but not necessarily system-designer-specified, criteria). Collections of genes of high fitness are then allowed to `breed', producing child collections that have characteristics of several parents and hopefully inherit the fitness-producing tendencies of each.

The typical implementation of an evolutionary GRN system uses a standard genetic algorithm in which each individual can be decoded into a GRN. The individual contains a specification for each `gene' that determines which `protein' it makes, the decay rate and binding affinity for that substance, and the set of substances that can interact with the `gene' (its regulators). For each regulator the individual represents the sign and degree of influence that substance has on the activation of the gene. Gene activation may be thresholded or continuously dependent on the regulators.

Evidence suggests that for complex behavioral results, evolutionary algorithms based on GRNs are rather more successful at producing good solutions efficiently than evolution of other representations, arguably because the GRN representation requires fewer evolved parameters to specify a complex behavior.

Reinforcement learning algorithms

Reinforcement learning algorithms are modeled loosely on the processes of natural conditioning described earlier. In precise terms, the goal of reinforcement learning is to construct a relationship between the current sensed state of the world and the action one should take when in that state in order to maximize the received reward over the agent's lifetime. The relationship between state and action is called the agent's policy.

Reward is generally taken to be a scalar value, that is, a number, with positive rewards being desirable and negative ones not. The reward may be given after every action or may be given intermittently when a `good' or `bad' situation is reached. The further apart rewards are, in terms of the number of actions between them, the harder the learning problem is: the learner has to decide which action(s) in a sequence are responsible for the received reward and which are irrelevant or unhelpful. This problem of apportioning credit or blame to individual actions is called the credit assignment problem.

Almost all reinforcement learning algorithms in practice use discounted reward rather than total actual received reward. This is largely for technical reasons: the actual reward collected by an agent with an infinite lifetime is infinite, while the discounted reward is finite and thus better behaved. The discounting of a future reward is achieved by multiplying its value by a discount factor (usually denoted γ) once for each step into the future – that is, a reward that will be received in three steps time is now only worth γ^3 of its value.

The value of γ must lie between 0 and 1, with values such as 0.9 being common. The closer γ is to 1, the further the agent's time horizon extends into the future (that is, the more distant a reward must be for its current worth to be negligible).

There are many reinforcement learning algorithms available, each with its own special properties. The key theoretical result is that it is possible to learn the optimal policy for any reward regime, given only an agent executing actions and receiving reward, provided that all states of the environment can be distinguished and that each possible action is taken in each available state infinitely many times.

Although there are many algorithms, there are two general approaches to reinforcement learning of which the algorithms are particular cases: model-free and model-based. In the former, the agent learns the optimal policy represented by the expected future discounted reward available from each state (the value function), while in the latter the agent also learns a model of its interaction with the environment, in the form of a table of transition probabilities between states and of the reward associated with state-action pairs. The tradeoff between these approaches is one of computation against experience: the model-based approach is generally able to make better use of experience than a model-free approach, at the cost of considerably more computation per experience.

One popular model-free algorithm is Q-learning (Watkins and Dayan 1992). In this technique, the agent represents the value function using a table of state-action pairs (S,A) and for each pair calculates an estimate Q of the discounted future reward it would obtain by taking action A in state S and following the optimal policy thereafter. In operation, an agent in state s takes an action a (one of those possible in state s) and finds itself in state s' and with a reward r (if reward is intermittent, r may be 0 unless the new state is one in which a reward is actually delivered). This is one experience. The agent then computes an estimate of the Q value of the pair (s,a) based on the reward received and the discounted maximum Q value of any pair containing the new state s'. The Q value of the pair (s,a) is then nudged a little toward this estimate. The nudge of the Q-value toward the estimate is controlled by a learning rate parameter usually denoted α. Technical conditions apply to the choice of the α parameter if the algorithm is to converge to the optimal policy.

An example of a model-based algorithm is Dyna (Sutton 1990). In this case, the agent keeps a model of the transition probabilities for each transition (s,a) → s' and of the reward associated with each particular (s,a) pair. These models are updated using each experience. However, the Q-value table is updated multiple times for each experience using the models to predict what transitions and rewards would occur for the updates not directly depending on the current experience. These additional updates can be chosen randomly or based on the current experience and models. An example of a non-random model-based (Dyna-like) method is prioritized sweeping.

A comparison of Q-learning, Dyna and Prioritised Sweeping (PS) can be found on-line at:

http://www-2.cs.cmu.edu/afs/cs/project/jair/pub/volume4/kaelbling96a-html/node29.html

where Leslie Kaelbling, Michael Littman and Andrew Moore demonstrates that for a well-chosen example the model-based methods require between 5% (PS) and 11% (Dyna) of the experiences that Q-learning needs for equivalent performance, while taking 2 (PS) and 6 (Dyna) times as much computational time.

Within any reinforcement learning algorithm, there is an explicit need to balance exploitation and exploration – stability and adaptability. Exploitation uses the knowledge collected so far to choose actions that maximize expected future reward, while exploration chooses actions that allow the agent to collect new information to improve the reliability of the future reward estimates. An algorithm which does no exploration cannot find the optimal policy, since it cannot visit every state and take every action there infinitely many times: it will after some period always choose the same sequence of actions – those that maximize the rewards received according to its partial knowledge. On the other hand, an algorithm cannot solely explore: to do so is to ignore the information on rewards being collected.

Two common techniques for this are □-greedy and Boltzmann-Gibbs action choice. The former chooses uniformly a random action possible in the current state with probability □, otherwise the action with highest Q-value in the current state. In a typical application might be 0.1. The latter computes a probability of choosing each action based on its Q-value and a `temperature' parameter T, such that each action's probability of being chosen is proportional to exp(Q/T) – this results in a Boltzmann-Gibbs distribution of action choices for each state. The higher the `temperature' the more nearly an equal random choice of action. The exploration parameter may be kept constant, or may be reduced over time (annealed) as the system learns.

For further introductory information on reinforcement learning, algorithms, theory and applications, see "Reinforcement Learning: A Survey" by Kaelbling, Littman and Moore at

http://www-2.cs.cmu.edu/afs/cs/project/jair/pub/volume4/kaelbling96a-html/rl-survey.html.

Bone regeneration and modular robotics

The generation of adaptive structures could take inspiration from bone generation. One could imagine a colony of modular robot units that perpetually attach and detach to form a particular structure whose shape and/or properties adapt to the environment by adjusting the ratio of between formation and resorption.

2.3 LOCOMOTION CONTROL

In this chapter, we discuss the balance between adaptability and stability is maintained on an organism level of biological hierarchy of complexity as in animal locomotion.

We present both neural and biomechanical aspects, and describe how these can lead to engineering principles for the design and the control of articulated robots, e.g. rovers to be used in planetary exploration missions. Finally we present a particular preliminary study which explores the feasibility of combining CPG-and-reflex based control with novel types of muscle-like actuators such as Ionic Polymer Metal Composites (IPMC) in an undulatory worm robot.

2.3.1 Problem statement

The control of locomotion represents an interesting framework to study the balance between adaptability and stability in invertebrate and vertebrate animals. During locomotion a good balance between the two has indeed to be maintained at multiple time scales: In the order of seconds, the animal has to produce a stable gait, i.e. a coordinated cyclic movement of all limbs, and maintain its posture while adapting to particularities of the terrain (e.g. with specific feet placement, dealing with obstacles, etc.). In the order of minutes and hours, the animal has to produce stable locomotion while changing gaits, speed and direction, and adapting to changes in body properties (fatigue and load). And in the order of weeks and months, the animal has to adapt to changes in body properties such as growth, injuries, and aging.

Neural control of locomotion

Many animals walk, flight and swim in a stereotyped way, adopting rhythmic patterns of movement, called locomotion patterns (or gaits). In a large variety of animals (from invertebrates to vertebrates (Orlovsky, 1999)) the neural control of these stereotyped movements is hierarchically organized, and has three components: (1) central pattern generators, (2) sensory feedback (reflexes), and (3) descending control signals from higher brain regions (e.g. the motor cortex in vertebrates). The Central Pattern Generator (CPG) is a key functional unit that contains all the mechanisms needed to generate the rhythmic pattern of movements. In vertebrates, it is located in the spinal cord and requires only very simple input signals to initiate and modulate complex and coordinated oscillatory patterns. The CPG essentially provides the feed-forward signals needed for locomotion even in the absence of sensory feedback and high-level control.

In terms of the stability/adaptability dichotomy, CPGs represent the tendency to adopt a stable stereotyped behavior, while their modulation by reflexes, sensory feedback and motor cortex represent the thrust toward adaptability. The overall resulting sensorimotor system is a complex adaptive system with strong bidirectional interactions between its components (CPG, cerebral cortex, sensors, and the musculoskeletal system) (Cohen, 1999). CPGs have the advantage of generating rhythmic movements that are stable, stereotyped, and especially useful for controlling rapid responses (such as flight or run) or involuntary movements (for example, breathing). From a control point of view, they have the nice property of reducing the dimensionality of the control problem, by requiring only low dimensional input signals to modulate multidimensional output signals. On the other hand, sensory feedback provides the capability to adapt the locomotion pattern to the environment (for instance, to walk on uneven terrain with holes, gaps and small obstacles) or fast responses to new situations (reflexes).

The extent and importance of sensory feedback modulation vary from animal to animal and at an extreme have led neurobiologists to formulate totally reflexive paradigms for locomotion control. Two examples of reflexive-based locomotion control are the stick insect (Cruse, 1998) and the nematode *Caernorhabditis elegans* (Niebur, 1993).

Role of the biomechanics

Another important aspect of the stability/adaptability dichotomy in locomotion is connected to the musculoskeletal system. Locomotion is the result of an intricate coupling between the neural dynamics and the body dynamics, and many fundamental aspects of locomotion control including gait transition, control of speed and direction, cannot be fully understood by investigating the locomotor circuit in isolation from the body it controls. A body has indeed its own dynamics and intrinsic frequencies with complex non-linear properties, to which the neural signals must be adapted for efficient locomotion control. As observed by roboticist Marc Raibert, the central nervous system does not control the body, it can only make suggestions (Raibert 1993).

The body is a redundant system with many muscles per joint, and several muscles acting on more than one joint. Muscles serve as actuators, brakes, stiffness regulators, and stores of elastic energy. During locomotion, the frequencies, amplitudes, and phases of the signals sent to the multiples muscles must be well orchestrated. In most vertebrates, complex coordination is required not only between different joints and limbs, but also between antagonist muscles which combine periods of co-activation for modulating the stiffness of the joint, and periods of alternation for actuating the joint.

In legged locomotion, the dynamics of a leg can be approximated by a pendulum model during walking, and a spring-mass model during running. These models allow one to relate several features, such as resonance frequencies, to the length and stiffness of the legs, and are able to describe the mechanics of legged locomotion surprisingly well in many animals (Blickhan 1993). The laws of mechanics are also useful to characterize the stability of gaits in legged animals. Gaits can either be statically stable, when the center of mass is maintained at all times above the polygon formed by the contact points of the limbs with the ground, or dynamically stable, when this rule is not maintained at all times and stability is achieved as a limit cycle which balances the moments, the gravitational forces, and the inertial forces over time. A large variety of gaits can be distinguished depending on the phase relation between limbs, such as the walk, the trot, the pace, and the gallop.

In terms of different balances between stability and adaptability, two extreme examples can be found among tetrapods (four-legged animals). The terrestrial turtle is an example of hyper-stability. Its morphology is such that it has a low center of mass over a large support polygon. The risks of falling over are almost zero, although ending up upside-down is fatal since a turtle on its back cannot return itself. Locomotion is quasi-static and can only be little modulated. The opposite case is the hare. Its locomotion is very adaptable, both in terms of velocity, changes of trajectory, and efficiency in complex terrain. The polygon of support is very reduced. The stability is dynamic, and is based on sophisticated mechanical principles such as self-stability. It appears that biomechanical parameters (such as the kinematics of the center of mass, the angle of attack of the limbs, and the -stiffness in the joints) make the hare's body self-stable. Perturbations are quickly dampened out without the need of neural control. The hare's biomechanics are therefore more sophisticated and more finely tuned than that of the turtle. The mechanical construction is more complex, but is also more adaptable and flexible, since it allows to rapidly and efficiently react to perturbations.

The importance of the mechanical properties of the body is furthermore illustrated by the work on designing passive walkers. Passive walkers are legged machines (some with knees and arms) which transform potential energy from gravity into kinetic energy when walking down a gentle slope. When correctly designed, these machines do not require any actuation or control for generating a walking gait, which in some cases, can be strikingly human-like (McGeer 1990, Collins 2001).

2.3.2 Principles and requirements

Let us consider as an example the exploration of a planetary surface. The main working requirements that a robot has to face with are:

- stability – the robot should be able to transverse uneven terrains and overcome large obstacles preserving its static/dynamic stability;
- robustness – the robot should operate in the harsh space environment;
- maneuverability/dexterity – the robot should maneuver efficiently, and manipulate dexterously (e.g. rock samples);
- energy efficiency – the robot should use as little energy as possible.

The requirements of stability and robustness have often led designers to prefer wheeled or caterpillar-tracked structures over more maneuverable legged robots, which have the potential of reaching areas inaccessible to wheeled structures. Moreover, the importance of stability and adaptability has led many robotic engineers to conceive for planetary exploration wheeled structure with adaptable/reconfigurable chassis (for instance, Mars rovers developed at JPL (Jet Propulsion Laboratory)).

In terms of control, legged robots require more sophisticated control algorithms than wheeled robots due to the additional degrees of freedom of articulated structures. Traditional control approaches in legged robot locomotion fall into two main categories: (1) approaches relying on the play back of pre-recorded trajectories and (2) heuristic control approaches. While a whole range of different techniques have been used, two main techniques respectively emerge in each of these categories: Zero Moment Point (ZMP) control [Vukobratovic 1990] and the Virtual Model (VM) control [Pratt et al 1997]. ZMP control is a trajectory tracking method, which relies on deriving and solving the dynamic equations of motion and using the solution to determine how desired joint angle trajectories affect the zero moment point, also known as the center of pressure. The goal is to ensure stability by keeping the ZMP inside the support polygons of the support feet over time. The desired trajectories can be derived by trial and error, or measured from human recordings. This type of control mechanisms or variants thereof has been used in many biped robots including the Honda robots [Hirai 1998], and the WL and Wabian robots developed at Waseda University [Yamaguchi & Takanishi 1997, Yamaguchi et al 1999]. Such trajectory tracking methods have four primary drawbacks. First, they are computationally intensive. Second, solving the dynamic equations requires a perfect knowledge of the characteristics of the robot and the environment. Third, trajectory-tracking methods require deriving the desired trajectories of the joint angles, which can be a long trial and error process. Finally, trajectory controlled robots are not robust to disturbances or changes in the environment, and require additional control mechanisms for dealing with them.

Among the approaches that use heuristic control algorithms, VM control [Pratt et al 1997] might be one that has had most success. The central principle in VM control is to use virtual elements (e.g. springs and dampers) placed at strategic locations on the robot to control the pitch, height and speed of the robot. The virtual forces applied by the elements are then mapped to physical torques at each of the robots joints (typically by computing the transpose of the jacobian relating the two attachment frames of the virtual element). This allows fast and relatively simple online control. To design a controller, one needs to define the virtual elements (e.g. a few springs to maintain an upright posture, and one to pull the robot forward), and a finite state machine for cycling through the various stages of the gait. This method has been successfully used with the MIT Spring Turkey and Spring Flamingo robots [Pratt & Pratt 1998]. Compared to trajectory tracking methods, VM control has the advantage that it is less sensitive to unexpected external forces due to unknown terrain, since these can to some extent be compensated with the virtual elements. The drawbacks of VM control are that (1) one still needs a perfect characterization of the robot's kinematics in order to compute the torques corresponding to the virtual elements, and (2) controlling the gait with a finite state machine is a rather rigid way of generating a cycle, which can not easily deal with problems such as contacts with obstacles that require fast and smooth modulations of the limb trajectories.

In this case study, we recommend to explore a new approach to locomotion control that takes inspiration from vertebrate locomotion control in order to develop CPG-and-reflex based control algorithms [Taga 1998, Kimura et al 1999, Ijspeert 2001, Arena 2002]. As discussed above, locomotion in many animals is controlled by the interaction of three components: (1) central pattern generators, (2) sensory feedback, and (3) descending supraspinal control. Control is organized such that the CPG generates the basic rhythmic patterns necessary for locomotion, and that higher control centers and sensory feedback modulate the CPG's activity for dealing with the environment and the behavioral requirements. The purpose of using CPG-and-reflex based control algorithms for legged robots is to obtain stable locomotion as the result from the interaction between the controller, the body, and the environment (such as to take the embodiment of the controller into account, as discussed in [Pfeifer and Scheier 1999]). From a dynamical systems point of view, locomotion becomes the limit cycle behavior of the controller-body-environment system. Small perturbations to the system are quickly forgotten and will not destroy the cyclic movements as long they remain within the basin of attraction of this limit cycle. Whenever strong external perturbations arise, actions of reflexes and corrections from higher control centers (e.g. a corrective step due to a side-push) are then meant to bring back the state of the system into the basin of attraction.

2.4 LOCOMOTION CONTROL - FEASIBILITY OF ENGINEERING SOLUTIONS

In order to create robots that can satisfactorily solve the balance of adaptability and stability and that meet the challenges of next generation space rovers, we propose to design robots that take inspiration from biology in terms of both control and structure. In particular, we propose to explore the feasibility of combining CPG-and-reflex based control with novel types of muscle-like actuators such as Ionic Polymer Metal Composites (IPMC).

To illustrate such an approach, we shall describe a first feasibility study of undulatory locomotion in a worm-like robot. Undulatory locomotion has several advantages satisfying several of the requirements analyzed above: stability (intrinsically satisfied), manoeuvrability on uneven terrain, efficiency, redundancy, and compactness. In fact, worms and snakes are ubiquitous on the Earth: they are able to move on uneven terrain in hostile environments, to overcome obstacles and to reach impervious places. In order to reach adaptability and stability, innovative techniques regarding both the control strategy and its integration in new smart materials must be faced. Therefore, in this work we propose a case study in which a high adaptive locomotion structure joint with a stereotyped central pattern generator is used. We take into account a worm-like robot as an example of trade-off between stability and adaptability: adaptivity is provided by undulatory locomotion, which is controlled by a CPG; moveover, for the considerations discussed above, the mechanical structure conjugates stability and adaptability issues. Some preliminary notes on undulatory locomotion are first discussed in Paragraph 2.4.1, while the idea underlying CNN-based CPG is introduced in Paragraph 2.4.2. In order to evaluate the performance of the robotic structure on Mars surface a mathematical model is introduced in Paragraph 2.4.3, the model is validated by using experimental data in Paragraph 2.4.4, the behavior of the model with respect to different gravitational fields is then studied in Paragraph 2.4.5, finally the suitability of the approach is discussed in Paragraph 2.4.6.

2.4.1 Notes On Undulatory Locomotion

Undulatory locomotion is defined as the process of generating net displacements of a robotic structure via a coupling of internal deformations to a continuous interaction between the robot and its environment (Ostrowski, 1998). This kind of locomotion is used by very different animal species (worms, snakes, fishes) in a huge variety of environments, because of the high adaptability to terrains, often inaccessible by other kinds of motion. This feature makes undulatory motion appealing also in robotics. Several kinds of undulatory locomotion are adopted in nature: for example, snakes use bending waves associated with the asymmetric friction of their scales to crawl on the plane (in this case the bending wave propagates from the head to the tail), whereas inchworms move in and out of the plane (in this case the bending wave travels from the tail to the head) (Alexander, 2003). Other worms (nematodes like the C. Elegans (Niebur, 1991)) have locomotion patterns resembling the typical motion of a snake, although the mechanisms used to generate the thrust in snakes and nematodes are very different. In undulatory locomotion a symmetric structure is actuated by periodic muscular pulses. Motion in one direction is made possible by one or more symmetry-breaking mechanisms (Mahadevan, 2004). Typically, these mechanisms are unidirectional waves and the inequality of static or dynamic friction.

Design case study

We take into account a worm model in which the symmetry-breaking mechanism is the inequality of dynamic friction. We develop the equations of motion of this model and investigate the suitability of the structure for planetary exploration.

In order to develop the model, we take as reference biological case the Lumbricus terrestris for which a huge set of experimental data has been collected (Quillin, 1998; Quillin, 2000). Moreover, we take as reference robotic structure a worm-like robot based on Ionic Polymer Metal Composites (IPMC) (Arena, 2002).

The actuation mechanism of the IPMC worm is based on a muscular wave propagating from the tail to the head of the body. There are several advantages of using IPMC:

- IPMC (Bar-Cohen, 2001; Bar-Cohen, 2000; Shahinpoor, 2005) are materials suitable for sensing and actuation integration. They are able to behave both as sensors and as actuators with good performance in terms of stress, strain, response time, reliability and life time. These polymers bend under the effect of a low electric field acting as motion actuators; moreover, when bent, a voltage is produced across the thickness of the strip between the two conducting electrodes attached; thus they act also as sensors;

- o IPMC are able to work under severe environmental conditions. They have been demonstrated able to work at low temperature (-80° ÷ -140°) and low pressure (in the order of few Torrs);
- o IPMC provide very low weight, low energy needs and efficiency arising from sensor/actuator integration.

The reference IPMC worm-like prototype is discussed in detail in (Arena, 2002). The structure of this robot is made of an IPMC strip, functionally divided in four identical segments. The whole strip length is L=0.1m, the total weight (including wiring contacts) is 1·10-3kg. The worm is actuated by using a CNN-based CPG (Arena, 2002; Arena, 2004). The CPG generates a pattern of four signals actuating the four IPMC strips, so that an activation wave propagates from the tail of the robot to the head, realizing locomotion. A schematic view of the IPMC worm is shown in Figure 2a while a schematic representation of the locomotion mechanism for a three-segment structure is shown in Figure 2b.

2.4.2 Cnn-Based Central Pattern Generators

In this Section the idea underlying CNNs for locomotion control is described. The key point is to use dynamical systems with simple nonlinearity (as CNNs are) and local connections to build an artificial CPG. The theoretical aspects underlying the CNN approach for implementing artificial CPGs were dealt with in more details in (Arena, 2002). The CPG of the artificial locomotion system is implemented by a network of coupled nonlinear oscillators, where each oscillator plays the role of a motor-neuron for each actuator of the robot. In fact, the periodic behavior of the nonlinear oscillator provides the rhythmic movements needed for a proper locomotion, for instance, of a segment of a worm-like robot. Focusing on this example, in our approach a robot with n actuating segments requires a CPG with n six motor-neurons (one for each segment). Connections between the motor-neurons determine the phase lags between the corresponding segment and thus the locomotion pattern of the robot. For instance, to implement undulatory locomotion, a constant phase lag (equal to $2\pi/n$) between motor-neurons controlling the adjacent segments is required.

Each nonlinear oscillator is a second-order CNN circuit with the following dimensionless equations:

$$\begin{aligned} \dot{x}_{1;h} &= -x_{1;h} + (1+\mu)y_{1;h} - sy_{2;h} + i_{1;h} + I_{syn;h} \\ \dot{x}_{2;h} &= -x_{2;h} + (1+\mu)y_{2;h} - sy_{1;h} + i_{2;h} \end{aligned} \tag{1}$$

where

$$y_{j;h} = 0.5(|x_{j;h} + 1| - |x_{j;h} - 1|)$$

with j={1,2} is the nonlinear output function; h=1..n indicates the segment number; x1 and x2 are the state variables of the motor-neuron; and $I_{syn;h}$ is the synaptic input given by the connections with the other motor-neurons.

The parameters for which a stable limit cycle is obtained are given in Table 1. The movements of the segment associated with the motor-neuron are controlled by one of the state variables x_1 and x_2.

The motor-neurons are connected by choosing synapses between them in a way that intrinsically fixes the locomotion pattern. Like in the biological case, these connections can be either excitatory or inhibitory; their choice can be accomplished by following the guidelines discussed in (Arena, 2002). As an example an excitatory synapse (from neuron k to neuron h) can be set by choosing either $I_{syn;h}=y_{1;k}$ (not-delayed) or $I_{syn;h}=y_{2;k}$ (delayed) in Equation (1). Another approach to fix the connections between motor-neurons in CNN-based CPGs is based on reaction-diffusion (RD) equations (Arena, 1999). Neurons can be connected in a ring-like structure to generate an autowave propagating in the ring. In this case, the equations of the motor-neuron can be rewritten as follows:

$$\begin{aligned} \dot{x}_{1;h} &= -x_{1;h} + (1+\mu)y_{1;h} - sy_{2;h} + i_{1;h} + D_1(y_{1;h+1} + y_{1;h-1} - 2y_{1;h}) \\ \dot{x}_{2;h} &= -x_{2;h} + (1+\mu)y_{2;h} - sy_{1;h} + i_{2;h} + D_2(y_{2;h+1} + y_{2;h-1} - 2y_{2;h}) \end{aligned} \tag{2}$$

where D_1 and D_2 represent the diffusion coefficients. Since equations (2) give rise to an autowave propagating in the ring, they are suitable to control undulatory locomotion, where a muscular wave propagates from the tail to the head of the worm-like robot. In fact, the autowave generated by the RD-CNN CPG directly elicit a "muscular" wave in the actuators of the structure.

2.4.3 Mathematical Model Of Lumbricus Terrestris/Ipmc Worm

Physically-based models are usually taken into account to model non rigid materials. These models can be continuous or discrete. Continuous methods are based on finite element analysis. Discrete methods consider the non rigid body as a set of points of mass connected by springs.

As concerns undulatory locomotion, Keller and Falkovitz (1983) proposed a continuous analytical model for the crawling of worms, taking into account friction and gravity; Dario et al. (2004) developed a simple continuous analytical model, based on the experimental data reported by Quillin (1998, 1999, 2000). An example of discrete models is the model of C. elegans introduced in (Niebur, 1991), which concerns locomotion undulatory on the plane.

In order to develop a discrete model of earthworm locomotion, a physically force-based model has been adopted.

Structure

Since the motion of the worm develops on the x-y plane, the body of the worm is a two-dimensional layer of points of mass, divided in N rectangular segments. The vertices of each segment are the points of mass, they are connected each other by springs and dampers (Figure 3).

Equations of motion

In this Section the equations of motion of the structure are briefly derived (Pavone, Thesis at University of Catania). The motion direction is the x axis. Each point of mass P at time t is subjected to elastic (longitudinal and lateral) forces, damping and frictional forces.

The elastic force between two points of mass $\mathbf{P}_j$ and $\mathbf{P}_k$ is given by the Hook's law:

$$\begin{aligned} \mathbf{f}_{jk}^{(e)} &= k_0 \left(\frac{L_{jk}}{L_{jk}^0} - 1 \right) \frac{\mathbf{P}_k - \mathbf{P}_j}{L_{jk}} \\ \mathbf{f}_{kj} &= -\mathbf{f}_{jk}, \quad L_{jk} = \left\| \mathbf{P}_k - \mathbf{P}_j \right\| \end{aligned} \tag{3}$$

where k_0 is the (longitudinal or lateral) spring stiffness, L_{jk} is the distance between two points $\mathbf{P}_j$ and $\mathbf{P}_k$, while L_{jk}^0 is the spring equilibrium length.

The damping between the two points $\mathbf{P}_j$ and $\mathbf{P}_k$ is given by the following expression:

$$\mathbf{f}_{jk}^{(d)} = D \frac{dL_{jk}}{dt} \frac{\mathbf{P}_k - \mathbf{P}_j}{L_{jk}} \tag{4}$$

The frictional force for the point of mass $\mathbf{P}_j$ is:

$$\mathbf{f}_j^{(f)} = -\varphi(\dot{\mathbf{P}}_j) \frac{\dot{\mathbf{P}}_j}{\left\| \dot{\mathbf{P}}_j \right\|} mg \tag{5}$$

where φ is defined as follows:

$$\varphi(\dot{\mathbf{P}}_j) = \begin{cases} \mu_{forward} & \text{if } \dot{P}_{jx} \geq 0 \\ \mu_{backward} & \text{otherwise} \end{cases} \tag{6}$$

where $\dot{P}_{jx} \geq 0$ is the component along the motion direction of the velocity of point $\mathbf{P}_j$.

Each given point $\mathbf{P}_j$ at time t is therefore subjected to this total force:

$$\mathbf{f}_j^{(tot)} = \sum_{k \in nr} k_k \mathbf{f}_{jk}^{(e)} + \sum_{k \in nr} D\mathbf{f}_{jk}^{(d)} + \mathbf{f}_j^{(f)} \tag{7}$$

where $j=1,2,\ldots,\ 2N+2$, nr is the set of neighboring points of mass and k_k is the appropriate elastic stiffness (longitudinal or lateral).

For each point of mass the Newton's second law can be applied to obtain the equations of motion. The whole dynamic motion equations are given by $2N+2$ second-order nonlinear differential equations as follows:

$$\begin{cases} \ddot{\mathbf{P}}_1 = \frac{1}{m}\mathbf{f}_1^{(tot)} \\ \ddot{\mathbf{P}}_2 = \frac{1}{m}\mathbf{f}_2^{(tot)} \\ \ldots \\ \ddot{\mathbf{P}}_{2N+2} = \frac{1}{m}\mathbf{f}_{2N+2}^{(tot)} \end{cases} \tag{8}$$

Actuation

The worm is actuated by acting on the equilibrium length of both longitudinal and lateral springs and letting the equilibrium length be a function of space and time, i.e. $L^0=L^0(x,t)$. Let us consider the longitudinal spring of the j-th segment, in this case x is the coordinate of the point A_j and a muscular excitation wave is applied as follows:

$$L_{long}^0(x_j, t) = L_{long}^0 + \psi(x_j, t) \tag{9}$$

where L_{long}^0 is a constant and $\psi(x_j, t)$ is a wave given by:

$$\psi(x\ ,t) = A\sin(2\pi\frac{x}{\lambda} + 2\pi\frac{t}{T}) \tag{10}$$

where λ is the wave number, T the wave period and A the amplitude of the activation wave.

As concerns lateral springs, they are actuated so that the area of each body segment is constant:

$$L_{lat}^{0}(x_j, t) = \frac{L_{long}^{0} L_{lat}^{0}}{\frac{L_{long}^{0}(x_j, t) + L_{long}^{0}(x_{j-1}, t)}{2}} \tag{11}$$

Parameters

The parameters of the model have been chosen taking into account experimental data on the *Lumbricus terrestris*. Moreover, several parameters have been chosen taking into account the characteristics of the robot and physically plausible values of them.

Structural parameters. The number of segments in the *Lumbricus terrestris* is $N\approx 150$. However, since a small number of segments is often used in robotic realizations, $N=4$ has been considered.

The mass of a *Lumbricus terrestris* varies in the range $0.01 \div 8 \cdot 10^{-3} kg$ (Quillin, 1998). The IPMC worm weights about $1 \cdot 10^{-3} kg$. So, it has been assumed a total mass $m=1 \cdot 10^{-3} kg$, which leads to a mass for each point equal to:

$$m_p = \frac{m}{2N+2} = 0.1 \cdot 10^{-3} kg \tag{11b}$$

The ontogenetic scaling for the length of the *Lumbricus terrestris* reported in (Quillin, 1998) is given by:

$$L = 102m^{0.34} \tag{12}$$

where the mass is expressed in grams. In our case, equation (12) gives $L = 102 \cdot 10^{-3}\, m$.

Therefore, each segment length is given by $L^0_{long} = L/N = 20 \cdot 10^{-3}\, m$. Analogously, the radius scales as $r = 5m^{0.34}$, an thus it has been chosen $L^0_{lat} = 5 \cdot 10^{-3}\, m$.

As concerns spring stiffness and damping coefficient, the following physically coherent values have been chosen $k=0.5N/m$ and $D=0.2Ns/m$.

Table 2 summarizes the structural parameters.

Environment parameters. For the acceleration due to gravity it has been assumed $g=9.81m/s^2$. A key point of the model is that forward and backward frictional coefficients are not equal. The same symmetry-breaking mechanism is for instance adopted by snakes. Frictional coefficients have been chosen as follows: $\mu_{forward}=0.2$ and $\mu_{backward}=5\mu_{forward}$. The coefficient $\mu_{forward}$ is very close to those of snake scales as reported in (Dowling, 1997) ($\mu_{forward}=0.3$). The results obtained with our model are quite independent of the ratio $\mu_{backward}/\mu_{forward}$.

Table 3 summarizes the structural parameters.

Actuation parameters. The actuation parameters, i.e. λ, T and A in equation (10), have been chosen taking into account the following considerations. Since the undulatory locomotion of *Lumbricus terrestris* is characterized by a wave along the body length, it has been chosen $\lambda=L$. The other two parameters have been chosen to match experimental data. Moreover, as concerns the parameter A, there is no direct correspondence with the biological case. It has been chosen $T=1s$ and $A=\frac{L_{long}^{0}}{20}$.

The model has been integrated by using the Euler method with a fixed step size Δt=10-4s. However, the results have been compared with those obtained with a method for moderately stiff problems (the Bogacki and Shampine method), which validates the approach.

Table 4 summarizes the actuation parameters.

2.4.4 Validation of the model

The model has been validated by comparing the velocity obtained in simulation with the velocity experimentally measured in *Lumbricus terrestris* (Quillin, 1999). Then, it has been valuated how the velocity scales with the mass either in the model or in experimental data.

The average speed of the model has been evaluated by taking into account a simulation of *30s* corresponding to 30 actuation periods. An average speed of $v=4.6\cdot 10^{-3}m/s$ comparable with the experimental data ($v=3.8\cdot 10^{-3}m/s$ (Quillin, 1999)) has been obtained.

The scaling law reported in (Quillin, 1999) gives the following expression for the velocity:

$$v = 3.8m^{0.33} \tag{13}$$

with the mass expressed in grams and the velocity in *mm/s*. Equation (13) has been compared with simulation results obtained with the model introduced above. The results are shown in Figure 4 (a) and (b) is a further confirm of the good matching between simulated behavior and experimental data: the normalized value of the speed with respect to the body length weakly depends on the mass and has values in the range $0.04\pm 0.02s^{-1}$ (the experimental value reported in (Quillin, 1999) is $0.04\pm 0.01s^{-1}$).

2.4.5 Study of the velocity with respect to the gravitational field

Once demonstrated the validity of the model, the behavior of a worm-like structure on different gravitational fields has been inferred in order to investigate the suitability of these structures for planetary exploration. It should be preliminarily noticed that the effect of the gravity is twofold and thus difficult to predict: on one hand, the frictional forces associated with the forward direction decrease, thus favoring an increase of the motion speed; on the other hand, the frictional forces associated with the backward direction decrease, thus weakening the anchoring mechanism of the worm.

The behavior of the model has been evaluated by carrying out simulations of the average speed at different values of g. The results are shown in Figure 5, where the average speed versus log g is reported. It is interesting to note that the curve shows a peak (according to the need of a trade-off between opposite effects of the gravitational field): surprisingly this peak is located near the values corresponding to the gravitational field of Mars ($g_{Mars} \approx g/3$).

The average speed at values corresponding to gravitational fields of Earth and Moon (it has been assumed $g_{moon} \approx g/3$) are also shown. Moreover, as it can be noticed the performance at g (Earth) are close to the peak, an opposite result would be counterintuitive.

2.4.6 Use of IPMC actuators under mars conditions

Besides the frictional forces examined in Section VI the use of IPMC leads to another positive effect. In fact, a greater bending results from a reduced gravitational field. To quantify this effect, an approximate model of the deformation of the IPMC actuators is introduced.

The IPMC actuator has been modelled as a two-link system with the points A and C constrained to the ground as shown in Figure 6. The forces to be considered are the gravitational force acting on the center of the structure and the deflecting force exerted by the IPMC and applied to the points A and C, orthogonally to the link itself. A further assumption is that the this force decreases linearly with the deformation.

With these assumptions, the following relationship for the parameter *A* can be derived:

$$A = A_0 \cdot \chi(g) \tag{14}$$

where A_0 refers to the value of the parameter *A* in equation (10) on Earth and $\chi(g) = c\frac{g}{g_{Earth}}$ with $c=-0.077$ and $d=1.077$.

By simulating the model and taking into account the effect of the gravity both on frictional forces and on the amplitude of the activation wave as in equation (14), it has been found that the velocity on Mars would be $v_{Mars} \approx 1.2 \cdot v_{Earth}$.

2.4.7 Conclusion

Adaptability and stability are important issues in locomotion, where, on one hand, static and dynamic stability and fast generation of stereotyped locomotion patterns should be assured, and on the other hand flexible and adaptive mechanical and control architectures are desirable. In this report an original strategy to artificially obtain adaptability is outlined. It derives from using the CNN paradigm for motion control and active polymers.

To fulfill these requirements, a design case study taking into account undulatory locomotion realized by a wormlike structure has been developed that is controlled by a CNN-based CPG. Its performance have been investigated under different conditions (gravity, friction, ...) through detailed mathematical models. The use of innovative smart materials based on IPMC has been successfully proposed in order to conceive an efficient locomotion system based on bio-inspired principles where the CNN motion control has been adopted.

This investigation highlighted the advantages of using such a structure under Mars conditions (gravity, friction, environmental conditions) and the benefits arising from a reduced gravitational field. The results obtained showed that under Martian gravitational field undulatory locomotion is still efficient.

2.5 IMMUNE SYSTEM

In this chapter, we discuss the balance between adaptability and stability is maintained on an sub-organism level of biological hierarchy of complexity as in immune system.

2.5.1 Problem statement

The immune system is a remarkable, and complex, natural defense mechanism, which responds to foreign invaders called pathogens. Organisms typically have two lines of immunity, innate (inherited at birth) and adaptive (also known as acquired) which develops over the lifetime of the organism. However this is not the case for all organisms, such as the shark, which has a very powerful innate immune system and no acquired immune system. The innate immune system has first contact with any pathogenic substance and in a large amount of cases, this is all that is needed to remove the pathogenic material from the organism. However, there are many cases where the innate immune system is insufficient and cannot remove the infection. If this is the case, then the pathogen is passed over to the adaptive immune system. The immune system demonstrates constantly the notions of stability and adaptability. With regards to adaptability, the function of the immune system will change in response to the environment, but will return to a stead-state after such events (stability). These properties have been exploited by the area of Artificial Immune Systems (AIS), where immunological properties are captured in computational solutions. Applications range from hardware fault tolerance (Canham and Tyrrell, 2002), email classification (Secker, Freitas and Timmis, 2003) to network security (Forrest, Hofmeyr and Somayaji, 1997).

The adaptive immune system primarily consists of B- and T-lymphocytes (cells). Through receptors on the cell, they are capable of binding with pathogenic material (antigens). Binding will occur between the receptors (paratopes) and antigen receptors (epitopes) if the affinity between the two is above a certain threshold. If a T-cell successfully binds an antigen this will cause the T-cell to stimulate B-cells through the emission of lymphockines. Additionally, B-cells can also bind with antigens, and therefore a notion of antigenic affinity is created. The B-cells receive stimulation from this interaction with the antigen. Through the combination of these two interactions (antigens and T-cells) a B-cell then becomes stimulated and reaches a threshold at which it transforms into a *blast cell*. These blast cells then produce large amounts of clones (in proportion to antigenic affinity: the higher the affinity, the larger the number of clones produced) and also a large number of free antibodies, which undergo somatic hypermutation to increase the diversity of the immune response. This whole process is known as affinity maturation and is part of the *clonal selection theory* (Burnet 1959), which is the term used to identify the process described above. These antibodies (with the assistance of killer T-cells) will remove the antigen from the system. The immune system maintains an *immune memory* of cells, so that when exposed to the same (or slightly different) antigen, a quicker secondary response can be elicited which results in quicker removal of the infection.

The immune system remembers encounters with antigenic material (Tizzard 1988). There are a number of theories on how the immune system remembers encounters with antigenic material, with the most favored view being that of clonal selection and memory cells (Burnet 1959). However, a theory first proposed in (Jerne 1974) suggested an Idiotypic network and the immune network theory. Although not widely accepted, this theory is interesting especially for computer scientists and is the model that will be discussed in more detail. The Idiotypic network was devised to explain the stimulation of B-cells in the absence of antigens. This is achieved by stimulation and suppression between cells via a network communicating via idiotypes on paratopes. The network acts as a self-organising and self-regulatory mechanism that captures antigenic information. Notable work in (Farmer, Packard et al. 1986) further explored the immune network theory and created a simple model of the Idiotypic network, which was further extended by (Perelson 1989). It can be noted that such a self-regulated system is akin to a homeostatic system, i.e. is capable of maintaining its own internal steady state.

Immune Responses

A primary response (Tizard, 1988a) is provoked when the immune system encounters an antigen for the first time. A number of antibodies will be produced by the immune system in response to the infection, which will help to eliminate the antigen from the body. However, after a period of days the levels of antibody begin to degrade, until the time when the antigen is encountered again. This secondary immune response is said to be specific to the antigen that first initiated the immune response and causes a very rapid growth in the quantity of B-cells and antibodies. This second, faster response is attributed to memory cells remaining in the immune system, so that when the antigen, or similar antigen, is encountered, a new immunity does not need to be developed. This is the notion of adaptability within the immune system. Figure 7 illustrates this process.

With this in mind, the immune system can be seen as a dynamic system, with many emergent properties, such as the protection of our bodies, the maintenance of the host, stable (steady) state operation and so on. The immune system is well renowned for protecting the body from infections, as outlined above. However, the immune system has an equally important role: one of maintenance. This maintenance involves (1) making cells grow and replicate; (2) making cells die; (3) making cells move; (3) influencing cell differentiation and (5) modifying tissue support and supply systems (e.g. building connective tissue, regulating blood vessel growth and supply). These processes may result from direct activation by various immune system cytokines. What is also interesting, is that these functions are present in embryological development of the body, but continue in a limited fashion controlled by the immune system. These activities are initiated through the turning on and off of genes (via the cytokines). It is common to distinguish between two principal types of genes: housekeeping genes and tissue-specific genes. Housekeeping genes are active in all cells at all times as the products of housekeeping genes are needed for the ongoing metabolism that all cells require. Tissue-specific genes are only expressed in certain cell types, when they are needed to carry out the specialised functions of the cell: hormone genes in endocrine cells, reproductive genes in germ cells etc. However, there is a third type, maintenance genes. These genes are different to housekeeping genes in that they are needed in times of crisis, as well as in an ongoing fashion. In order to perform maintenance effectively, the immune system has a program of three parts:

- recognition: identifying when things are right and wrong;
- cognition: interpreting the signs, evaluating the results and making decisions;
- action: doing the job.

Immune Memory

It is possible to identify two main philosophical avenues that try to explain how immune memory is acquired and maintained (Tew & Mandel, 1979), (Tew et al, 1980), (Ada & Nossal, 1987) and (Matzinger, 1994): 1) clonal expansion and selection, and 2) immune network. Throughout the lifetime of an individual, it is expected to encounter a given antigen repeatedly. The initial exposure to an antigen that stimulates an adaptive immune response is handled by a spectrum of small clones of B-cells, each producing antibodies of different affinity. The effectiveness of the immune response to secondary encounters is considerably enhanced by storing some high affinity antibody producing cells from the first infection, named memory cells, so as to form a large initial clone for subsequent encounters. Thus memory, in the context of secondary immune responses, is a clonal property (Coutinho,1989). Another theory is the theory first proposed by Jerne (Jerne, 1974) and reviewed in (Perelson, 1989) called the Immune Network Theory. This theory postulates that B-cells co-stimulate each other via portions of their receptor molecules (idiotopes) in such a way as to mimic antigens. An idiotope is made up of amino acids within the variable region of an antibody or T-cell. A network of B-cells is thus formed and highly stimulated B-cells survive and less stimulated B-cells are removed from the system. Both theories allow for the explanation of a meta-dynamical system that is capable of stability, adaptability and reconfigurability.

Interactions Between Immune, Neural and Endocrine Immune, neural and endocrine cells express receptors for each other

This allows interaction and communication between cells and molecules in each direction. It appears that products from immune and neural systems can exist in lymphoid, endocrine and neural tissue at the same time. This indicates that there is a bi-directional link between the nervous system and immune system. Therefore, it would seem that both endocrine and neural systems can affect the immune system. There is evidence to suggest that by stimulating areas of the brain it is possible to affect certain immune responses, and also that stress (which is regulated by the endocrine system) can suppress immune responses: this is also reciprocal in that immune cells can affect endocrine and neural systems. The action of various endocrine products on the neural system is accepted to be an important stimulus of a wide variety of behaviors. These range from behaviors such as flight and sexual activity to sleeping and eating.

The primary function of the immune system is to defend the body against foreign invaders and malfunctioning cells. There are a wide variety of components that are used to achieve this, ranging from the bone marrow to lymph nodes. The immune system displays a number of interactions with other biological systems including the following: immune cell populations have receptor profiles for modulators such as neurotransmitters and endocrine hormones; and immune products also exist in neuroendocrine tissues.

The nervous system's functions are the reception of stimuli, with the transmission of nerve impulses and activation of muscle (or effector) mechanisms. The nervous system has a number of interactions, which can be summarised as follows. Neural cells express receptors for cytokines, hormones and neuro-transmitters. The brain can stimulate defense mechanisms against infection, thus engaging the immune system. The hypothalamus within the brain, controls the pituitary and other endocrine glands and it is known that neural products coexist in immune and endocrine tissues.

Finally, the endocrine system's function is to secrete hormones into the blood and other body fluids, with the aim being to regulate metabolism, growth etc. There are a large number of components that make up the system including glands such as the thyroid, pineal and the thymus. These glands are closely related to three fundamental activities which are of interest: growth, release of hormones to the brain, and immune system development. There are a number of interactions that the endocrine system is involved with: endocrine cells express receptors for cytokines, hormones, and neuro-transmitters; hormones provide feedback to the brain that affect neural processing; hormones including the reproductive hormones also affect the development of the nervous system. Again, endocrine products also exist in both immune and nervous tissue.

In terms of adaptability the immune system must constantly keep abreast not only of infectious pathogens, but also internal stressor signals associated with tissue stress and the like. In response to these signals, the immune system produces a diverse range of antibodies (through mutation) as to endow the immune system with the ability to adapt to the situation. Of course, this is not a static situation, and is dynamic and noisy. The challenging nature of how and where the immune system operates means that a careful balance needs to be stuck between the adaptive nature of the immune system, but also maintaining a stable state. The interactions between immune cells gives rise to immune homeostasis, (thus a notion of stability) but the interactions between the three system immune, neural and endocrine also give riser to homeostatic at the host level.

2.5.2 Principles and requirements

Several general principles can be extracted from the immune system and applied to creating a system capable of stable, yet adaptive, behavior (Bersini, 2002):

- Principle 1: The control of any process is distributed around many operators in a network structure. This allows for the development of a self-organising system that can display emerging properties;

- Principle 2: The controller should maintain the viability of the process being controlled. This is keeping the system within certain limits and preventing the system from being driven in one particular way;
- Principle 3: While there may be perturbations that can affect the process, the controller learns to maintain the viability of the process through adaptation. This learning and adaptation requires two kinds of plasticity: a parametric plasticity, which keeps a constant population of operators in the process, but modifies parameters associated with them; and a structural plasticity which is based on the recruitment mechanism which can modify the current population of operators;
- Principle 4: The learning and adaptation are achieved by using a reinforcement mechanism between operators. Operators interact to support common operations or controls;
- Principle 5: The dynamics and metadynamics of the system can be affected by the sensitivity of the population;
- Principle 6: The system retains a population-based memory, which can maintain a stable level in a changing environment.

2.6 IMMUNE STSTEM - FEASIBILITY OF ENGINEERING TECHNIQUES

In an attempt to create a common basis for AIS, work in (de Castro and Timmis, 2002) proposed the idea of a framework for AIS. The authors argued the case for proposing such as framework from the standpoint that in the case of other biologically inspired approaches, such as artificial neural networks (ANN) and evolutionary algorithms (EAs) such a basic idea exists and helps considerably with the understanding and construction of such systems. For example, (de Castro and Timmis, 2002) consider a set of artificial neurons, which can be arranged together so as to form an artificial neural network. In order to acquire knowledge, these neural networks undergo an adaptive process, known as learning or training, which alters (some of) the parameters within the network. Therefore, the authors argued that in a simplified form, a framework to design an ANN is composed of a set of artificial neurons, a pattern of interconnection for these neurons, and a learning algorithm. Similarly, the authors argued that in evolutionary algorithms, there is a set of “artificial chromosomes” representing a population of individuals that iteratively suffer a process of reproduction, genetic variation, and selection. As a result of this process, a population of evolved artificial individuals arises. A framework, in this case, would correspond to the genetic representation of the individuals of the population, plus the procedures for reproduction, genetic variation, and selection. Therefore, the authors adopted the viewpoint that a framework to design a biologically inspired algorithm requires, at least, the following basic elements:

- a representation for the components of the system (known as shape space);
- a set of mechanisms to evaluate the interaction of individuals with the environment and each other. The environment is usually simulated by a set of input stimuli, one or more fitness function(s), or other mean(s) and;
- procedures of adaptation that govern the dynamics of the system, i.e., how its behavior varies over time. Adopting this approach, (de Castro and Timmis, 2002) proposed such a framework for AIS. The basis of the proposed framework for is therefore a representation to create abstract models of immune organs, cells, and molecules, a set of functions, termed affinity functions, to quantify the interactions of these "artificial elements", and a set of general-purpose algorithms to govern the dynamics of the AIS.

The framework can be thought of as a layered approach as shown in Figure 8. In order to build a system, one typically requires an application domain or target function. From this basis, the way in which the components of the system will be represented will be considered. For example, the representation of network traffic may well be different that the representation of a real time embedded system. Once the representation has been chosen, one or more affinity measures are used to quantify the interactions of the elements of the system. There are many possible affinity measures (which are partially dependent upon the representation adopted), such as Hamming and Euclidean distances. The final layer involves the use of algorithms, which govern the behavior (dynamics) of the system. Here, in the original framework proposal, algorithms based on the following immune processes were presented: negative and positive selection, clonal selection, bone marrow, and immune network algorithms.

When constructing such an AIS, there are many computational and practical issues to consider. The first is computational complexity of the approach. This relates to the time and space required to generate the suitable number of detectors (members of a population) that are required for the job. For example, there are a number of works that outline the unacceptable computational complexity of the negative selection approach from AIS, as there is an exponential relationship between the size of the data set to be used, and the number of detectors that it is possible to generate. However, other approach within AIS, such as clonal selection and immune networks, do not suffer the same problem. The second aspect to consider is the data to be used. If one abstracts away from the system components and uses state machines, then one has to be careful that there is an accurate mapping between the state machine and the actual system, and ensure that the state machine adequately scopes the space to be immunised. Consideration here also has to be given to the way in which data is represented. The shape space paradigm proposes varying ways of data representation and interaction. However, when dealing with discrete values, such as those found in embedded systems, the method of defining affinity (i.e. seeing how similar one item is to another) is not as clear-cut as it may seem. This is coupled with the fact that mutation, even what might be thought of as a small amount, could have a huge impact on the meaning of the data. Should a binary shape space be employed, the mere flipping of one bit could indicate a huge shift in meaning of the state, rather than the small shift that may be desired. In both of these situations, domain knowledge can play a pivotal role in the success or failure of such as system.

2.6.1 Adaptation within a System

Infecting antigens drive the development of antibodies within the immune system. These agents of change can be considered to be external infecting antigens, which are driving the immune system to protect the host body from infection, and drive the immune system to adapt to changing antigenic infection. However, when one considers embedded systems, one has to consider whether they really evolve, in the sense previously mentioned. Embedded systems are in their very nature self-contained systems, hence they should not, in principle, be considered as evolving systems. They do however interact with the outside world, which could affect the system. Factors such as electro-magnetic noise, radiation, vibration and temperature may affect the normal operation of the system and thus potentially causing faults. In addition to the above factors, components might fail, which can affect the operation of the system, system consumables may become exhausted, and abnormal human interaction could also affect the system in some way. Any artificial immune system for embedded systems should be able to cope with all of these agents. However, these are not the only agents of change, there might also be changes in the physical components of the system e.g. replacement of faulty parts, upgrade of components or the addition of new components. The concept of adaptation is therefore important.

Any immunised embedded system will need to be able, firstly, to detect such consequences from the agents of change, and secondly, to adapt to them and possibly new ones. It should be noted that the AIS does not have to detect a change in components, but merely the consequences of that change. An analogy can be made with the immune system. Should a host have an organ transplant, the immune system does not know this is a new organ, merely that something has changed and it is no longer recognisable as self, i.e. it has detected the consequence of the change. In an AIS, this can be viewed at two levels. At one level, there are minor adaptations of a system to the environment, e.g. if a component fails, the system should be able to detect the consequences of this failure and reconfigure for continuing to provide a degraded service when available redundancies do not permit the continuation of delivery of the original service. At the other level, there are the issues of possible families of embedded systems, where a whole host of similar embedded systems are developed over time, with similar or different components. What is desired here, therefore, is a system that can have an immune system capable of adapting to new components, new operating conditions etc, without the need to retrain it, but use the immune knowledge of existing embedded systems. This then naturally leads to two areas of reconfiguring the AIS. The first is at design time. A new embedded system (the first of its kind) can have a set of detectors generated that should be capable of working with that system: in essence this is a static generation of detectors, off line, first attempts of this (using a very simple robot system) were tried in Canham & Tyrrell 2002. However, the second area for reconfiguring the system is at run-time. The system should allow new components to be introduced, removed and so on and be able to adapt to these changes: having to re-learn a new set of detectors from scratch is not practical they need to be evolved from the knowledge the AIS already has and can capture from its new hardware/software or environment. It may also be possible to introduce new detectors with a new component, therefore a new component is already endowed with its own immune system, which is then integrated into the systems immune system.

We present a brief review of 2 dynamic AIS applications, one in the context of continual learning and classification (in the context of email filtering) and the other in the area of context aware systems.

2.6.2 Dynamic Learning

The Artificial Immune System for Email Classification, AISEC (Secker, Freitas and Timmis, 2003) seeks to classify unknown e-mail into one of two classes based on previous experience. This system captures both the adaptive nature of the immune system (in being able to adapt its response to emails entering the system) and achieve a stability of dynamically maintaining a set of detectors at a steady level. It does this by manipulating the populations of two sets of artificial immune cells. Each immune cell captures a number of features and behaviors from natural B-cells and T-cells but for simplicity we refer to these as B-cells throughout. These two sets consist of a set of naïve (sometimes called free) B-cells and a set of memory B-cells. Once the algorithm has been trained each B-cell represents an example of an uninteresting e-mail by containing words from that e-mail's subject and sender fields in its feature vector. New e-mails to be classified are considered to be antigens and so to classify an e-mail it is first processed into the same format of feature vector as a B-cell and then presented to all B-cells in the algorithm. If the affinity between the antigen and any B-cell is higher than a threshold, the B-cell is said to recognise the antigen and thus classified as uninteresting. If this antigen is later confirmed by a user to represent an uninteresting e-mail, the B-cell which classified it as such is useful and is rewarded by promotion to a long-lived memory B-cell (assuming it was not already). At this time it is also selected to reproduce by clonal selection. This constant reproduction combined with appropriate cell death mechanisms are features that afford our algorithm its dynamic nature. The user feedback will be given asynchronously to classification but on a regular basis. As the algorithm is designed to address concept drift over long periods, reasonable pauses in this feedback should not cause an undue drop in classification accuracy.

Representation

A B-cell receptor is represented as a two-part vector. One part of the vector holds words contained in the subject field of an e-mail, the second holds words contained in the sender (and return address) fields. The actual words are stored in the feature vector because once set this vector will not require updating throughout the life of the cell. This can be contrasted to the common practice of using a vector containing binary values as the receptor, each position in which represents the presence or absence of a word known to the algorithm. As words are continually being added and removed from our algorithm each cell's vector would have to be updated as appropriate when this action occurs. The two sub-vectors are unordered and of variable length. Each B-cell will also contain a stimulation counter used for aging the cell.

Affinity Measure

The affinity between two cells is a measure of the proportion of one cell's feature vector also present in the other. It is used throughout the algorithm and is guaranteed to return a value between 0 and 1. The matching between words in a feature vector is case insensitive but otherwise requires an exact character-wise match. Given bc1 and bc2 are the cells we wish to determine the affinity between,

Algorithm

The AISEC algorithm works over two distinct stages: a training phase followed by a running phase. This running phase is further divided into two tasks, that of classifying new data and intercepting user feedback to drive evolution. During the training stage the goal is to populate the gene libraries, produce an initial set of memory cells from training examples, and produce naïve B-cells based on mutated training examples. As the B-cells in the AISEC algorithm represent only one class the training set, contains only e-mails the user has explicitly selected as uninteresting. Now the algorithm has been trained it is available to begin the classification of unknown e-mail and population manipulation processes based on user feedback. During this running phase the algorithm will wait for either a new e-mail to enter the system and so be classified or an action from the user indicating feedback. Upon receipt of either of these the necessary procedure outlined below will become invoked. To classify an e-mail, an antigen is created in the same form as a B-cell, taking its feature vector elements from the information in the e-mail and an assignment to a class is made. To purge the population of cells which may match interesting e-mails, the AISEC algorithm uses a two signal approach. The system assumes that signal one has occurred, that is the antigen generated from the classified e-mail has already stimulated a B-cell to have been classified. Signal two comes from the user in the form of interpreting the user's reaction to this e-mail. It is during this stage that useful cells are stimulated and unstimulated cells are removed from the algorithm.

Results

Table 5 summarises the results over a continuous test set. Precision is the percentage of messages classified as uninteresting that really are uninteresting, and recall is the percentage of uninteresting messages classified as uninteresting. AISEC shows a better balance between these two measures. When compared with a naïve Bayesian classifier, it achieves a higher precision at the expense of recall. This demonstrates the naïve Bayesian classifier blocks fewer uninteresting messages, but the ones it does clock are more likely to be uninteresting and is due to a Bayesian classifier's bias towards assigning the majority class to an example. Even though, overall, AISEC yielded the slightly higher accuracy we do not claim it classifies with higher accuracy in general. Instead we believe it is reasonable to conclude that the algorithm performs with accuracy comparable to that of the naïve Bayesian algorithm but with somewhat different dynamics.

2.6.3 Context Aware AIS

A Context Aware AIS is being designed to assist the user through the provision of a user friendly context-aware system that provides an assessment derived from their current context. The system should be implemented on a resource constrained device and must be effective even in the absence of connectivity. In practice, context-aware software running on mobile devices needs to work in a range of networking environments with the real possibility that it must spend a proportion of time working with no connectivity. There are some benefits to autonomy and keeping more information locally on the user's device, particularly if privacy of sensitive contextual data is an issue. The system proposed in (Mohr, Ryan and Timmis, 2004) capitalize in the adaptation and stability properties of clonal selection and the maintenance of memory cells via a simple interacting network. The immune metaphors employed allow for a system that can adapt the contents of the memory structure to an ever changing environment despite noisy data. In addition, the system has the property of stability, in that common occurring patterns will be retained over a long period of time, and perturbations in the input space will not adversely effect the maintenance of memory. In the context of the AIS framework.

Representation

The system's inputs consist of the user's context and possible options (e.g. different activities such as "lunch" or "meeting"). The user's context is represented by an attribute vector, a1,a2, ..., an , which contains attributes along with their attribute identifier — note that attributes can appear in an arbitrary order. Possible options are also represented by attribute vectors, one for each option (options may comprise of an arbitrary number of attributes). Each of these vectors constitute a detector.

Affinity Measures

Affinity is the mechanism by which the distance between two elements is calculated. We use the affinity measure to determine how similar two detectors are, if they are close enough to be neighbors, and how much one can stimulate the other. Measuring the distance between GPS co-ordinates is fairly straight-forward as standard Euclidian distance can be used, but measuring the difference between non-numeric attributes is more difficult, e.g. the difference between two mobile phone cell IDs

Algorithm

The user's behavior is learned from the continuous input of local context. Each attribute in the context attribute vector is presented to the system. First a check is performed to see if the dimension exists to which the attribute belongs, if it does then the attribute stimulates all existing detectors within this dimension — stimulation depends on the distance to all detectors within this dimension, which is calculated using the appropriate affinity function. If the distance to all detectors is greater than a threshold, it is converted into an and added to the dimension and a further check is made to evaluate if that detector should become a neighbor of other detectors. After all the attributes have been considered, cross-dimensional-links between them are created or, if they already exist, are stimulated (in proportion to the affinity value of the link). Both levels decrease due to decay functions.

Results

Figure 9 shows the output after the algorithm iterated through 200000 points. Detectors are represented by small circles and their resource level is visualized by the darkness of the circle. These points where collected over a period of four months for the same journey from a one point to another. The map shows a high activity at both points, and a low to medium activity in between — a standard averaging technique showed a very similar result. The algorithm reduced the points to about 150, and due to the use of detectors the information about the lost points is retained by the stimulation levels of the detector. Furthermore, noisy data which was mostly caused by occasional inaccuracies in the GPS measurements is eliminated by the decay function.

2.7 ROBOT SWARMS

In this chapter, we discuss the balance between adaptability and stability is maintained on an super-organism level of biological hierarchy of complexity as in robot swarms.

2.7.1 Problem statement

We need to trust robots, remotely and in places difficult of access, to do their jobs properly, quickly, creatively, autonomously and reliably. A swarm of small relatively simple cheap robots (distributed system) is more reliable and damage-tolerant than a single expensive and complex robot. Some current advantages of robots in a distributed system are:

- they can remain more or less static; instead of moving, the robots can communicate within the system, thus saving energy and simplifying the design;
- they are simultaneously compact and distributed;
- they can harvest energy and material for themselves and for the entire swarm;
- they are self-repairing, adaptively homeostatic/homeoresic as a group;
- the group can be immortal with regular replacement of individual robots.

In addition it would be advantageous if:

- they could be informed and operate, both globally and locally;
- they could have different working modes in different environments/under different conditions;
- they could save energy if only a few robots were needed for a particular task and the rest could "sleep";
- they could work under harsh conditions due to their adaptability as individuals and as a group.

Nature's prototype for such robots is a colony of social insects (e.g., ants or bees): it is self-organized, self-dependent, self-adapted and self-regulating. To understand the idea of "self-", we need to know how "self-" works. There are two main concepts (Johnson, 2001; Bonabeau, Thiraulaz, 2000).

Our proposal is within the framework of the most pressing problems of adaptive intelligence and artificial life (Boden, 1996). Biological systems are massively parallel and distributed, they use disposable components, they are robust to perturbations in their environment, they learn innovative solutions to problems, and their global structure and behavior are not predictable from simple inspection. The favorite prototype for this kind of system is ants. Colonies of social insect are are not only well adapted (Wilson, 1971; Zakharov, 1980; Deneubourg et al., 1978, Couzin,2003), but show their great abilities in adaptability and optimization.

To develop control mechanism for adaptable robot team behaviour we need to quantify group cohesion so that we can track its development, adaptation and performance of its main function. To investigate and model adaptability and stability we need a method of its detection and "a ruler" to measure it. Adaptability is such type of system quality for which there is no scale for measurement. To improve this situation we propose as an universal "a ruler" index of nominal entropy (INE), measuring diversity of any suitable parameter in a system.

2.7.2 Index of Nominal Entropy

To measure the balance between adaptability and stability, an index of Nominal Entropy can be defined as the relation between group entropy to the maximum possible entropy of the group (all agents are equal). Modelling will then be based on an estimation of chaos and order relations in any parameter of individual and group behaviour, that is relevant to system adaptation to environment.

Entropy can be a measure of several quantities:

- uncertainty;
- uniformity;
- scatter of a distribution;

- o quantity of information (in texts) or of order and disorder – C. Shannon (1949).It was suggested for description of structural information changing in a system in a process of communication.

$$H = -\sum_{i=1}^{n} P_i \lg P_i \tag{15}$$

The main quality of this interpretation is that H is maximum when all probabilities are equal to each other. H_{ma}-H=I – surplus (abundant) information in a system;

- o diversity. In this case H_{max}= lgn – is potential diversity, H is current diversity;

- o predictability – risk in decision making process. In this case entropy is the mean value expectance (first momentum) of $\log P_i$. The unpredictability of event will be $1/\log P_i$;

- o the interpretation of H as mean value expectance (first momentum) of disorder in a system;

Disorder = unpredictability; Structure = predictability

$0 \le H \le \log n$ and depends on the number of elements in a system. We do not need this for monitoring system development when n is changing constantly, or for comparing different systems.

To enable us to deal with different n we will use nominative (normal) entropy:

$$h = \frac{H}{H_{\max}} \tag{16}$$

For any of probability distribution h [0;1] and is the measure of order in a system.

All possible states in a system can be described as structural (rules) or chaotic (free will, creative solutions, which ruin structures in a system, increase unpredictability of its behaviour, etc.) The number of all possible states in a system is Hmax=logn or hmax = 1. In this case h is the relative measure of disorder in a system (chaos) and 1-h is the relative measure of order in a system.

This method works well if we want to see system dynamics as it shown on the example of estimation of nominal entropy (group coherence) in ant hierarchy (Figure 10). We tried this method for investigation of ant management adaptability and found that there are adaptable and non-adaptable, but very adaptive. We chose the ant rank in the hierarchy as a parameter to estimate their collaboration (management).

We found that the most efficient way to keep group coherence is a Triple Elementary Unit. It allows large numbers of ants to join the same team without loss the group cohesion and automatic change (increase or decrease) in the number of ants performing a particular task. The number of hierarchical levels is 5 (the highest number to which ants can count) after which group loses integration. (Bogatyreva, Shillerov, 1998).

For the formation of a command unit, leaders and subordinates have quantifiable differences in their behavior (Figure 11 a):

- leaders do not look for contacts; they pay little attention to subordinates and contact them only when they need to;
- subordinates actively look for contacts with leaders and with each other.

If each ant/robot is given a rank according to the percentage of contacts it rejects, it is possible to draw diagrams of their interactions (Figure 11 b). These concepts also apply to humans in management systems (Bogatyreva & Shillerov, 1998).

The Group Nominal Entropy cohesion index is a working, empirically tested cohesion index and can be used as a basis for our model of adaptable distributed swarm robot behaviour.

So, for robotic swarm, designed for other planets exploration we definitely would like:

- o Self-organisation but under our control;
- o Adaptability but not beyond predictability.

To achieve this, we need Biomimetic approach that gives us a hope for predictable adaptability because, comparing with physical (non-living) systems changes in life are never random (system would not take off in bizarre new direction) and they are open for environmental impacts only at some sensitive moments of their development (semi-open systems).

2.8 CONCLUSIONS

In Paragraph 2.3 we discussed how the balance between adaptability and stability is solved in animal locomotion, and have presented a particular preliminary study which explored the feasibility of combining CPG-and-reflex based control with novel types of muscle-like actuators such as Ionic Polymer Metal Composites (IPMC) in an undulatory worm robot.

We recommend following up on this with a concrete project aiming at developing and constructing a variety of robots for planetary exploration based on similar principles. Such a project would be carried out along four main axes:

- o further simulation studies. Realistic dynamic simulations have a key role to play in the design of both control algorithms and robot structures. We suggest to extend the simulation studies presented above … extensive simulation studies with respect to different parameters of the model;
- o extensive tests with the current worm robot. Study of the friction;
- o design and realization of new prototypes, and application to modular robotics. application of the synergetic IPMC/CNN-based approach to other types of bio-inspired robots, Use of materials that serve both as actuators and as sensors. That replicate the self-stabilizing properties of the musculoskeletal systems of animals. Extension of the approach to robots that reconfigure themselves, e.g. that are made of multiple units which can dynamically attach and detach;

- o design of adaptive CPG-based control. Several important points remain to be investigated in relation to the controllability of the CPG-based mechanisms. In particular the controllers should be able to smoothly adjust the speed and direction of locomotion, and, if needed, the type of gait.

In Paragraph 2.4 we discussed artificial immune systems have much to offer dynamic real-time environments, particularly when homeostasis-type properties are required – it seems to us that this applies to all space missions. Note that in traditional mission equipment the expected outcome of component failure is system failure! That is there is often never the idea that component failure can be used in a dynamic way which might also act as its own recommendation organizer and alter the system, again dynamically (possibly using its memory of previous similar events), to prevent (i) the need for system shut-down and worse (ii) complete system failure.

State-of-the-art technology should provide unparalleled insight into the health of systems by sensing the spectrum of conditions including: electrical, mechanical, temperature, stress etc. These data would feed into our AIS to perform system checking, long-term failure prediction and immediate system recovery (if the prediction is good even before the failure has occurred!)

We have identified a number of properties of the immune system that we wish to incorporate into a system but also issues to solve:

- o Embodiment. Ideally this should be in an actual physical device. Within the general context of AIS, this is the essence of the process and it is likely that any system produced will be part of the embedded system being considered. There are many issues regarding appropriate sensors, inputs, outputs, actuators? etc to be decided;

- Dynamic system and environment. Within any set of processes that require any type of monitoring, there will be many inherently dynamic components. Such system would normally be required to be capable of continual operation and may change the notion of what are appropriate (correct?) conditions over time (for example during peak periods of operation [whatever that might be] certain processes may be "switched in", causing related increases in power consumption as requirements change, these are "switched out" and power consumption reduced);

- Learning and adaptation. The system will have to continually adapt as requirements on the processes change dependant in some respects on an unpredictable environment Certainly akin to adaptable immune system characteristics and memory;

- Communication. There are large amounts of communication going on in the immune system, both in the immune system (cytokines etc) and between systems (immune, neural and endocrine). This is where it becomes possible to maybe captialise on the interactions between innate and adaptive immunity. Considering a distributed set of sensors, danger may not be appropriately identified by a single device; multiple sensors may cooperate in diagnosing the condition of the system. It would seem appropriate that this might be self organized and distributed with sensors "presenting" their data to the components in the monitoring system?;

- Timescales. Within the context of such systems we certainly will require different timescales in the processing, presentation, learning and adaptation of the information – generally electronic items change "quickly", whereas environmental aspects, temperature, speed etc might change at a slower rate?.

Bringing all of these ideas together into a single project, would give the opportunity to test some innovative, potentially ground breaking, principles and making a real leap towards future embedded technology.

In Paragraph 2.5 we discussed robot swarms. A specific of a distributed system is that its parts work more or less independent comparing with integrated system where parts and different functions are fixed together and with each other. In both cases environment brings unpredictability in system behaviour, but in the case of distributed system there is a place for a chance even within a system itself – input of each part into different functions can be different. So, this model is suitable for integrated systems as well (with fixed parts and their functions) but the possibility for adaptation will be less.

According to the conditions the meaning of different parts or events behaviour can be crucial for a system – a single agent/event can save/ruin the system. To develop a method for a description of a goal (program) directed behaviour affected by unpredictable situations is a challenge.

3 CONCLUSIONS

The work performed in the framework of the Bionics and Space System Design has lead to the identification of four case studies which have been analyzed and described in the previous chapters. Based on the different and complementary expertise of the Biomimicry Expert Group, D'Appolonia has assigned each of the case studies selected by ESA to a different working team. In order to facilitate the management activities, a responsible has then been selected within each working group. The work has lead to a better understanding of the biological principle, together with a first attempt of an engineering solution.

As described in the previous sections, the case study on the balance between adaptability and stability has concentrated on three aspects of such topic: locomotion control, artificial immune systems and robot swarms taking into account all possible levels of hierarchy in biological systems: sub-organisms (immune systems), organism (locomotion) and super organisms (swarm).

RDL/DMZ/SMC/AB:ad

REFERENCES

Aigeldinger, T. and F. Fish, Hydroplaning by ducklings: overcoming limitations to swimming at the water surface. J Exp Biol, 1995. 198(7): p. 1567-1574.

Alexander R. M. (2003) Principles of Animal Locomotion, Princeton Univ. Press.

Anderson C & Franks (2004) NR Teamwork in animals, robots and humans. *Adv. in Study Social Behav. 33.*

Arena P., C. Bonomo, L. Fortuna, M. Frasca, (2002) "Electro-active polymers as CNN actuators for locomotion control", In Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on, vol. 4, pp. 281–284, 2002.

Arena P., L. Fortuna, M. Branciforte (1999) "Reaction-diffusion CNN algorithms to generate and control artificial locomotion", IEEE Trans.CAS I, Vol. 46, No. 2, pp. 253–260.

Arena P., L. Fortuna, M. Frasca, and G. Sicurella (2004) "An Adaptive, Self-Organizing Dynamical System for Hierarchical Control of Bio-Inspired Locomotion", IEEE Transactions On Systems, Man, And Cybernetics Part B: Cybernetics. August; 34: No. 4. 1823–1837.

Arena P. , L. Fortuna, M. Frasca, "Multi-template approach to realize Central Pattern Generators for artificial locomotion control", Int. J. Circ. Theor. Appl., 30, pp. 441–458, 2002.

Ashby, W.R. (1953). Homeostasis. In H. von Foerster (ed) Cybernetics: Transactions of the 9th conference. New York: Josiah Macy, Jr. Foundation.

Backen SJ, Sendova-Franks AB & Franks NR (2000) Testing the limits of social resilience in ant colonies. *Behav. Ecol. Sociobiol.* 48, 125 - 131.

Bar-Cohen Y. (2000), "Electro-Active Polymers as Artificial Muscles-Capabilities, Potentials and Challenges", Handbook on Biomimetics, Yoshihito Osada (Chief Ed.), NTS Inc., Sect. 11, Chap. 8, "Motion" paper 134, Aug.

Bar-Cohen Y. (2001), "Electro-Active Polymer (EAP) Actuators as Artificial Muscles", SPIE, Topic 1.

Blickhan R. and R.J. Full (1993) Similarity in multilegged locomotion: bouncing like a monopode. Journal of Comparative Physiology, 173:509–517.

Boden MA (1996) (Ed) *The philosophy of artificial life.* Oxford Readings in Philosophy, OUP.

REFERENCES
(Continuation)

Bogatyreva OA (2002) Management in ant colonies: natural analogue of the resolution of social contradictions using TRIZ tools. *ETRIA Conference, Strasbourg*, November , 17-27.

Bogatyreva OA Bogatyrev NR (1987) Socio-metric approach to the dynamics of the individual activity of social insects with the solitary type of foraging. *Proc. 8th All-Union Symp. Novosibirsk*, 162-166.

Bogatyreva OA Shillerov AE (1998) *Social Synergetics*. Novosibirsk, 292pp. (in Russian).

Bonabeau E Thiraulaz G (2000) Swarm Smarts , *Scientific American*, March, 73-79.
Bonser, R. and J. Rayner, Measuring leg thrust forces in the common starling. J Exp Biol, 1996. 199(2): p. 435-439.

Cannon, W.B. (1932) The wisdom of the body. Norton, New York.

Cohen A. H., D. L. Boothe (1999) "Sensorimotor Interactions During Locomotion: Principles Derived from Biological Systems", Autonomous Robots, 7, 239-245.

Collins S.H., M. Wisse, and A. Ruina. (2001) A 3-D passive-dynamic walking robot with two legs and knees. The International Journal of Robotics Research, page In press,.

Couzin ID Franks NR (2003) Self-organized lane formation and optimised traffic flow in army ants. *Proc. R. Soc.* Ser. B 270, 2003, 139 - 146.

Cruse H., T. Kindermann, M. Schumm, J. Dean, J. Schmitz (1998) "Walknet – a biologically inspired network to control six-legged walking", Neural Networks, 11, 1435-1447.

Deneubourg J Goss S Franks N Sendova-Franks A Detrain C Chretien L (1990) *The dynamics of collective sorting: Robot-like ants and ant-like robots, simulation of adaptive behaviour: from animals to animats*, MIT Press, Cambridge, MA.

Di Paolo, E.A. (2005). Autopoiesis, adaptivity, teleology, agency. Phenomenology and the Cognitive Sciences (In press).

Di Paolo, E. A., (2002). Fast homeostatic oscillators induce radical robustness in robot performance. Proc. of SAB'2002, MIT Press

Dowling K. J. (1997) "Limbless Locomotion: Learning to Crawl with a Snake Robot", PhD. Thesis. December.
Gorman, RP & Sejnowski, TJ (1988a) Analysis of hidden units in a layered network trained to classify sonar targets. Neural Networks, 1, 75-89.

REFERENCES
(Continuation)

Gorman, RP & Sejnowski, TJ (1988b) Learned classification of sonar targets using a massively parallel network. IEEE ASSP, 36, 1135-1140.

Hertz, J. Krogh, A & Palmer, RG (1991) Introduction to the theory of neural computation, Persuesu books.

Hirai, K., Hirose, M., Haikawa, Y., and Takenaka, T. (1998) The development of honda humanoid robot. In Proc. IEEE Int. Conf. Robot. Automat. (Leuven, Belgium, May 1998), IEEE, 1321-1326.

Ijspeert A.J. (2001) A connectionist central pattern generator for the aquatic and terrestrial gaits of a simulated salamander, Biol. Cybernetics 84(5): 331-348.

Iversen, S., Iversen, L. and Saper, C.B. (2001). The autonomic nervous system and the hypothalamus. In: Principles of Neural Science, 4th edtion. Eds: Kandel, Schwartz, Jessell. 960-981.

Jet Propulsion Laboratory: http://marsrovers.jpl.nasa.gov/home/index.html.

Johnson S (2001) *Emergence. The Connected Lives of Ants, Brains, Cities and Software*, Allen Lane The Penguin Press.

Keller J.B. and M. Falkovitz (1983), "Crawling of worms", Journal of Theoretical Biology. 104: 417–442.

Kimble G. A. (1961). Hilgard and Marquis' Conditioning and Learning. The Century Psychology Series, Appleton-Century-Crofts, New York.

Kimura, H., Akiyama, S. and Sakurama, K. (1999) Realization of dynamic walking and running of the quadruped using neural oscillators, Autonomous Robots 77(3): 247-258.

Le Cun, Y. Boser, B, Denker, JS, Hendersen, D., Howard, RE, Hubbard, W. and Jackel, LD. (1989) Backpropagation applied to handwritten zip code recognition, Neural Computation, 1, 541-551.

Lippmann, RP (1987) An introduction to computing with neural nets. IEEE ASSP magazine, April 1987, 4-22.

Mahadevan L., S. Daniel, and M. K. Chaudhury, (2004) "Biomimetic ratcheting motion of a soft, slender, sessile gel", PNAS, January 6, vol. 101, no. 1, 2326.

REFERENCES
(Continuation)

Maturana, H.R. and Varela, F.J. (1980) Autopoiesis and Cognition. Dordrecht, Holland: D. Reidel.

McGeer T. (1990) Passive dynamic walking. International Journal of Robotics Research, 9(2):62–82,.

Mingers, J. (1994) Self-producing Systems: Implications and Applications of Autopoiesis. New York: Plenum.

Mowrer O. H. (1956) Two-factor Learning Theory Reconsidered, with Special Reference to Secondary Reinforcement and the Concept of Habit, Psychological Review 63:114-128

Neal, M. & Timmis, J. (2005) Once more unto the breach, towards artificial homeostasis?. Preprint at http://www.cs.kent.ac.uk/pubs/2005/1948/content.pdf.

Niebur E. and P. Erdos (1991), "Theory Of The Locomotion Of Nematodes: Dynamics of undulatory progression on a surface" Biophys. J.; 60: 1132–1146.

Niebur E. and P. Erdos (1993) , "Theory of locomotion of nematodes: control of the Somatic Motor Neurons by Interneurons", Math. Biosci. 118:51-82.

Orlovsky G. N. T. G. Deliagina S. Grillner (1999) Neural Control of Locomotion. Oxford Press.

Ostrowski J. P., J. W. Burdick (1998), "The geometric mechanics of undulatory robotic locomotion", Int. J. Robotics Research 17 (7), pp.683-701.

Pavlov I. P. Pavlov (1927) "Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex", OUP (English edition translated and edited by G. V. Anrep, 1960)

Pavone M., "Modelli matematici per soluzioni biomimetiche nell'ingegneria aerospaziale", Thesis at DIEES, University of Catania.

Peter Eggenberger Hotz (2003), Genome-physics Interaction as a New Concept to Reduce the Number of Genetic Parameters in Artificial Evolution - Proceedings of IEEE Congress on Evolutionary Computation 2003, IEEE Press.

Pfeifer, R. and Scheier, C. (1999) Understanding intelligence, MIT Press, Cambridge.

Pratt, J. and Pratt, G. (1998) Intuitive Control of a Planar Bipedal Walking Robot.

REFERENCES
(Continuation)

Pratt, J., Dilworth, P., and Pratt, G. (1997). Virtual Model Control of Bipedal Walking Robot. In Proceedings ICRA, 193-198.

Quick, T (2003). http://www.cs.ucl.ac.uk/staff/t.quick/autopoiesis.html

Quillin K. J. (1883), "Ontogenetic Scaling Of Hydrostatic Skeletons: Geometric, Static Stress And Dynamic Stress Scaling Of The Earthworm Lumbricus Terrestris", The Journal of Experimental Biology. May 1998; 201: 1871–1883.

Quillin K. J. (1999) , "Kinematic Scaling Of Locomotion By Hydrostatic Animals: Ontogeny Of Peristaltic Crawling By The Earthworm Lumbricus Terrestris", The Journal of Experimental Biology. February 1999; 202: 661–674.

Quillin K. J. (2000), "Ontogenetic Scaling Of Burrowing Forces In The Earthworm Lumbricus Terrestris", The Journal of Experimental Biology. August; 203: 2757–2770.

Raibert M.H. and J.K. Hodgins. (1993) Legged robots. In R.D. Beer, R.E. Ritzmann, and T.M. McKenna, editors, Biological neural networks in invertebrate neuroethology and robotics, pages 319–354. Academic Press.

Rumelhart, D.E. McClelland, J.L. and the PDP research group (1986). Parallell Distributed Processing, Explorations in the microstructure of Cognition,. Volume1 , Foundations. MIT press.

Shahinpoor M. (1995) , "Micro-electro-mechanics of Ionic Polymer Gels as Electrically Controllable Artificial Muscles", J. Intelligent Material Systems and Structures, Vol. 6, pp. 307-314.

Shannon (1949) , C.,*Bell System Tech. J.,* 1948. Reprinted in Shannon CE & Weaver W *The Mathematical Theory of Communication,* University of Illinois Press, Urbana.

Sutton R. S. (1990), Integrated Architectures for Learning, Planning and Reacting based on Approximating Dynamic Programming, Proc. 7th International Conference on Machine Learning, San Mateo.

Taga, G. (1998) A model of the neuro-musculo-skeletal system for anticipatory adjustment of human locomotion during obstacle avoidance, Biological Cybernetics 78(1): 9-17.

Taylor Tim (2004), A Genetic Regulatory Network-Inspired Real-Time Controller for a Group of Underwater Robots" Proceedings of Intelligent Autonomous Systems 8, IOS Press, Amsterdam,

REFERENCES
(Continuation)

Torsten Reil (1999), Dynamics of Gene Expression in an Artificial Genome -- Implications for Biological and Artificial Ontogeny, Proceedings of European Conference on Artificial Life. Springer

Vukobratovic, M., Borovac, B., Surla, D., and Stokic, D. (1990) Biped locomation. Dynamics, stability, control and application. Springer.

Watkins C. J. C. H. and Dayan P. (1992) "Q-Learning", Machine Learning 8:3-4, 279-292,

Wiener, N. (1948). Cybernetics. MIT Press.

Wilson EO (1971) *The insect societies*. Cambridge, Mass. Harvard Univ. Press.

Yamaguchi, J. and Takanishi, A. (1997) Development of a Biped Walking Robot Having Antagonistic Driven Joints Using Nonlinear Spring Mechanism, in Proc. of the IEEE International Conference on Robotics and Automation (ICRA'97), IEEE, 185-192.

Yamaguchi, J., Soga, E., Inoue, S., and Takanishi, A. (1999) Development of a Bipedal Humanoid Robot -Control Method of Whole Body Cooperative Dynamic Biped Walking. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA'99), IEEE, 368-374.

Zakharov AA (1980) Observer ants store of foraging area information in *Formica rufa* L. (*Hymenoptera, Formicidae*). – *Insectes Sociaux*., 27, 3, 203-211

TABLE 1
PARAMETER VALUES OF SYSTEM

μ	s	i_1	i_2
0.5	1	-0.3	0.3

TABLE 2
STRUCTURAL PARAMETERS.

Number of segments N	$N=4$
Total mass m	$m=0.001$ *Kg*
Equilibrium length of longitudinal spring L^0_{long}	$L^0_{long} = 0.025m$
Equilibrium length of lateral spring L^0_{lat}	$L^0_{lat} = 5 \cdot 10^{-3}\, m$
Worm length L	$L = N \cdot L^0_{long} = 0.1m$
Stiffness k_{long}	$k_{long} = 0.5$ *N/m*
Stiffness k_{lat}	$k_{lat} = 0.5$ *N/m*
Damping coefficient D	$D=0.2$ *Ns/m*

TABLE 3
ENVIRONMENT PARAMETERS

Gravity g	9.81 m/s^2
Forward friction coefficient $\mu_{forward}$	0.2
Backward friction coefficient $\mu_{backward}$	5 $\mu_{forward}$ =1

TABLE 4

ACTUATION PARAMETERS.

Wavelength λ	$\lambda=L=0.1m$
Period of the actuation wave T	$T=1s$
Amplitude of the actuation wave A	$A=\frac{L_{long}^{0}}{20}$

TABLE 5
PREDICTIVE ACCURACY FOR CONTINUOUS LEARNING TASK

Algorithm	Classification Accuracy	Recall	Precision
Bayesian	88.05%	67.76%	93.93%
AISEC	89.09% ±0.97	81.13 ±4.71	82.20% ±2.63

D'APPOLONIA

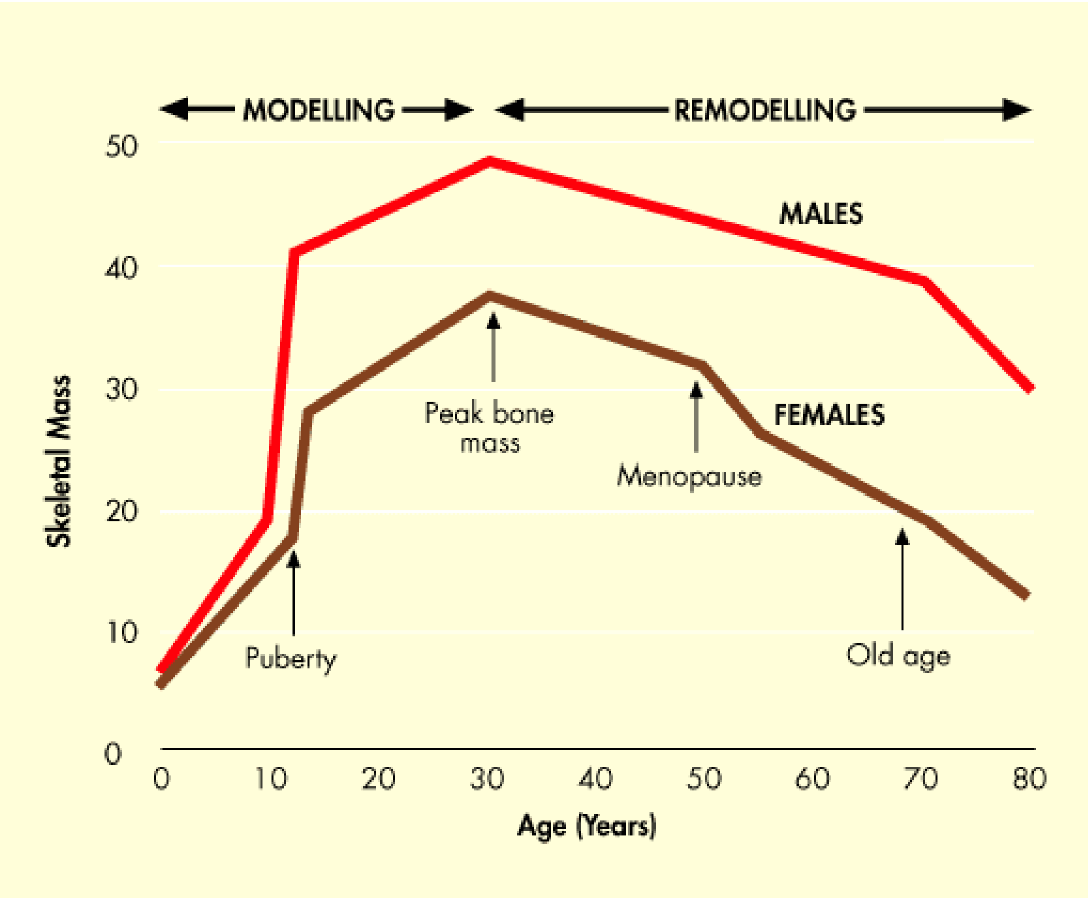
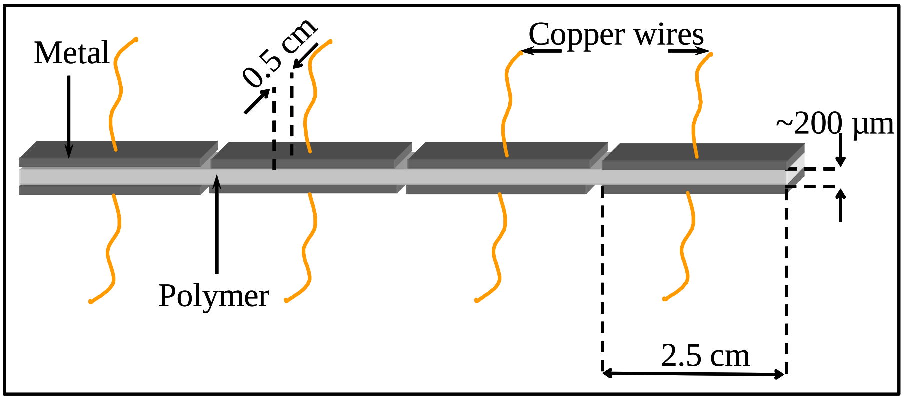


FIGURE 1

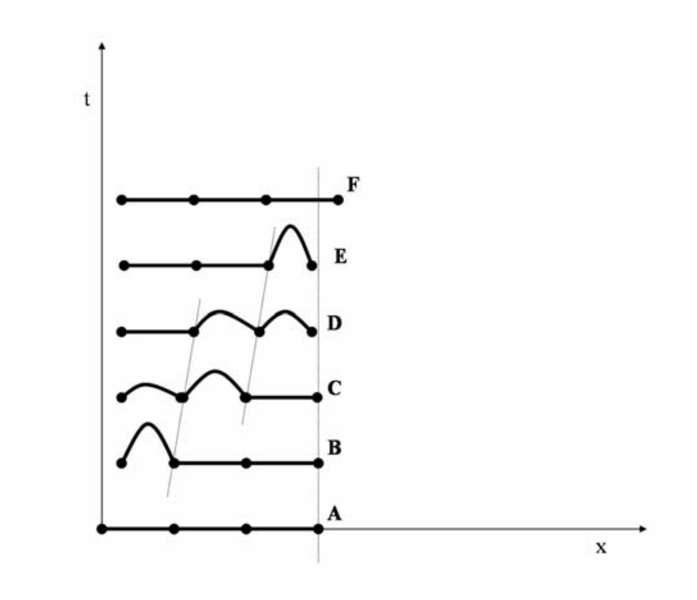
EVOLUTION OF THE BONE MASS
(SOURCE: THE ASSOCIATION OF THE
BRITISH PHARMACEUTICAL INDUSTRY)

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands



(a)



(b)

FIGURE 2

STRUCTURE OF THE IPMC WORM AND SCHEAMATIC REPRESANTATION OF THE LOCOMOTION MECHANISM

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands

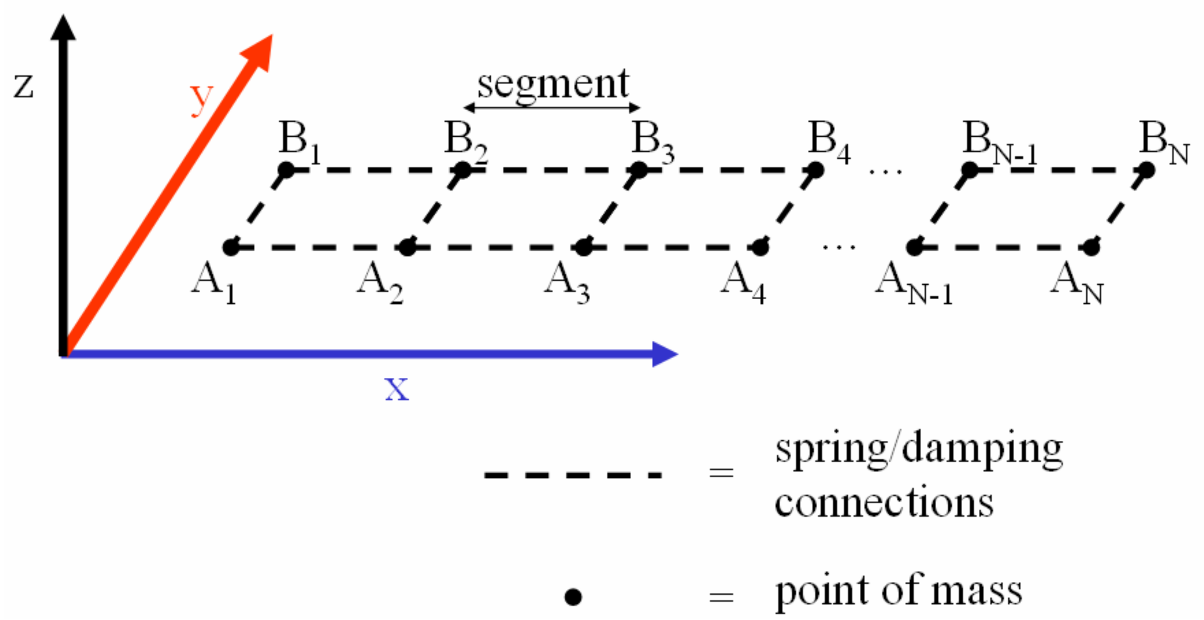
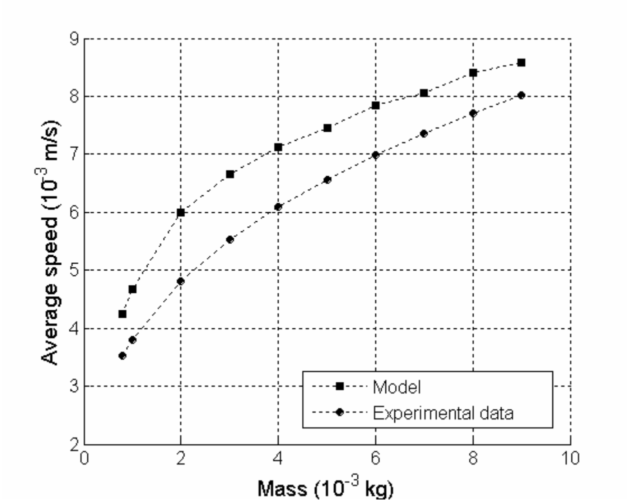


FIGURE 3

STRUCTURE OF THE WORM MODEL

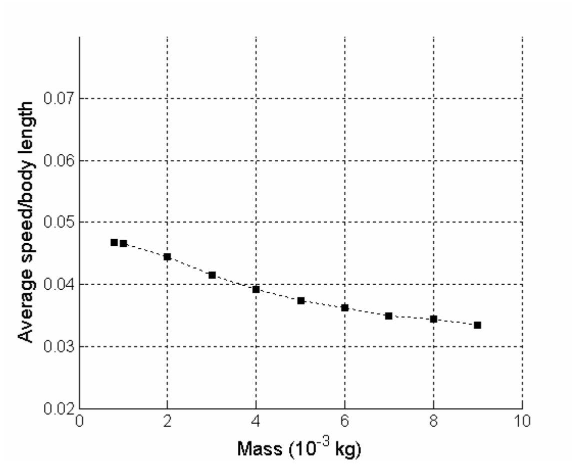
PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands



(a)

Comparison between model and theoretical scaling predicted by Equation (13)



(b)

The average velocity normalized with respect to the body length shows a weak dependence on the mass.

FIGURE 4

AVERAGE VELOCITY VS. MASS

D'APPOLONIA

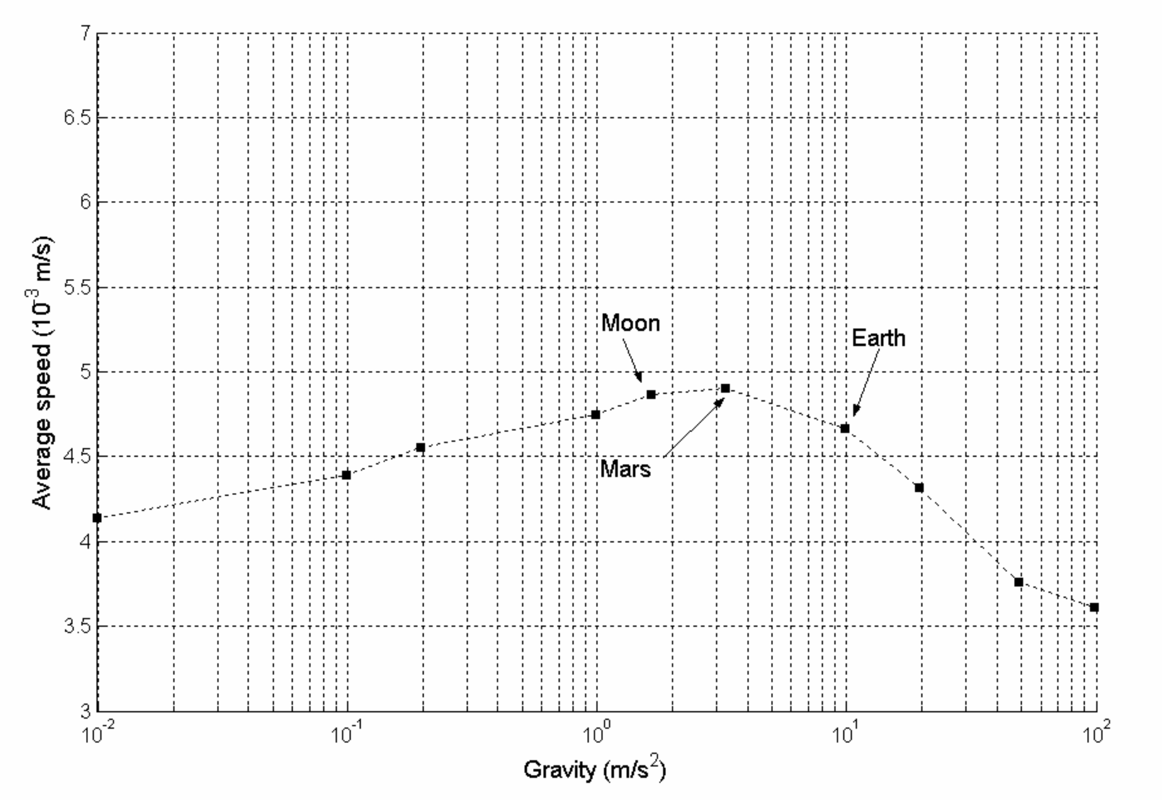


FIGURE 5

VELOCITY VS. GRAVITY ACCELERATION

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands

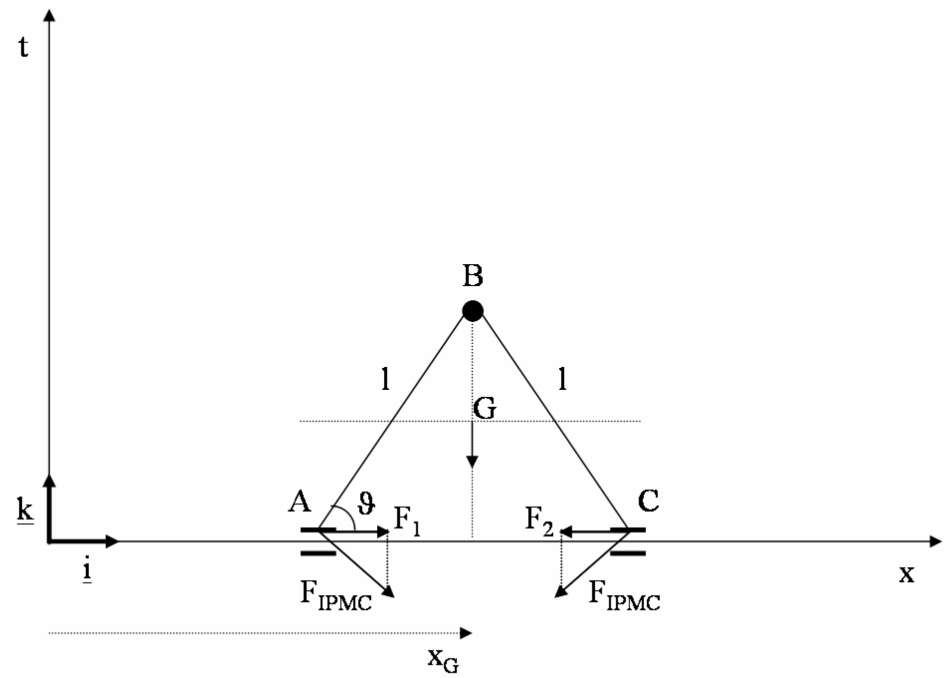


FIGURE 6

SCHEMATIC REPRESANTATION OF THE IPMC ACTUATOR

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands

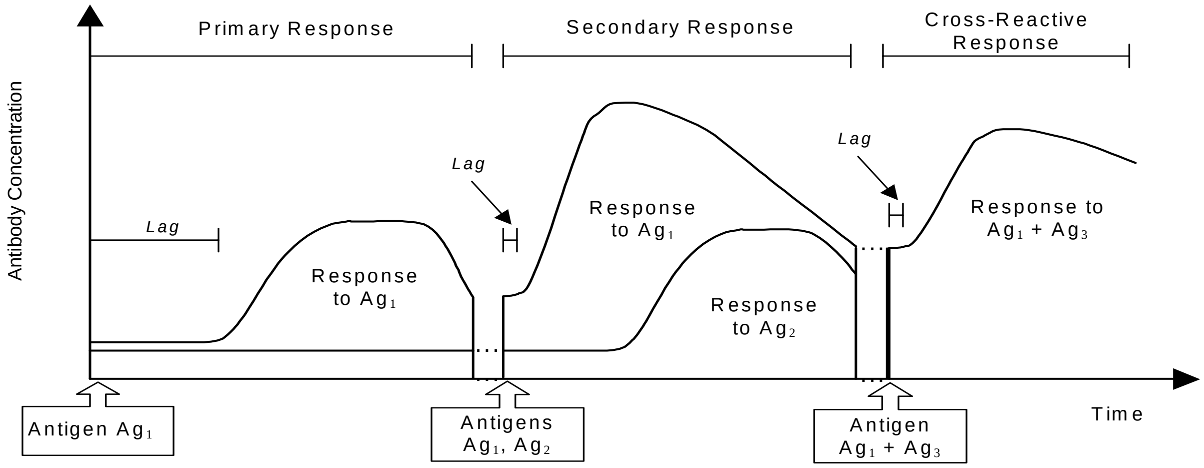


FIGURE 7

LEARNING IN THE IMMUNE SYSTEM

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands

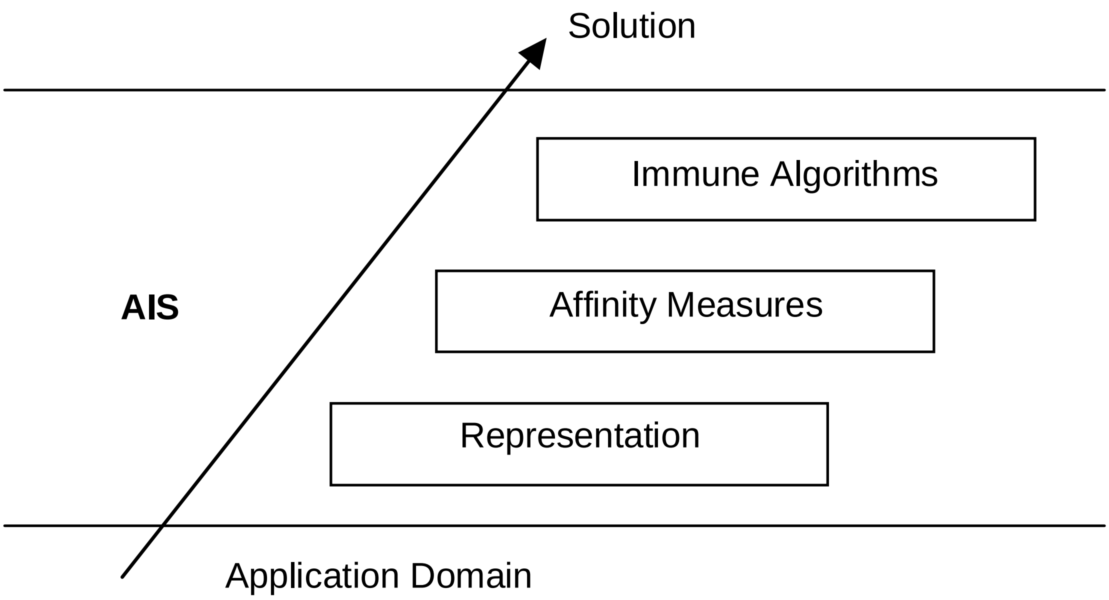


FIGURE 8

A FRAMEWORK FOR AIS

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands

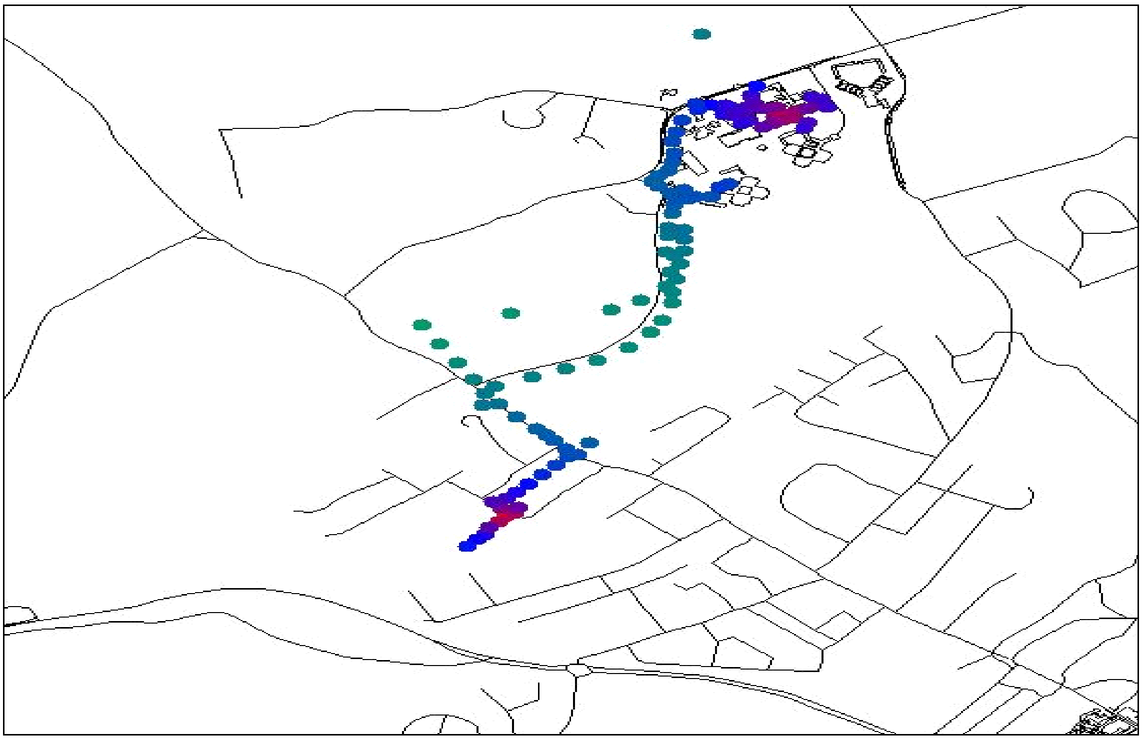
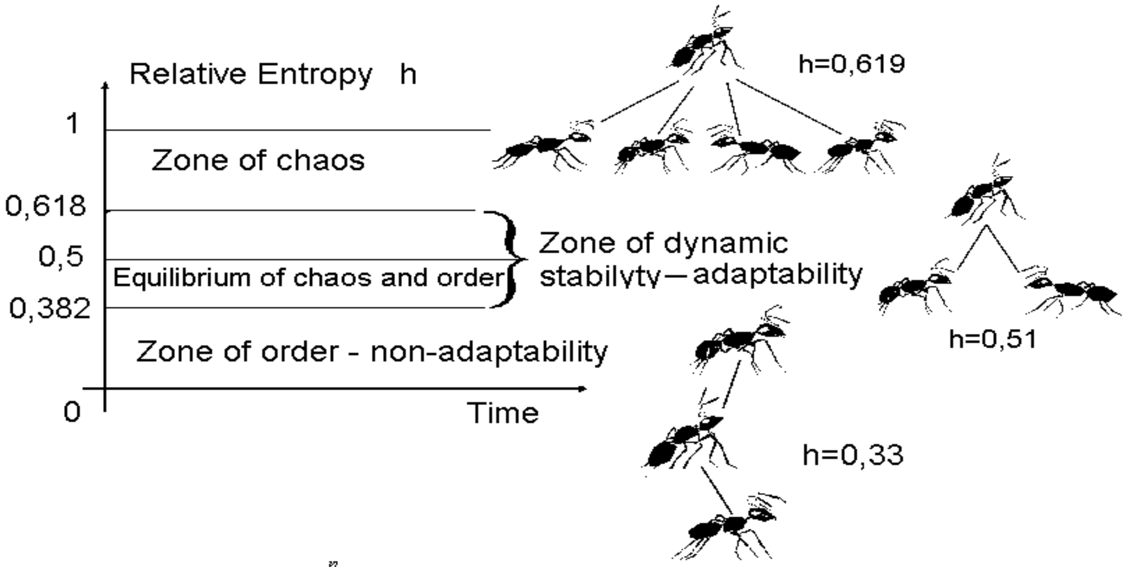


FIGURE 9

IMMUNE NETWORK FOR TRACKING
GPS DATA

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands



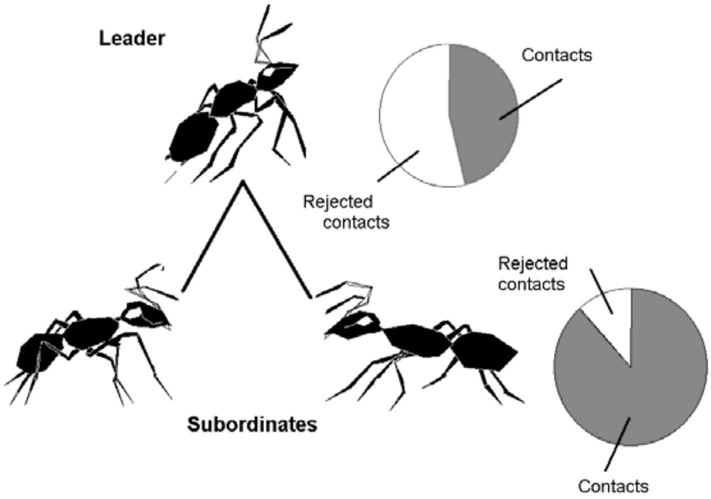
$$h_\alpha = \frac{H_\alpha}{H_{\alpha\,\max}} = \frac{\sum_{i=1}^{n_\alpha} \frac{r_1}{r_\alpha} \log \frac{r_1}{r_\alpha}}{\log \frac{1}{n_\alpha}} \qquad 0 \le h_\alpha^1 \le 1$$

FIGURE 10

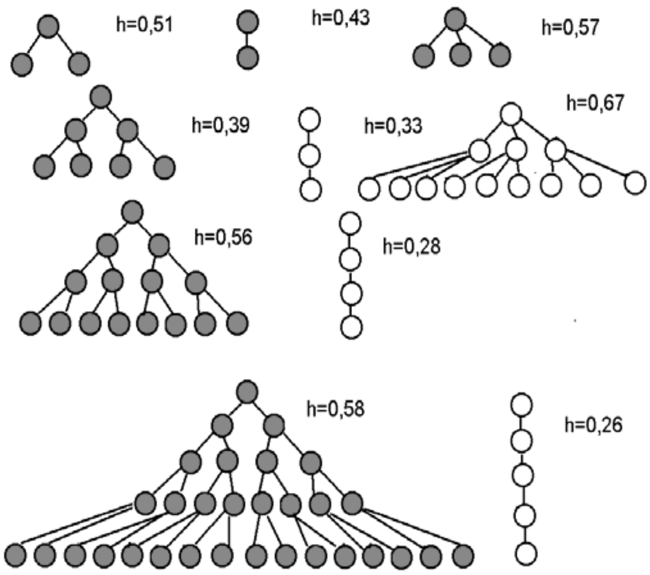
ESTIMATION OF ADAPTABILITY,
INSTABILITY AND NON-ADAPTABILITY,
USING THE NOMINAL ENTROPY INDEX

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands



(a)



(b)

FIGURE 11

FORMATION OF A COMMAND UNIT

PREPARED FOR

ESA, ESTEC
Noordwijk, The Netherlands