Towards navigation without sensory inputs:
modelling Hesslow’s simulation hypothesis in artificial cognitive agents

HIS-IKI-MD-04-008

Alberto Montebelli

Submitted by Alberto Montebelli to the University of Skovde as a dissertation towards the degree of M.Sc. by examination in the School of Humanities and Informatics [2004-10-28].
Abstract

In the recent years a growing interest in Cognitive Science has been directed to the cognitive role of the agent's ability to predict the consequences of their actions, without actual engagement with their environment. The creation of an experimental model for Hesslow's *simulation hypothesis*, based on the use of a simulated adaptive agent and the methods of evolutionary robotics within the general perspective of radical connectionism, is reported in this dissertation. A hierarchical architecture consisting of a mixture of (recurrent) experts is investigated in order to test its ability to produce an 'inner world', functional stand-in for the agent's interactions with its environment. Such a mock world is expected to be rich enough to sustain 'blind navigation', which means navigation based solely on the agent's own internal predictions. The results exhibit the system's vivid internal dynamics, its critical sensitivity to a high number of parameters and, finally, a discrepancy with the declared goal of blind navigation. However, given the dynamical complexity of the system, further analysis and testing appear necessary.

Keywords:
simulators, simulation hypothesis, evolutionary robotics, situated robotics, embodied cognition, mixture of experts.
Acknowledgements

I would like to express my gratitude to many. I will restrain my Mediterranean attitude by limiting such a long list of people to the following. First and foremost I have to thank my supervisors, Tom Ziemke and Henrik Jacobsson, who supported my first steps towards the discovery of such an inspiring research field and magically transformed my ignorance and confusion in sounding computer simulations; my father, who, when necessary, substituted for the missing Italian welfare system; Henrik Svensson, Nicklas Bergfeldt and Andreas Hansson, whose academic role as critics of my dissertation resulted in stimulating suggestions and helped to make my report smoother. Finally, from the number of relevant omissions I will mention Nataša Jankovic, whose bright curiosity made her a perfect non-technical reader.
a Nicoletta,
a mia madre Franca
INDEX

Chapter 1.
INTRODUCTION ............ 7
1.1 The quest for the basic mechanisms of mind ............ 8
1.2 Motivations ............ 10
1.3 Overview ............ 10

Chapter 2.
THEORETICAL BACKGROUND ............ 12
2.1 The Computational Theory of Mind ............ 13
2.2 Different ontologies ............ 15
2.3 The conceptual origins of Evolutionary Robotics ............ 16
2.4 Different representations ............ 18
2.5 Hesslow's simulation hypothesis ............ 20

Chapter 3.
METHODS: EVOLUTIONARY ROBOTICS ............ 22
3.1 A 'reductionist' insight in Evolutionary Robotics ............ 23
3.2 The Khepera robot ............ 23
3.3 A kick-start to Artificial Neural Networks ............ 26
3.4 Genetic Algorithms ............ 33
3.5 What can ER teach us? ............ 37

Chapter 4.
EXPERIMENTAL BACKGROUND ............ 38
4.1 Blind navigation at Skövde ............ 39
4.2 Tani and Nolfi's joint venture ............ 44
### Chapter 5.
**EXPERIMENTAL SET-UP**

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Problem description</td>
<td>50</td>
</tr>
<tr>
<td>5.2 The environment</td>
<td>51</td>
</tr>
<tr>
<td>5.3 The agent's sensory-motor system</td>
<td>52</td>
</tr>
<tr>
<td>5.4 The agent's navigation unit</td>
<td>53</td>
</tr>
<tr>
<td>5.5 The agent's prediction unit</td>
<td>55</td>
</tr>
</tbody>
</table>

### Chapter 6.
**EXPERIMENTS: RESULTS AND ANALYSIS**

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 EXPERIMENT 1: one vs many</td>
<td>58</td>
</tr>
</tbody>
</table>

### Chapter 7.
**DISCUSSION**

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1 Accuracy and effectiveness of the prediction</td>
<td>72</td>
</tr>
<tr>
<td>7.2 Which gating mechanism?</td>
<td>73</td>
</tr>
<tr>
<td>7.3 Which experimental protocol? How many experts?</td>
<td>74</td>
</tr>
<tr>
<td>7.4 About distal interpretations and cultural biases</td>
<td>75</td>
</tr>
</tbody>
</table>

### Chapter 8.
**CONCLUSIONS**

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>77</td>
</tr>
</tbody>
</table>

**BIBLIOGRAPHY**

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>81</td>
</tr>
</tbody>
</table>

**APPENDIX**

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>85</td>
</tr>
</tbody>
</table>
CHAPTER 1:

INTRODUCTION
1.1 The quest for the basic mechanisms of mind

The overall commitment for the work reported in my dissertation is the study of that complex phenomenon we could somewhat vaguely refer to as mind or intelligence. The state of the art in neuroscience is currently exploring and successfully penetrating the secrets of physiological details underlying the functioning of biological nervous systems. Sophisticated quantitative structural modelling for whole families of synaptic ionic channels [Nicholls et al., 2000], sharp anatomical observations at both the cellular and the inter-connective level, chronic implants of extracellular micro-electrode arrays in the brain of primates followed-up through several years of acting and learning [Wessberg et al., 2000], functional analysis through fMRI, PET and SPECT are only examples of the frantic ongoing activity daily producing a rich harvest in terms of scientific data. On the other hand the qualitative approach of ethology, psychology and psychiatry tries to sketch and give a reason for the huge phenomenological landscape of embodied nervous systems acting in the context of their environment. What is missing is in between: at an exponential rate during the last century the scientific community has been endowed with a number of details at the sub-microscopical level while almost completely lacking for a satisfactory explanatory model about how such mechanisms could be mapped in terms of macroscopic observations even when reducing our field of interest to the simplest forms of biological behaviours. The disciplines who intentionally try to fill the gap (i.e. neuro-psychology, psycho-physiology) do nothing but drawing statistical correlations between neural and mental events, missing the point of the actual dynamics of the bridging phenomena. This critical gap is constituted by the cultural concept of mind, borderline between a material world described in physical terms and another that seems to reject such a language [Parisi, 1999]. Drawing a mechanical metaphor it sounds like we were studying a stroke engine accurately exploring its physical structure, the details of its design and alloys; furthermore measuring its performances, power and speeds under different operative conditions, but with no hints about the thermo-dynamical process which generates such performances given the system. The quest is for the 'basic laws' from which intelligence emerges. Maybe the human flight is quite different from the biological one and therefore we do not need a complete reimplementation of the biological intelligence in order to achieve artificial forms of intelligence [Whitby, 1996], but both humans and birds are able to fly since they
actually share (hopefully at a different level of awareness) the exploitation of common basic physical mechanisms. In this sense we are still so far from reasonable answers about the basic mechanisms of the mind that the distinction between 'general intelligence' and 'biological intelligence' is just functional to determine methodological commitments. In the former case an engineering-like approach where the goal, namely implementing intelligence, is unbounded to any operative limitation. In the latter a strong, direct biological plausibility is perceived as a rich source of inspiration. This is at least partly true: even in this case the most characteristic phenomena of biological nervous systems are mostly neglected or only sketchy embedded in the current models: unreliable time dynamics, failures, redundancy, energetic considerations, specific and recurrent through species architectures, neural morphogenesis and maturation. Can we really forget about all that? Can we simulate intelligence using mathematical models thus transducing it in computational terms? Or is it a necessarily 'bloody and juicy' biological phenomenon, carrying some sort of biological causality as Searle, somewhat impressionisticly, claimed [Searle, 1980]? Embarrassingly nobody has a general answer about these questions, thus enriching this field with the best soil for awkward conjectures. Furthermore, after centuries of traditions in science and philosophy of mind, we still lack a reasonable and satisfactory definition for intelligence. I will skip the problem of providing such a definition or to appraise the feasibility itself for such a definition, although I do not consider this omission irrelevant at all.

In what follows I will generally refer to any possible methodology developed in order to implement an artificial form of intelligence as Synthetic Intelligence (SI; [Haugeland, 1985]). I suggest the domain of genuine SI should definitely include the mathematical modelling for any of the most simple nervous systems developed during the biological evolutionary history, scaling up to the most complex. Although the biological history should be considered a valuable source of inspiration, 'should include' does not mean 'should be limited to', thus resulting in a 'science of the possible' more than in a merely descriptive science. Even neglecting the applicative perspective, the potential importance of this research field is definitely unquestionable for the study of mind: any theory of mind should be experimentally grounded in order to achieve a fully scientific relevance. And such fully scientific models with explanatory and predictive powers, inherently described in unambiguous
mathematical terms, would fit the necessary requirements to work as effective keys to a deeper comprehension of the biological nervous systems at any given level of complexity, in a sort of resonant coupling. The experimental methods of Evolutionary Robotics [Nolfi and Floreano, 2000] should be considered part of this wide theoretical perspective.

1.2 Motivations
In the recent years a growing interest in Cognitive Science has been directed to the cognitive role of the agent's ability to predict the consequences of their actions, without actual engagement with their environment. The creation of an experimental model, based on the use of a simulated adaptive agent and the methods of Evolutionary Robotics within the general perspective of radical connectionism, is reported in this dissertation. A hierarchical architecture consisting of a mixture of (recurrent) experts [Jacobs et al., 1991] is investigated in order to test its ability to produce, by life-long learning, an 'inner world', functional stand-in for the agent's interactions with its environment. Such a mock world should be rich enough to sustain 'blind navigation', which means navigation based solely on the agent's own internal predictions, once its inputs have been cut off. Only challenging but partial results in this direction have been reported in the literature so far [Ziemke, 2003]. This constitutes a model for the simulation hypothesis formulated by Hesslow [Hesslow, 2002].

1.3 Overview
In chapter 2 the main theoretical references for my work are introduced, contrasted to the dominant paradigm for the study of the mind (computationalism) which represents the theoretical reference for the traditional Artificial Intelligence: the genesis of constructivism and, consequently, of radical connectionism (the theoretical framework for ER), the idea of constructivist representations (emulators/simulators) and finally Hesslow's simulation hypothesis. The basic methods of Evolutionary Robotics extensively used in my experimental work are addressed in chapter 3. Chapter 4 focuses on the main suggestions emerging from the experiments already reported in literature and directly related to my experimental set-up. The following chapters (chapters 5, 6, 7 respectively) present the detailed description of my
implementation, the results achieved so far and their critical discussion, together with a number of suggestions about necessary possible developments. Finally, chapter 8 reports my general conclusions.
CHAPTER 2:

THEORETICAL BACKGROUND
2.1 The Computational Theory of Mind

Any experimental work is triggered by some personal belief: my implicit assumption is that the computational modelling of cognitive processes is feasible (meaning: it is possible to model cognition in mathematical terms). This statement will not take my work closer to the traditional positions of the so-called *Computational Theory of mind* (CTM) though. This scientific paradigm combines a specific account for mental states- *Representational Theory of Mind* (RTM)- with a specific account of reasoning- *Computational Account of Reasoning* (CAR)- [Standford Encyclopedia of Philosophy]. According to the RTM, statements with a specific semantic value can be built on the basis of a set of amodal, arbitrary symbols carrying both semantic and syntactic value. Intentional states (such as beliefs or desires) are but functional relations between the thinker and the symbolic representation of the statement. The CAR, derived from the developments of analytical philosophy and logical empirism during the late 19th and early 20th century, namely the *formalist program*, developing consistent mathematical reasoning systems where all derivations were strictly grounded in a set of explicitly formulated axioms and syntactic rules of inference with no room for the mathematicians' semantic intuition. The valuable successes of this program together with the notion of *computable functions* (by operational definition the class of functions evaluable in a finite number of steps by a *Turing Machine*) and the reference in a powerfully emerging technology which implemented an approximation for the Turing Machine itself, rose in short time up to the level of a well established 'computational metaphor of the mind'. The CTM paradigmatic arguments rest on the hypothesis that reasoning and cognition is an activity of formal symbol manipulation according to a formal syntax where every operator processes each statement regardless of its semantic value, generating new statements with consistent semantic values. Once free from any commitment to semantic contents the step towards a practical implementation of purely syntactic rules is rather trivial. In the specific historical context electronic technologies offered general purpose programmable machines (computers) as an effective substrate for such a development. The CTM has been the dominant paradigm since the dawn of Artificial Intelligence research in the 1950s, and we can now consider equivalent the two labels within the broader SI research activities. The CTM roots its foundations on a necessarily objectivist ontology: an “objective reality exists and is what it is
independently of any relation with a cognitive subject” [Stewart, 1996], according to a common sense conception of the experienced world. In the formalist program the community of mathematicians endowed the formal systems with their own interpretative capabilities, granting the semantic content for the symbols. Analogously in the CTM the designers of the systems are the negotiators of the correspondence between the representing and the represented world, given that some stable and objective state of things in a 'real world' does exists. Therefore CTM systems are inherently 'heteronomous', critically relying on the designers' intervention in order to remain semantically grounded in the world [Stewart, 1996]. This determines the essence of the symbol grounding problem: the semantic value of the symbols remains a constitutionally meaningless concept to the machine which manipulates them [Searle, 1980].

The terms under which Newell and Simon stated their Physical Symbol System Hypothesis within the CTM leave little room to attacks, although that does not make their argument stronger: the necessary and sufficient condition for a system to exhibit general intelligent action is to be a Physical Symbol System (PSS) [Fodor and Pylyshyn, 1988]. Thus any physical system exhibiting general intelligence is an instantiation of a PSS (necessary condition). Conversely a PSS with 'proper internal organization' will exhibit general intelligence (sufficient condition), although we have no hints about what should be considered 'proper' internal organization. Thus what should be a logical valuable achievement grows up to the role of a dogmatic truth: whenever general intelligence is lacking the community of CTM people will still be confident in the correctness of its own approach, still claiming that what is missing is simply proper internal organization. And this lack became rather evident during the last decades when the traditional approach to AI failed most of its ambitious programs (or we should simply say 'natural programs', given the kind of problems biological intelligence effectively faces), for example when dealing with the domain of perception.

Thus the CTM inspired research retreated from the purpose to replicate human-level intelligence to specialized sub-domains (i.e. Knowledge representation, natural language understanding, vision, truth maintenance, plan verification) nursing the "feeling that one day all these pieces will fall into place and we will see 'truly' intelligent systems emerge” [Brooks, 1991].
2.2 Different ontologies

“Contemporary cognitive science, it is fair to say, displays a deep-seated commitment to a representational view of the mind. According to such a view, intelligence is largely a matter of problem solving, and problem solving is carried out via computations defined over internal representations of salient real-world structures, facts and hypotheses” [Clark and Grush, 1999; emphasis added].

This scientific paradigm, briefly introduced in the previous section as CTM, is so widely accepted in the scientific community that not only the AI researchers, but even the work by most of the 'classical' connectionists (e.g. the famous PDP research group consisting of McClelland, Rumelhart, Hinton, etc.) could be considered sharing a common overall sensitivity [Clark and Grush, 1999].

In the last decades, though, a number of new facts led to the rise of a new sensitivity, in the form of a new and independent paradigm. The work of cognitive scientists and philosophers like Gibson, Piaget, Von Uexküll, Von Glaserfeld and the important contribution by Ernesto Maturana constituted the theoretical ground for an alternative scientific paradigm: constructivism [Stewart, 1996; Ziemke, 2001]. Maturana and Varela set a milestone for the new theoretical perspective introducing the term autopoiesis to designate the circular organization common to both the living systems and the cognitive processes, that they identified as different sides of the same coin [Stewart, 1996]. Autopoietic organisms are:

- operationally closed: autonomous and intrinsically self-referential (the nature of their internal dynamics depend on their own internal organization solely).

- thermo-dynamically open: an autopoietic organization relies on a continuous exchange of matter and energy between the organism and its environment.

Therefore this emerging paradigm is strongly committed with a non-objectivist ontology. According to Maturana: “perception should not be seen as grasping an external reality, but rather as a specification of one” (quoted in [Stewart, 1996]). The subject and object of knowledge become mutually constitutive, distinct but inseparable entities. Rejecting objectivism does not imply the creation of arbitrary
worlds: the resulting inner world can be viable only when it produces a high degree of coherence to the external events. For example in the form of sensory motor correlations: in a complex organism the movement of the eye's motor actuators will not be confused with the movement of the observed object, thus generating perceptual invariants whose creation is "the basis for the emergence of a stable external world populated with objects that exist as such in the cognitive repertoire of the organism itself" [Stewart, 1996]. The nature of these 'objects' critically depend on the peculiar nature of the specific organism and of the activities it engages in its environment.

Some theoretical extensions apply quite naturally to this perspective [Ziemke, 2001; Chrisley and Ziemke, 2002; Clancey, 1997]:

- **situated cognition**: any cognitive agent acts within a physical (and social) environment and cognition largely relies on nature of the cognizer-environment coupling.

- **embodied cognition**: cognition is instantiated in a physical system. The presence of the body is central to most of the cognitive processes. Cognition is the emergent property of a physical system and supports the activities of the body.

This conceptual corpus perfectly matched with a further theoretical referent: the non-linear system theory [Serra and Zanarini, 1990]. This offered the opportunity to exploit the availability of powerful ideas expressed in mathematical (therefore unambiguous) terms: cognitive phenomena could be thought as complex dynamics emerging from elementary neural mechanisms (see for example [Skarda and Freeman, 1987]). The same way logics offered a 'natural' language to AI, mathematics offered a natural language to an emerging class of constructivist models of cognitive processes.

### 2.3 The conceptual origins of Evolutionary Robotics

A radically different approach with respect to CTM is represented by the *evolutionary robotics* (ER). This new methodological approach emerged during the last decade following the way opened by Brooks and others who claimed: "... we believe that human-level intelligence is too complex and too little understood to be correctly decomposed into the right sub-pieces at the moment... Furthermore we will never
understand how to decompose human-level intelligence until we've had a lot of practice with simpler intelligences” [Brooks, 1997]. Hence he suggested to approach the problem of implementing intelligent systems through a bottom-up and incremental design process. Contrasted to the traditional AI robotics approach of functional decomposition (where sensors feed a perception module that interacting with a modelling module provides the system with a symbolic description of the world; a planning module consistently operates on this symbols and trigger an action through an execution module driving the actuators [Fig. 1- top]) Brooks introduced the concept of decomposition by activities leading to a domain-dependent, hierarchically layered architecture. Each activity represents a specific pattern of interaction with the world (i.e. wondering in the environment, obstacle avoidance, seeking for a light source [Fig. 1- bottom]) and is independently implemented in a dedicated layer interacting with the others through messages carrying no implicit semantics (subsumption architectures). Each internal state represent what the agent is currently doing. “In Maturana's terms, Brooks' robots accomplish and sustain a structural coupling between the robot and the walls” [Clancey, 1997], mediated by the robot's sensory motor system.

Fig. 1: the traditional approach through functional decomposition (top); example of subsumption architecture (bottom) (courtesy from [Ziemke, 2001]).

Each layer determines a simultaneous parallel ongoing process: the lower layers
continue their operating sensitive to, although not constrained by, the states in the higher levels. Notice that all the mechanisms for this sensory-motor system are pre-wired and therefore statically built, thus explicitly following the designer's ontology and conception of the relations between the agent and its world. Nevertheless the emphasis on internal representations which characterizes both traditional AI and robotics fades: the agent deals with indexical representations (its 'here and now' following its own specific frame of reference), representing but the ongoing interactions with its environment and thus reacting rather directly to its sensors; each layer extracts only the relevant aspects of the world in the specific context (“...projections of a representation into a simple subspace” (Brooks, quoted in [Clancey, 1997]) while “the world is its own best model”. The complexity of the observed behaviour is mostly nothing but mirroring the complexity of the surrounding environment, given extremely simple mechanisms. Intelligence is an emerging property assigned by the observer: “Intelligence is determined by the dynamics of interaction with the world” [Brooks, 1997].

Robotics is a necessary choice: only an embodied agent physically interacting with the environment through a set of sensors and actuators can keep semantically grounded its processes. Furthermore agents must be situated, acting in the real world that provides both the necessary substrate for interacting and reliable continuity.

The whole set of theoretical implications of the new paradigm were finally integrated and coherently formulated in the pamphlet of Radical Connectionism [Dorffner, 1997], which offered a first possible implementation of constructivism through an experimental, bottom-up approach to cognitive phenomena. The use of situated and embodied agent, self-organization of the cognitive system in free interaction with its environment through a direct and continuous sensory motor loop, rich connectionist state space, constitute the foundations of this operative approach.

2.4 Different representations
In his classical meta-theory of representations, built as a theoretical reference to the design of classical a-modal, arbitrary, symbolic representations, Palmer defines a representation “first and foremost, something that stands for something else” [Palmer, 1978]. The availability of two related but functionally separated entities, a 'represented world' and a 'representing world', is therefore a basic requirement.
Furthermore, given that the representing world is nothing but a model for the represented world, not all but at least some relations in the latter should be structurally preserved in the former. These requirements are so general that not surprisingly a huge polemic about representations spread through the years. The strong emphasis on symbols and symbol manipulation, and the heavy amount of unresolved questions it historically produced, generated a strong reaction against representations. Brooks who explicitly referred to 'intelligence without representations' [Brooks, 1991] and Maturana were among the first to claim that the notion of representation was not a necessary condition for 'elementary cognition' [Stewart, 1996]. The strict application of such an argument, according to the meaning suggested by Palmer, would of course reduce the ambitions of the constructivist paradigm. Stewart suggests that a 'constructivist representation' should “represent the anticipated consequences of an organism's action for its future perception”. Such representations, inseparable from the subject, create the ground for intentional actions: by setting a desired perceptual configuration ('goal') an organism could use mere mental activity to determine the correct sequence of actions to achieve the desired state, without taking the risk of acting among the potential dangers of the real world [Stewart, 1996].

The fundamental role for such internal representations is also defended by Clark and Grush [Clark and Grush, 1999]. They argue that analytically traceable instances of emulator circuitry should be able to enhance real time action by providing the cognitive agent with mock information about the ongoing activity (i.e. during fast voluntary goal-directed movement, where the proprioceptive feed-back signal about position, orientation and trajectory of the arm is physiologically slower than needed for control purposes) and constitute a minimal form of representations, standing-in for specific, usually extra-neural, states (minimal robust representations). This would not only endow the cognitive system with the opportunity of rapid control, but also could support behaviour in total absence of real world feedback, allowing the agent to practice its environment without the need a of physical engagement with it. This mechanism could scale up to more sophisticated representations, authentic functional surrogates for and de-coupleable from the real world. According to the authors, representations of this sort should be considered the necessary condition for any genuinely cognitive agent, and the key to more sophisticated forms of cognition: “the
capacity to think about the distal and absent is grounded in the use of systemic stand-ins and emulation-based strategies” [Clark and Grush, 1999].

2.5 Hesslow's simulation hypothesis

In the recent years the debate on the conceptual role of functional stand-ins for the external world became object of experimental study in neuroscience (for a list of references see [Hesslow, 2002]). In his simulation hypothesis, Hesslow creates the theoretical basis for an account of the richness of the inner world based on internal simulation of perception and behaviour, according to neuro-physiological evidence (and some further reasonable conjectures). His position about the mental activity could be synthesized in his own words: “Thinking consists of simulated interactions with the environment” [Hesslow, 2002].

This hypothesis rests on the postulated existence of three basic mechanisms which give an account for the creation of the inner world:

- **Simulation of behaviour**: similar neural activity is evoked for overt as well as for covert behaviour. The main difference between the two rests in the fact that in the latter the primary motor cortex output is inhibited.

- **Simulation of perception**: similar neural activity is evoked for both actual sensory inputs and mental imagery.

- **Anticipation**: the neural activity in the frontal lobes responsible for covert behaviour can elicit, through associative mechanisms, an approximation for the same sensory state resulting from overt behaviour.

The literature in neuroscience offers a growing body of experimental evidence about the first two mechanisms, whereas indirect evidence and reasonable conjectures based on anatomical data support the confidence in the existence of an anticipatory mechanism.

Sketching the basic mechanism of the interaction of an agent with its environment during overt behaviour [Fig. 2-left] we can see how the actual sensory input \( S \) produces overt behaviour \( R \) through activity elicited in the sensory cortex neural
state \(s\) which leads to motor response preparation in the motor cortex \(r\). The execution of the motor action will produce a new sensory state \(S'\), generating a new motor response \(R'\).

In the case of covert behaviour [Fig. 2-right], the motor response preparation \(r\), through the postulated anticipation mechanism, elicits in the sensory cortex a neural state \(s'\) that in turn determines the motor response preparation \(r'\) and so on. Through these mechanism it is possible to simulate the previous stimulus response sequence by relying on a chain of probable perceptual consequences coherently with the covert motor activity: with no overt behaviour the agent experiences the sensory consequences for a chain of its potential actions.

Interestingly, stating his hypothesis and emphasizing the role of associative mechanisms, Hesslow refuses to participate the polemic about representations: “The simulation hypothesis requires no assumptions about the existence of 'images', 'representations' or other mental entities”. Some theoretical concepts can grow so problematic to constitute a limitation more then a source of inspiration for further research.

![Fig. 2: A stimulus-response sequence during overt (left) and covert behaviour (right). S represents the actual sensory stimulus, R the overt behaviour, s the sensory cortex neural state, r the motor response preparation (courtesy from [Ziemke, 2002]).](image-url)
CHAPTER 3:

METHODS: EVOLUTIONARY ROBOTICS
3.1 A 'reductionist' insight in Evolutionary Robotics

The following sections introduce the basic tools (specifically the core concepts and methods in evolutionary robotics) that I will exploit all along my experimental work, although limited to their most basic and intuitive aspects. Firstly the robot Khepera, that is the physical experimental referent for my simulations, will be briefly presented. Then will follow some basic concepts about the family of the control systems used in ER to endow the robot with a proper behaviour, namely the Artificial Neural Networks (ANNs). Finally will be briefly presented the main ideas underneath the Genetic Algorithms (GA), routinely used in ER to develop the control system of the robot given the specific task.

3.2 The Khepera robot

Khepera is a low cost mobile robot first developed at the Microprocessor and Interface Laboratory (LAMI) at the Swiss Federal Institute of Technology Lausanne (EPFL) and currently produced and commercially available by K-Team S.A. [K-Team, 2002]. Together with the usual practical criteria which should be respected within any project of development of robotic hardware (mechanical robustness, efficient delivery of power supply) the miniature robot Khepera was engineered in order to fit some basic properties specifically oriented to facilitate the design and analysis of behavioural experiments [Nolfi and Floreano, 2000]:

![Khepera robot](image.jpg)

*Fig. 3: a 'greedy' Khepera robot endowed with the gripper module enjoys sugar.*
- Miniaturization: Khepera measures 55 mm in diameter and weighs 70 gr in its basic configuration. It is therefore possible to provide the robot with rather complex environments even when only limited spaces are available and the light weight bounds the mechanical momentum to low values preventing 'clumsy' control systems from producing critical damages to the physical framework of the robot.

- Modular open architecture: commercially available linear, matrix and video vision turret, a gripper with two degrees of freedom and further tools can modularly expand the sensory-motor basic configuration of the robot thank to its extension bus. Furthermore I/O and radio modules enhance the opportunity to collect valuable data for the analysis process.

- Interface: Khepera works under different modalities. It can be fully autonomous when its control system is downloaded on its available memories (128 Kbytes EEPROM and 256 Kbytes RAM) and processed by its own CPU board (Motorola MC68331) while the four NiCd rechargeable batteries provide it with an autonomy of about 30/40 min. It can be used in interactive mode by relying on the RS232 serial-line miniature connector which supports data and control transmission (thus a distal computer can run the control system, store and analyse the experimental data while exploiting the robot's sensory-motor capabilities). Finally Khepera can work in hybrid mode, combining aspects from the two former modalities.

- Compatibility with larger robots: a family of robots with different physical characteristics (namely Koala and K-Alice by K-Team) although endowed with similar sensory-motor capabilities is available in order to test the possibility to scale up the same control system to altered bodily structures.

In its basic configuration Khepera is endowed with a sensory-motor board and a CPU board [K-Team, 2002]. The latter, even in case of interactive mode functioning, provides the system with A/D conversion of the signals coming to and from the former, which in turn controls:
• the *motor system*, consisting of two lateral wheels driven by two independent DC motors, while the robot's body is equilibrated by frontal and posterior sliding pivots;

• the *sensory system*, consisting of eight infrared sensors distributed around the its body, six on one side and two on the other, following a bilateral symmetry with reference to the sagittal section.

These sensors can be independently programmed according to two different modalities [K-Team, 2002]:

• in *active mode* they emit a beam of infrared light and measure the quantity of reflected light, depending on reflectance of the observed surface, that is approximately inversely proportional to the distance within a range of 5 cm;

• in *passive mode* each sensor simply measure the amount of infrared light coming from the environment, roughly proportional to the amount of visible light.

Such a naive sensory-motor system constitutes an obvious oversimplification with respect to even the simplest living forms. Nevertheless it should be explicitly noted that ER drastically changes the problem of design/analysis compared to traditional robotics. In the latter the designer's explicit analysis and understanding of the specific domain is a fundamental first step: his implementation critically rely on the accuracy of his model, the representation of the environment he plainly codes. Only the processes he will explicitly reduce to statements in a proper algorithm describing the ongoing interactions between agent and environment are supposed to drive the behaviour of the system. All the rest should be considered an unintentional space of unpredictability. Traditional engineering never showed any tolerance towards such an unreliable space. The following processes of system implementation and data analysis can be considered rather automatic contrasted to that first phase. Conversely in the case of ER, as will be more clearly expressed in the following two sub-sections, the most delicate initial step consists of the choice of the criteria which measures the
appropriateness of the behaviour of the system facing the given task. The system self-organizes in order to find an effective solution merely enlightened by that driving suggestion. An ambiguous, redundant and contradictory evolutionary space is tolerated. The result of the adaptation process is a complex system where (at least at a first level of analysis, see the following section 3.4- b) no imperative rule given by a designer in the traditional sense of computer science plays a role: the process of analysis (test phase) follows the complexity of the system and is possible and certain only within a limited degree. In a way at this stage of development, while lacking proper general mathematical tools for the analysis of such complex systems, ER should be considered sort of synthetic ethology: the results we can achieve through the analysis of an oversimplified sensory-motor system are valuable and meaningful whether finalized to the extraction of general properties. As a further element of complexity: “Evolved controllers cannot be fully understood by isolating them from their environment and looking at their structure because their functioning has become intimately related to the physical interactions between the robot and the environment” [Nolfi and Floreano, 2000]. A meaningful difference between traditional approaches and ER does exists: is it possible to bound the complex problems presented by real environments (and specifically the ones whose solutions we label as 'intelligence') within a proper exhaustive set of positivist explicit rules we can draw by merely relying on our rationality? Or is it a better way to try to fix the elementary mechanisms and let self-organization of the system do the rest? Rarely reality reacts so quietly towards our attempt to tame it. In a way the solutions generated by an explicitly designed system of rules are never surprising, innovative; in other words never 'intelligent'.

3.3 A kick-start to Artificial Neural Networks

a) A brief introduction to the neuro-physiology of the nervous system

The most impressive feeling we achieve after a first glance on the biological nervous systems derives from the complexity and variety of behaviours emerging from apparently simple and intra/inter-species stereotyped simple units, namely the neural cells. Given the limited computational power which characterizes each isolated neuron, the complex architectures of the biological assemblies of these basic elements are largely responsible for the powerful properties of the evolved nervous systems. A
number of the order of $10^{11}$ neural cells with proper organization constitutes the human nervous system. When exploring such assemblies, for any single cell we have to consider two simultaneous architectural key strategies:

- convergence: each neuron receives signals from about $10^4$ other neurons through its dendritic branches;

- divergence: each neuron affects by its own activity a similar number of other neurons which reaches through its (often relatively long) axonal projections.

The cellular membrane is an efficient selective physical barrier, separating the intra and extra-cellular environments. In normal condition it is possible to measure an electric potential between the two, due to a different ionic concentration, with the former slightly negatively charged with respect to the latter (typical values are -60/-70 mV: resting potential). Whenever opportunely stimulated, the axon can produce a stereotyped signal, a brief depolarization about 2 ms long, which propagates along its length (action potential). Distally along the axons are the interface elements between different neurons (synapses). We can sketch a typical interaction between 2 different neurons: when the action potential propagating along the pre-synaptic axon reaches the synapse, specific chemical agents (neurotransmitters) are released into the synaptic gap and can bind to selective receptors on the dendritic membrane of the post-synaptic neuron, formed by the emerging portion of a protein crossing the membrane lipid bilayer and constituting a sort of channel between the internal and external compartments. This binding modulates the opening of the channel and therefore the permeability to the specific ionic species for whom the channel is selective. The varied distribution of charge determines a local variation of the electronic potential across the membrane. When a number of different local phenomena sum up to reach a critical value (threshold) an action potential will be triggered and start propagating along the post-synaptic axon. Any synapse has a different efficiency: an incoming pre-synaptic action potential can be more or less effective in contributing to the generation of the post-synaptic action potential, depending on the local chemical conditions, often determined by the frequency by whom the synapse is used.
Obviously this is but a sketchy hyper-simplification (for a detailed description [Nicholls et al., 2000]): different types of neural cells with different morphologies and slightly different physical/chemical properties; different types of synapses (chemical, electric, mechanical); a number of different neurotransmitters and specific receptors with completely different space and time dynamics; some synapses (inhibitory synapses) instead of contributing to the generation of an action potential determine an obstacle to its triggering.

The purpose here is but to offer a qualitative description and framework to host the biological meaning of the elements which inspired the the mathematical modelling of the neural systems: therefore not an ambitious reproduction of the complex biological reality, but a way to emphasize its fundamental aspects.

**b) From biology to the model: the neuro-cybernetic project**

The history of artificial neural networks starts around the middle of the 20th century by the fathers of *neuro-cybernetic* along with their commitment towards an operational, practical and implementable definition for intelligence. Since its dawn the corpus of ideas, theories, computational models for this research field found an explicit and direct inspiration in the mechanisms experimentally observed in the biological nervous systems. The physical substrate that the nervous system offers to the mind has always been considered, within the neuro-cybernetics, a fundamental element to the development of intelligent behaviour. The biological inspiration which characterized the mathematical modelling of the neural systems, currently a rich set of models with different level of accuracy, survived along the years although huge variations in the acknowledgement and credit from the scientific community, and with different degrees of strictness depending on the different fields of application, ranging from the most technical (i.e. Signal processing, temporal series prediction) up to the creation of hypothetical-deductive models dealing with elementary cognitive processes (i.e. natural language understanding, pattern classification, navigation in simple environments) and the emulation of social phenomena. In what follows I'll focus on a specific class of models, known as Artificial Neural Networks (ANNs; for a technical description [Haykin, 1998]).
c) The atomic unit: the artificial neuron model

The commonly used model for the artificial neuron constitutes a further simplification with respect to the complexity of the already simplified neuro-biological model just sketched. The scalar product of an n-dimensional input vector and a weight vector with identical dimension represent the argument for a non linear activation function (typically a sigmoid or step function) that produces the output of the artificial neuron. In analytical terms:

\[ y = f \left( \sum_i X_i \cdot W_i + b \right) \]

where \( y \) (a scalar value) is the final activation value, the vectors \( X \) and \( W \) are respectively the input and the weight vectors and \( b \) is the bias.

![Fig.4: the standard ANN model for the neuron.](image)

The analogy with the biological model is obvious: the input vector represents in fact the set of incoming stimuli in form of neurotransmitters; the weights stand-in for the synaptic efficiency; finally the activation function determines the activation state of the neuron following the overall stimulation on the dendrites summed up to the bias: whenever the threshold (the point of sudden change in the activation function) is reached an high activation value represents the triggering of an action potential propagating along the axon.
d) Building up and training the network: emergent global properties

The computational potential for a single unit is definitely inadequate to make a useful object out of it. The powerful properties of the ANNs emerge as global characteristics of a number of interconnected units, so that the output from one unit represents the input for a different one. The network topology, or architecture, constitutes the framework of the interconnections among all the units. Expertise is the only general tool the designer can rely on in order to draw a proper architecture since are almost completely lacking universal mathematical tools able to guide such a choice. Typical architectures are composed by a set of different layers of units, following a feed-forward or recurrent processing flow: a processing flow proceeding linearly from the input to the output layers in the former [Fig. 5- left] is contrasted to a recycling of part of the pre-processed information in the latter [Fig. 5- right].

![Diagram](image)

*Fig. 5: feedforward (left) and recurrent networks (right). The white and black arrows represent the inter-levels and intra-level processing flow respectively.*

Once defined the characteristics on the units (specifically the activation function) and the architecture of the network there is only a degree of freedom left in order to adapt the system's behaviour to the given task, namely the synaptic weights. Such an adaptation is achieved through a learning phase that usually precedes the operative
life of the network (although in some cases operative life and learning are completely overlapped). The learning algorithms used for this purpose can be of very different types: the most commonly used (supervised learning) is constrained to the existence of a 'tutor', an entity able to identify a proper set of couples formed by an input and the related output (training set) through whom the designer primes the network with the correct output for the specific input in the training set. When the faced problem is feasible for the network, an iterative learning process developed through repeated priming of the whole training set (epochs) and relying on proper learning rules is able to modify the weights of the system minimizing the error produced by the network exposed to the training set up to an arbitrary accuracy. The choice of the training set is a further critical element that stresses the capacities and experience of the network designer. The ANNs exploit their plasticity in order to adapt to the general requirements the designers implicitly suggest through the training set. An ambiguous or ineffective training set will produce a network adapted to some form of regularity it discovered during the learning phase and different from the one the designer meant to implement. Further algorithms (unsupervised learning) do not require the presence of the tutor. An important example follows in the next subsection (genetic algorithm).

e) Properties of the ANNs

- Plasticity (learning): ANNs, particularly in unsupervised learning, self-organize their solutions. This has to be contrasted to other common methods of machine learning that, following the dominant paradigm of explicit programming, constrain the machine to learn nothing but what is explicitly considered in the program; and mostly but declarative ('know what' about riding a bike) rather than procedural ('know how' riding a bike) knowledge is allowed.

- Generalisation: the system can go beyond what explicitly taught during the learning phase and is able to coherently process inputs it never saw before.

- Inference: the network specializes through extracting regularities presented during the learning phase, even above the designer's explicit intentions. For example Elman trained a recurrent network to predict the next word in a body of simple natural language sentences. The analysis of the network following the training
phase showed how words with similar syntactic and semantic value were mapped into activation state vectors with closer Euclidean distances within the internal layer (i.e. The words 'plate' and 'glass', breakable inanimate nouns, were mapped close and rather distant from the couple formed by 'woman' and 'girl', animate human nouns) [Elman, 1990; Elman, 1993]. This self-organized, distributed representation is eventually more likely able to produce an extension of the properties associated with 'plate' to 'glass' than to 'girl'.

- Flexibility: the same architecture, after re-learning, has a huge range of applicability.

- Parallelism: the high interconnectivity of the network exploits the basic elementary operator (functional transformation of a weighted sum) splitting the overall computational process in a number of coherent information flows. Fast processing of complex functions becomes thus feasible relying on the naive computational power of the single units.

- Distributed properties: the whole structure holistically participates the learning process and there is no equivalent to a local memory cell to store data or instructions as in traditional computers. This gives the network a peculiar robustness: even a progressive loss of connectivity within the network determines a graceful degradation in the operative performances of the network, at least until a certain critical limit is reached.

It is important to explicitly notice that some practical and theoretical difficulties (especially relative to the phase of analysis) still slow down the exploitation of the full dynamical power of recurrent ANNs. Actually it should be noted that the current use of ANNs is mostly limited to arguably rather trivial architectures with no or very basic dynamical properties.
3.4 Genetic Algorithms

a) A metaphor: 'architectonic' kick-start to GA

Milano is famous at least for two reasons: the first one is its thick fog (in that huge, boringly flat area of Italy it can be so dense up to reduce the visibility to a range of less than 20 m!), the second is its Duomo, an impressive cathedral whose construction started in 1386, inspired to an explicit commitment to vertical shapes, whose dizzy, fractal-like development lays within the frame of the purest gothic architecture [Fig. 6]. Putting the two attractions together let's imagine we park our old-fashion mongolfier right above the big church. Our goal is to determine the height for the highest architectonic element around, although we actually ignore everything about the building and the thick fog does not afford us with any hint about its external aspect. The only tools we are provided with for our measurement are a finite number (let's say 10,000) of radio probes we can drop from our flying site and they'll feed us back with the height of the point they reach. We are moreover endowed with a launch device which accurately determines both the direction and the speed for the drop of each probe (NOTE: in classical mechanics the complete vectorial information about initial position and speed univocally determine the trajectory for a mass collapsed in a point). Up to this point the only reasonable strategy we can adopt is to be confident in chance, throwing each of the 10,000 probes in independently chosen random directions with random initial speed.

Fig. 6: Milano, Il Duomo.
The highest feed-back value we will receive will be our estimation of the height of the building and its accuracy will critically depend on the probability of shooting the highest surfaces versus lower ones. But consider now we can integrate our simple feed-back about height and the initial conditions relative to the drop of the related probe, namely its initial direction and speed. A probe able to feed us back with its own identification label together with the height is all we technically need in order to associate the height to the parameters used by the launch device. We can start dropping 1000 probes (generation 0 of the algorithm) using randomly generated initial conditions and then map the altitude reached for any of them as a function of its initial conditions. We can now select the initial conditions relatively to the probes which reached the 100 highest heights, and then use each of them to set the launch device and drop 10 more probes, producing the next generation of 1000 drops. Unpredictable random events (slight variations in the physical properties of the device, a blow of wind slightly shaking our mongolfier, etc.) will slightly affect the ideal settings, thus dropping the 10 probes following some gaussian distribution around the selected initial conditions. Furthermore we could randomly extract a couple of initial conditions and generate a number of initial conditions as a function of them (i.e. a linear combination, thus exploring a direction connecting two promising points). The procedure is iterated until all of the 10.000 available probes have been dropped. What we obtain is still a random search for the highest point, although chance is now driven by some criteria of efficiency: contrasted to a merely blind random exploration, we exploit all the information available after each generation in order to focus the drops at the next generation. Both the random shifts in the ideal behaviour of the launch device and the combination of the properties of different selected initial conditions allow the maintenance of a certain degree of variability, functional to an effective wide spatial exploration, within the population. A mathematical analysis of this process shows how in 'non pathological conditions' (lack of peaks, singularities in the explored surface) the algorithms gives better results than other random search techniques.

b) GA and ANNs

Within the family of unsupervised learning algorithms ER has a specific commitment to the just introduced general optimization methods, commonly referred to as genetic
*algorithms* (for a detailed description: [Goldberg, 1989]). ANNs architectures can be trained using this approach in order to determine the proper set of weights that effectively fits the task. The GA starts with a given architecture, endowed with an initial random population of synaptic weights and exposes it to the specific task (the analogous of the set of initial conditions for the launch device in the introductory metaphor). The designer of the experiment provides a *fitness function*, a criterion that quantitatively estimates the efficiency the system shows in accomplishing the task (equivalent to the height of the point reached from the dropping probes). Once selected the best individuals (a number of different selection criteria has been reported in the literature), the next generation is created on this basis by introducing slight variations in the set of weights (the random error in the launch device) and eventually combining characteristics from two networks randomly extracted from the fit population (the analogous for creating new initial conditions as a function of the selected ones). This two operations constitute the fundamental operators for genetic algorithms: respectively the *mutation* and *cross-over* operators. The algorithm iteratively explores the multi-dimensional space of all the possible sets of weights searching for the highest value of the fitness function.

Obviously the ANNs' synaptic weights are not the only parameters which can be evolved: the architectures, the activation functions and even the body morphology of the agent can be evolved as well. Furthermore some experimental approaches simultaneously exploited the synergy between phylo-genetic and onto-genetic learning, effectively evolving a proper set of unsupervised learning rules and letting them operate during the robot-environment interaction in order to achieve pheno-typical adaptation for the ANN control system of a mobile agent. Even the role of morphogenesis has been partially investigated [Parisi, 1999].

The genetic algorithms proved their power as tools for mathematical optimization in almost two decades of theoretical maturity. The role of the designer is often undervalued: he actually triggers the whole artificial evolution process and, particularly by stating the fitness function, he becomes a sort of 'demiourgos' of the whole simulated world, whose destiny is critically dependent on that choice.

The concept of *fitness space* provides an interesting qualitative framework for defining and comparing different classes of fitness functions on a 3-dimensional ordinal function scaling [Nolfi and Floreano, 2000]. We can design the fitness
function by choosing its components along three different dimensions [Fig. 7] and contrasting on each of them two qualitative extremes: internal vs external (informations directly available from the automata perspective vs information measured by the observer); functional vs behavioural (a detailed description of oscillatory motor activation in order to achieve walking vs a generic reward for moving in the environment); implicit vs explicit (the generic reward some general behaviour vs a detailed list of rewarding conditions).

![Fitness Space](image)

*Fig. 7: Fitness Space (adapted from [Nolfi and Floreano, 2000]).*

The choice of explicit, functional fitness function, eventually based on data not directly accessible to the automaton, is an effective way to push the control system to achieve the goal. This solution, though, is not conceptually too distant from the imperative and invasive approach of the traditional AI methods. On the other hand an implicit, behavioural fitness function, based on informations directly sensed by the automaton is the best way to let the system free to evolve and self-organize in order to fully exploit its own potentialities in relation to the environment, resulting in a less designer-biased solution [Nolfi, 1998].
3.5 What can ER teach us?
Hence ER is a method for the creation of autonomous control system in robots. It is based on the use of artificial neural networks (ANNs) which learn by genetic algorithms, thus largely relying on self-organization in order to fit the specific task with minimal human intervention, usually according to the general principles of Radical Connectionism.

Summarizing and generalizing the hints suggested by the experimental evidence [Nolfi and Floreano, 2000]: in most cases the solutions found through artificial evolution are qualitatively different from solutions partially or totally hand-crafted by human designers. “Artificial evolution: a) tends to produce a variety of different solutions for a single problem; b) tends to select solutions which rely on emergent behaviour resulting from the interaction between the agent and the environment, and less on internal mechanisms” [Nolfi and Floreano, 2000]. The common-sense bias of the observer/designer can be ineffective (or even misleading) compared to the power of self-organizing systems. Underlying these methods is a concept of mind as a complex process, adaptively shaped as an effective control system, deeply exploiting the potential of the the agent's body and sensory system, of its (artificial) nervous system and its set of adaptive mechanisms, of its environment. All the possible interactions among them all are also taken in account, according to a widely holistic perspective.

Starting with the next chapter will be shown how the general properties of the evolutionary robotics methods can be used in order to address the specific aim of my work.
CHAPTER 4:

SIMULATORS & EMULATORS: EXPERIMENTAL BACKGROUND
The search for the grounding of the simulation hypothesis in the experimental context of evolutionary robotics is not a new idea at the University of Skövde indeed. Two published works [Jirenhed et al., 2001; Ziemke et al., 2003] and a former, unsubmitted Master Project [Hjeml, 2003] represent both the first steps towards that goal and initial, challenging achievements. In this section I will briefly sketch the main ideas and results presented in those papers, in order to clarify how my current work might be interpreted as a contribution to a common ongoing development. I will also introduce a couple of related experiments [Nolfi and Tani, 1999; Tani and Nolfi, 1999] that represent the direct source of inspiration for my specific experimental set-up and finally I will describe the specific problem I am going to address in my research.

4.1 Blind navigation at Skövde
The work reported in 2001 [Jirenhed et al., 2001], represents a first, tentative effort in implementing the simulation hypothesis in an experimental context by realising a minimal inner world and using it to control overt behaviour. A simulated robot endowed with eight standard infrared sensors produced both motor control and prediction about the sensory state at the next time step. The controller was evolved using a standard GA, encoding the connection weights for a recurrent ANN with static architecture (8 inputs, 3 hidden and 3 context units, 2 output units directly driving the two motors and 8 output units representing the prediction for the input sensed at the next time step). The evolution was carried out by relying on a twofold selection criterion: a behavioural fitness function rewarded straight and fast movement and obstacle avoidance; a prediction fitness function rewarded the accuracy of the prediction. The former was used for the selection of 60 controllers from a population of 150; the latter to a further refinement accomplished by selecting from this reduced sample the 30 best predictors that reproduced using the mutation operator only (following, e.g., Meeden [Meeden, 1996]).

The tested resulting behaviour was quite poor. The robot navigates in its environment during a context building phase [Fig. 8- frame 1]; then the sensory input from the environment is cut off and substituted with the robot's own prediction about the next sensory input (internal simulation phase); finally the robot is given the opportunity to recover its orientation (context rebuilding phase). By monitoring the alternation
between internal simulation and context rebuilding [Fig. 8- even and odd frames, respectively] we can easily recognize many wrong actions executed during the internal simulation phase (i.e. frames 4, 6, 8). Interestingly the analysis of the activation of the hidden units shows how this failures are not associated with a lack of internal dynamics triggering context independent behaviours [Fig. 8- bottom]. The qualitative analysis of the results makes explicit at least one reason for the low performances of the system: the robots appears to achieve good predictions for the most active sensors according to its specific navigation strategy (i.e. right side sensors in a right hand wall following strategy), but the limited range of the sensors inhibits the creation of any meaningful prediction about the remaining sensors, although they provide fundamental signals about the occurrence of some new situation.

Fig. 8: the agent simulated by Jirenheid et al. performs its 'context building phase' (frame 1) and then alternates blind navigation (even frames) and context rebuilding (odd frames). The plot of the activation of the hidden units (bottom) illustrates the rich internal dynamics (with permission [Ziemke, 2003]).
For example, adopting the right hand wall following strategy the frontal sensors, which nevertheless trigger the avoidance/turning behaviour during context building and rebuilding phases, are active so seldom that for the EA it does not pay off to predict them accurately.

A further development of the previously reported work was accomplished by the same authors by removing the critical limitation on the sensor range [Ziemke et al., 2003]. In a second set of experiments the sensory input came from a rod sensor covering a visual field of about 30 degrees, feeding the ANN with a more smoothly varying information about the distance of the tracked rod within a range of about 30 cm. In the experiment this information was mapped to the 10 input units of a purely feed-forward ANN [Fig. 9]. The connection weights were evolved in order to develop the control system of the robot's motor and prediction outputs in a simple, symmetrical environment, consisting of four corridors of equal length and shape [Fig. 10- left]. The walls remain below the visual horizon of the rod sensor and thus remaining invisible to the robot (supposed to avoid the contact with invisible obstacles), differently from the four rod placed outside and at the end of each corridor, that represent the only cue actually perceived within the environment. A main difference
from the above first set of experiments consists in the fact that a single behavioural fitness function, rewarding straightforward motion and obstacle avoidance, was used here to separately train the sensory-motor and the prediction module. Therefore the ability to predict the next sensory motor input was not committed any more with the accuracy of the prediction itself, but with its viability (in other words with the quality of the behaviour according to the general requirements of straightforward and collision free movements) when real sensors were replaced by the prediction itself, thus internally simulating 'blind' navigation. A prediction module viable during blind navigation has to necessarily capture at least some of the sensory properties of the actual environment: this is of critical importance in order to keep the agent's behaviour coherently grounded to its world. The satisfactory functional overlapping of the behaviour of an evolved agent during a first lap performed using actual inputs and a second one executed with the agent merely relying on its own predictions [Fig. 10- left] shows how the formation of viable predictions is possible even for long time intervals, although depending on the sensory sequence associated with the specific behaviour (i.e. it is feasible in case the sensory motor module of the agent exhibits smooth turns and not in case it evolved sharp ones).

**Fig. 10:** qualitative analysis of the agent's behaviour in the experiment by Ziemke et al. Left: the agent's trajectory during the first lap, when the robot receives information from the actual physical sensors (black), overlaps with the one during a second lap performed in 'blind' conditions (red). Right: most of the errors checked during blind navigation are 'naive' rotational errors (with permission [Ziemke, 2003]).
Interestingly the viable predictions showed low or no similarity with the actual sensory inputs, although qualitative differences are not surprising at all: when the agent receives external input it can rely on an indexical representation about the 'here and now' of the real world, acting in a purely reactive way; on the other hand, when it has to behave following its own predictions, it needs some elementary form of internal representation of the current engagement with the environment. In fact during blind navigation the hidden units of viable individuals are activated like a sort of internal timer, providing some kind of short-term memory about the ongoing activity. Even the analysis of error conditions is encouraging further research: most of the failing agents exhibit a naïve form of inaccuracy in the timing of turns during blind navigation and consequently a simple shift in their trajectories (which could be easily recovered by allowing some steps for context rebuilding) that still denote a satisfactory internal representation of the environmental general features [Fig. 10-right].

To summarize, this experiment shows how even a simplistic reactive stimulus-response mapping permits a viable prediction at the level of individual time steps, although the scaling up to more complex behaviours and environments, eventually relying on anticipation and planning at a higher level of abstraction) seems problematic with such an elementary architecture.

A further, although partial, attempt is reported in the un-submitted work by Hjelm [Hjelm, 2003]. A simulated Khepera robot was endowed with a long range 'laser' sensor providing the robot with distance information about the obstacles. Two functionally distinct networks were used in order to implement the control system in different simple environments: a sensory-motor unit, trained by GA to produce reliable collision-free navigation; a prediction unit whose outcome represented a prediction for the sensory input at the next step, trained in different comparative experiments by GA, back-propagation and back-propagation through time. Both the environment and the overall network architecture were similar to the ones described in the preceding section. With respect to Ziemke's experiment, where the agents crashing into the walls determined a sudden end of their life-time, here they were allowed to recover from mistakes, thus facilitating the evolution of navigation skills. Training using GA resulted in poor predictions of the sensory state at the next time step (the prediction for the corners are completely omitted), while back-propagation
methods gave better results (both corridors and corners are predicted, although with rough accuracy).

4.2 Tani and Nolfi's 'joint-venture'
An architecture whose potential internal dynamics seem articulated enough and therefore appealing for the prediction of the sensory inputs in non trivial environments was reported by Tani and Nolfi [Tani and Nolfi, 1999]. A simulated robot, endowed with a pre-given basic navigation behaviour (stereotyped wall following) pre-cabled in a navigation module and whose sensory information comes from a belt of long range 'laser' sensors providing the agent with a measure of the distance from the objects in its neighbourhood, moved freely in its environment, consisting of two rooms and whose geometry was not completely trivial (i.e. irregular corridors where each of the corners is designed with a quite different angle). The navigation module independently determines the motor control.

The agent is provided with a hierarchical architecture consisting of two layers and implementing a prediction module [Fig. 11]. Each layer is composed of a mixture of RNN experts, inspired by the work by Jacobs et al. [Jacobs et al., 1991]. Each expert at any level is a RNN, competing with the others in order to adaptively specialize in predicting its input-signal at the next time step during navigation in an initially unknown environment. At the lowest level the sensory-motor information coming form the sensors and from the output of the navigation module constitutes the input. The dynamic gating mechanism produces a weight for each expert's output, dynamically and non linearly rewarding the experts at its level for lower errors in prediction over a specific time window, in order to determine the overall output for the prediction module. The same dynamic values control the adaptive process within each RNN expert too, whose internal weights are updated according to the back-propagation through time algorithm (BPTT) as soon as the actual target signal is available from the environment, that is at the next time step. The dynamical gating mechanism works such that the best predictor over a suitable time interval tends to contribute to the overall output and to learn exclusively. At the highest level the incoming input-signal is in the gating dynamic of the immediately preceding hierarchical level.
Therefore a hierarchically structured set of adaptive systems working as dynamical function generators, namely the experts that dynamically compete with each other both in adapting their predictions to the sensory-motor characteristics of the given environment and in contributing to the overall prediction, is available in the architecture. In general, the self-organized switch of the gate from one best performing expert to another determines a segmentation of the observed input dynamics: the continuous input-flow is “perceived as articulated into sequences of meaningful representative modules” [Tani and Nolfi, 1999]. The fundamental point is
that a switch at the lower level corresponds to changes in the observed dynamical structure of the sensory motor flow, rather than to temporal differences in the sensory-motor state. This can be interpreted as the building process of a structured bottom-up chain of abstractions, a self-organized collection of intra-level symbols emerging from the interactions between the agent (including its physical body, its specific sensory-motor system, its navigation behaviour) and its environment.

While the robot navigates in the environment, the processes of prediction and prediction learning run in parallel: as the agent is located in different positions of the environment the gating mechanism selects the best sensory predictor from the set of the available experts [the winner's identification number is marked in the environment in Fig. 12- left, (a) and (b)].

![Fig. 12: Different experts are the best predictors for different position in the rooms (left). Gate dynamics at the lower hierarchical level (right (a) and (b) for room A and B respectively) and at the upper one (right bottom).](image_url)
After a transitory period (not shown) the gate dynamic at the lower level becomes stable and movements in specific areas of the environment trigger the reliable selection of a related expert: i.e. corridors correspond to expert 4, corners to expert 3 in room A [Fig. 12- right, (a)]. After the robot experiences room B [Fig. 12- left, (b) and right, (b)] corridors still correspond to expert 4, while corners to expert 3 or expert 2 when turning left or right respectively. At the higher hierarchical level the experts observed the dynamic of the gate dynamics at the lower level, which looks different in the different rooms. Two of them developed the ability to predict with some degree of accuracy the gate opening dynamical flow in one of the two rooms each, therefore permitting a sharp self-organized disambiguation of the two different situations, emerging at an higher conceptual level: after a first transitory phase where also expert 0 plays a role, expert 1 and expert 4 take charge for highest activation during, respectively, navigation in room A and B [Fig. 12- right bottom]. Since the learning process is 'totally dynamic' (meaning that the association between expert and specific signal segment is dynamically created on-line), this disambiguation can emerge as a stable result at the higher level only after the stabilization at the lower level has already been reached.

A more detailed analysis of the phase trajectory of the hidden units for each expert at the higher level showed how the gating mechanism selects the intrinsic dynamics (which means the activation dynamic of the hidden units when a feedback loop is installed from the prediction output to the input) which becomes more coherent with the ongoing coupled dynamics (namely the phase trajectory with the actual input feeding the expert) in the two different rooms [Fig. 13]. According to the authors: “Learning to predict the next sensation implies that the system must acquire some analogical model of the observed target”. Such an acquisition is well synthesised by the activation dynamic of the hidden units. These result represents the most direct source of inspiration relatively to my project, whose underlying question might be formulated in the following terms: is such a dynamical 'analogical model' (or in short, although with some risk to trigger the historically old and endless polemic, 'representation') of the environment a sufficient substrate in order to implement an agent capable of effective blind navigation, that is navigation based merely on its own predictions, i.e. with no actual sensory inputs?
Fig. 13: Intrinsic and coupled attractors developed by the hidden units of the five experts at the upper level.

The availability of an architecture exhibiting a set of complex dynamics does not represent a mere technical detail once considered my declared goal (blind navigation). The same authors reported a work on a similar, although dynamically simpler, hierarchical architecture [Nolfi et al., 1999], whose predictive ability at the lower level, based on simple infra-red proximity sensors and on the use of a single RNN for each layer, resulted more focused on the next sensory state than on the dynamical structure of the incoming sensory-motor flow. An analysis of the prediction output showed how it reached a high global accuracy although the lower level completely neglected most of the meaningful changes in the sensory-motor flow during the exploration of a simple environment divided in two rooms (i.e. predictions associated to corners). The situation is similar to the one previously described in section 4.1: rare events, although fundamental to a reliable adaptation to the environment, do not pay off sufficiently to motivate adaptive systems towards accurate predictions. The strategy to directly commit with the global internal dynamics peculiar to the specific environment sounds as a promising alternative in order to escape local minima.
CHAPTER 5:

EXPERIMENTAL SET-UP
5.1 Problem description

The purpose of the current section is to clarify the domain I will address with my work. A simulated robot, endowed with a long range set of sensory inputs which provides the agent with distance information from the surrounding objects in its environment, navigates following a simple, pre-given stereotyped behaviour (wall following). Furthermore, during a learning phase a set of Recurrent ANNs, whose organization is inspired by the architecture presented in the experiment by Tani and Nolfi, compete to produce an accurate prediction of the next sensory input. After training is completed (namely the outcome of the competing networks is sufficiently accurate and stable), a testing phase follows, during which alternatively the actual sensory input [Fig. 14- (A)] or the current best prediction [Fig. 14- (B)] are used to feed the motor unit that produces navigation (details about the gating mechanism among different experts during both the learning and operative phase will follow below). The purpose for my research is testing the feasibility of such a 'blind navigation', relying on the self-organized robot predictions of the next sensory state solely, therefore on a dynamical anticipation mechanism which is presented here as a tentative model for Hesslow's simulation hypothesis. The choice of the specific experimental set-up is functional to the goal: the problems reported in the literature suggest the use of long range sensors [Jirenhed et al., 2001] and more complex architectures showing richer internal dynamics [Ziemke et al., 2003; Hjelm, 2003].

![Diagram of the simulated control system](image)

*Fig. 14: the simulated control system normally operates relying on the sensory input provided by the distance sensor (A). During 'blind' navigation the system input is self-generated by the prediction unit (B).*
My project will be developed according to objectives converging to the declared aim:

- implementation of an experiment inspired to the work presented by Tani and Nolfi;
- analysis of the accuracy of the predictions at the lower hierarchical level (prediction on the next sensory input);
- analysis of the system dynamics when the predictions are used in closed loop to feed-back the prediction unit, while no actual inputs are provided to the network;
- testing of the navigation capabilities when merely sensory prediction is used to feed the motor unit. The performance achieved by a simpler system consisting of a single RNN predictor will be contrasted to the performances of a multi-layer architecture of multiple RNN experts.

Navigation should be here considered as a prototype problem. Independently on its effectiveness, during the test phase the robots will navigate by relying on some form of internal representation of the environment, inherently different from the symbolic representations in classical AI. In this case the representation should be the outcome from a dynamic system which is selected each time its activity becomes coherent with the sensory information coming from the external world [Tani and Nolfi, 1999].

The learning process does not take place in a vacuum: a global, complex context consisting of a number of different crucial contributors (the agent's body with its peculiar sensory system, its specific pre-given navigation, its plastic artificial nervous system and its adaptive potential governed by the operating learning rules; the environment and its noisy feedback as a source of shaping experiences) is, at any given time, both its cause and effect, offering learning potentialities and receiving further shape from learning itself.

5.2 The environment

As directly inspired by Tani and Nolfi [Tani and Nolfi, 1999], the environment for my experiments consists of two rooms obtained within a squared space of 1m x 1m. The 'Stockhol' and 'Roma' rooms [Fig. 15] have been designed as two distinct environments, possibly offering the agent a quite different sensory experience. The Stockholm room (S in the following pages) is a proportioned, regular, squared environment, whose sensory cadences should smoothly evoke 'relaxed' system's
reactions, whereas Roma (R) is a messy, irregularly shaped space, a narrow alley that would offer any 'intelligent enough' artificial agent at least the opportunity to condense a dim wish of escape. The two rooms are separated by a wall, and a door can be opened in order to allow the agent's migration from one to the other.

![Diagram of Roma room](image)

*Fig. 15: the simulated Khepera robot breathes again entering the 'Stockholm' room (top) after a shocking journey in the 'Roma' room (bottom). The beams explicitly visualize the activity of the distance sensors.*

5.3 The agent's sensory-motor system

The agent used in my simulated environment consists of a simulated Khepera robot in its basic configuration. The simple motor system is constituted by the 2 motors only. The usual set of infra-red input sensors is here substituted with an elementary, specifically developed long range distance sensor. It consists of a belt of sensors located in a semi-circle defined by the agent's coronal plane passing through the symmetry axes. Considering the symmetry axis and with respect to the sagittal plane, 7 beams explore the space according to the following directions: 0°, +/-22°, +/-55°,
+/-90°. For each beam is returned the distance of the closest intercepted object (namely: the closest internal or external wall) referred to the robot's rotation axis (the robot's radius is 2.75 cm). The maximum range for each beam is wider than the largest dimension of the environment; therefore any distance can be accurately measured within the chosen environment, generating a linear outcome in the interval [0, 1]. Noise can be injected according to an uniform distribution, as a percentage of the measured distance (therefore the range of the measure error is relative to the specific measure itself). The sensory dynamic achieved through the distance sensor during (and function of) the navigation in the two rooms is presented in Fig. 16.

![Fig. 16: The sensory dynamics in the room S (left) and R (right) during one lap draw a quite different image of the two environments. The beams from the leftmost (-90°) to the rightmost (+90°) are mapped proceeding top-down along the rows.](image)

5.4 The agent's navigation unit

Similarly to the original work by Tani and Nolfi, the agent is endowed with a pre-given, independent navigation behaviour, in the form of a reliable 'right hand wall following' behaviour, that autonomously drives the agent's left and right motors. After a few attempts were made to hand-craft the behaviour this skill was finally evolved in a simple feed-forward network whose input consists in the 7 signals coming from the distance sensor, injected with a +/-2% relative noise; 4 hidden units represent the intermediate processing level towards the final stage of the 2 motor outputs, ranging in the interval [0, 1]. In order to develop the behaviour, a simulated robot was given a fixed starting position in the centre of the upper leftmost (quite short!) corridor of the S room, with downwards orientation. The network was trained by a standard evolutionary algorithm, with all weights initially set equal to zero and the random
number generator seed to 225. For each of the 6400 generations 100 individuals were evolved on a single epoch 800 time steps long (each time step in the simulated world represents 100 ms in the real analogue). A mix of 'tournament' and 'elitism' criteria was used: the 20 best individuals were copied to the next generation, and the 80 further off-springs were selected by an intra-generation tournament and finally updated by adding a random value extracted from a gaussian distribution with unitary standard deviation. No crossover operator was used. The agent's life was suddenly interrupted in case it crashed against a wall or the distance measured from the beam with +90° offset (the one on the extreme right) resulted higher than 8 cm. The former punishment was introduced to prevent the evolution of uneffective behaviours, the latter to encourage actual right hand wall following. The fitness function summed up along the whole life the values for the motor activation scaled by a variable factor. The factor was equal to 1 if the agent measured a distance equal or less then 5 cm on its right side from the exploring beam at +90° and then decayed to zero according to a quadratic law as the distance approached the maximum allowed distance (8 cm).

The evolution was interrupted when the system was still showing a clear performance increment, after the ability to reliably generalize to a new situation, where the physical division between room S and R was removed, was tested. A tendency to collide during the navigation along the longest external wall of room S encouraged further evolution. A further evolutionary opportunity of 3 epochs of 8000 time steps was offered to any individual for 50 generations; the noise level being increased to +/-4% and all of the other evolutionary parameters confirmed. The agents spent the first epoch in room R, the second in room S and finally the third in the open environment after the door separating the two rooms was opened (in the first and last situation the agent started from the middle of the leftmost corridor in room R, whereas in the second from the middle of the leftmost corridor in room S; in all cases with downwards orientation). The best individual, showing a reliable right hand wall following navigation skill in each of the three possible environmental configurations [Fig. 17], was finally selected.
5.5 The agent's prediction unit

Following Tani and Nolfi the architecture used in the present project is articulated in two similar layers, in hierarchical relationship. Each layer [Fig. 18] is endowed with five random generated Elman RNNs, consisting at the first and second hierarchical level of 9 and 5 inputs, and 7 and 5 outputs respectively, and 3 hidden neurons with the relative context neurons. The activation function is a standard sigmoid. The complete sensory-motor information (7 signals from the distance sensor with a +/-2% relative noise injected, 2 signals from the motor output of the Navigation Unit) constitutes the input to each expert at the first level, which during autonomous navigation is trained to predict the sensory state at the next time step by a standard back-propagation algorithm [Haykin, 1999]. Each expert within the layer is continuously competing with all the others in order to produce the best prediction. This 'fight' is ruled by a dynamical gating mechanism structured over \( n \) levels: for each expert a gate value is selected from a set of \( n \) possible values ranging between 0 and 1 and then used as an expert-specific scale factor for the learning rate when back-propagation is applied. For each time step the best predictor is promoted to the next level until it reaches the maximum value, whereas all the others move to the closest
lower level until they reach the lowest one [Fig. 19]. Therefore a stable best predictor over \( n \) consecutive time steps will surely receive a full learning process, whereas a poor expert over the last time steps will receive a fraction of the error correction only; no correction at all when its gating value finally reaches 0. This simple mechanism encourages a strong competition, more than a cooperation, between the experts; a condition whose importance is emphasized and recommended by some authors [Jacobs et al., 1991], although Nolfi and Tani followed a slightly different direction, producing an output as a linear combination weighted by the dynamical gate state.

The experts at the second level receive as an input the array of the gating state at the first level and learn to predict the next state by undergoing a training phase according to the same rules and implications already described for the lower level. The major difference lying in the fact the gate and the experts at the second level are updated every 5 time steps.

Further details and deeper insight over the learning process are illustrated in Appendix A, which introduces to the temporal sequence adopted in the simulation for the routine calls during the learning process at the two distinct architectural levels.

---

**Fig. 19:** the general rule for the dynamic gating mechanism determines a jump towards adjacent (higher or lower) levels until the gate value for the specific expert reaches one of the extremes (gate value equal to 0 or 1).
CHAPTER 6:

EXPERIMENTS: RESULTS AND ANALYSIS
6.1 EXPERIMENT 1: one vs many

The experimental protocol
During the 30000 time steps of the training phase the RNN experts at the lower level competed for the best prediction of the sensory state and the ones at the higher level for the prediction of the lower level gate state at the next time step. During this time period the actual sensory inputs, injected with a 2% relative noise, were used to feed both the navigation and the prediction unit. The navigation unit also provided the prediction unit with the motor state as a further input [Fig. 14- left]. The agent initially moved in the S room and then, after 10000 time steps (1700s in terms of simulated time), the door separating the two rooms opened allowing the transition to the R room. After further 10000 time steps in the R room the door opened again and closed after the agent returned in the S room, where it ran 10000 time steps more. The learning phase was accomplished by supervised training using a standard back-propagation algorithm. The training parameters were set to 0.8 and 0.002, for the momentum and learning rate respectively. During the first 400 time steps (40s in simulated time, roughly, the necessary time to complete one lap) all of the experts received a correction based on the whole learning rate. Afterwards the learning rate was scaled by a factor according to the gating mechanism described in section 5.5. The number of the levels is one of the independent variables considered in the following experiments.

Further 500 time steps during which the actual sensory inputs were cut off constituted the testing phase for the navigation in blind conditions. After a brief, qualitative analysis of the experts' error, the output of the best predictor at the higher level was selected to control the experts' switch at the lower level: its highest predicted gate value determined the selection of the prediction by the related expert at the lower level. This prediction represented the input for both the navigation and the prediction unit [Fig. 14- right].

Case A: single expert
In the first experiment five independent RNNs, with the same architecture described for the experts of the hierachical architecture described in Fig. 18, were trained on the prediction task of the sensory state at the next time step by standard back-propagation.
Fig. 20: the global prediction error curve plotted over the time steps during the initial transitory period of the learning phase (first three laps in room S). The five colours represent five different tested RNNs.

Fig. 21: the sensory prediction error curve (blue trace) for the best single RNN predictor plotted over the agent’s space trajectory (yellow and thicker trace) during the last lap in room S (top) and in room R (bottom).
Obviously in this situation where every expert is trained independently from the others the higher level plays no role. During the first three laps all of the RNNs show a clear decrease of the global prediction error [Fig. 20]. The difficulty of the prediction task in the environment is clearly not uniform: specific locations (i.e. corners) remain harder to be predicted all along the experiments, giving rise to the characteristic time/space periodicity of the error curve [Fig. 21]. The prediction global error fades and stabilises quite soon, although it never actually drops in the most challenging sites (for example the corners). The prediction error in room S and R shows quite different time patterns [Fig. 22] and navigating from one room to the next during the learning phase does not produce a critical increment in the prediction error (not shown).

![Graph](image)

Fig. 22: the sensory prediction error from the five tested RNNs plotted over the time steps during the last laps after the first presentation of room S, of room R and after the second presentation of room S, from the top to the bottom respectively.
On the other hand a qualitative appraisal of the sensory prediction denounces poor performance: the outputs developed from the best predictor shows only dim hints about the sensory dynamics characteristics of the two rooms. Only the frontal beam captures some of the peculiar sensory flow in the S room (i.e. peaks and ramps), whereas the other rows merely exhibit some faint signs of the spatial-temporal periodicity [Fig. 23]. Given these premises the poor outcome during the testing phase, when the agent was asked to move relying on its own predictions only, becomes quite obvious. The trajectory immediately diverges from the available arena cycling along a huge circular attractor right above the proper field [Fig. 24]. It is also interesting to observe the peculiar dynamics of the activation of the hidden units in the two different rooms. The trajectories in the phase space show a slight separation in the S and R room, although the shape of the resulting attractors is qualitatively very similar [Fig. 25].

![Fig. 23: sensory prediction developed by the best tested RNNs during the last lap in the S (left) and R (right) lap in the training phase. The frontal beam exhibits some stronger similarity to the actual input in the S room (left, central row).](image-url)
Fig. 24: After the agent's actual sensory inputs are cut off it leaves its trajectory within the room (the quasi-rectangular lower loop) to engage with its self-sustained free trajectory (circular loop). Walls have been removed from the simulation during the test phase.

Fig. 25: phase space plot of the activation of the hidden units at the lower level during the last two laps in the room S (blue) and R (red). Each column represents a different RNNs; each row a different combination of activation values for the hidden units. The attractors developed by the best tested RNN are displayed in the rightmost column.
Case B: 5 expert; 3 gate levels
The complete architecture, structured in two layers of competing experts as described in section 5.4 was then tested. In a first attempt the gating mechanism was structured on three levels (in other words three possible gate states, respectively mapped in the multiplicative factors 0.001, 0.501 and 1.001, were allowed). A far more complex and harder learning process emerges, especially after the agent experiences training in the R room [Fig. 26]. The main effect of the presentation of the R room, where the green expert results most of the time as the best predictor, seems to lie in the fact that a richer switch among the experts emerges as soon as the S room is presented again (not shown).

Fig. 26: sensory prediction error at the lower level of the architecture as a function of the time steps during the last laps in room S (top), R (centre) and S again (bottom). Different colours map the performance of different experts. The global performance is the best prediction locally available over the five experts.
After a transitory phase the experts stabilize in determining a clear segmentation of the sensory flow: the green expert tends to take charge for giving a prediction in the corridors, the blue one for situations when there is open space on the agent's left side (approaching or leaving a corner, crossing the space between two blocks), the yellow expert when the agent approaches a corner following a long wall [Fig. 27- top]. In the R room, differently from other tests not reported here where right and left turns were discriminated in two different experts, higher level concepts are not preserved in this case. The only higher level concept developed in room R seems to be 'corridor' [Fig. 27- bottom]. When it comes to the analysis of the gate prediction at the higher level, the system fully unwraps its complexity: the rather ordered framework the learning
process shows during the first presentation of room S and R, suddenly develops in a messy and less accurate dynamic when room S is presented for a second time [Fig. 28 and Fig. 29- left]. Comparing Fig. 29- right and Fig.27- left we can see how the gate prediction is accurate when the green expert is providing the best prediction, poor when this is done by the yellow and blue ones. The quality of the sensory prediction is anyway quite disappointing: far from the bright result shown along the frontal direction in Fig. 23, the prediction only displays faint sign of the environmental periodicity (not shown).

Fig. 28: gate prediction error at the higher level of the architecture as a function of the time steps during the whole training phase. Huge qualitative changes affect the system passing from room S (first row, 1 to 10000 time steps) to room R (second row, 10001 to 20000 time steps). The changes become dramatic returning to room S (third row, 20001 to 30000 time steps).
Fig. 29: the switching among the best experts along the whole training phase alternates a frantic change after the second presentation of room S (fifth and sixth row left). The gate prediction error curve obtained using the best available local prediction plotted over the space trajectory during the last lap in room S (right), displays accurate predictions periodically followed by poor ones.

Testing the system in 'blind conditions' after the prediction by the black expert at the higher level triggers the experts' switch at the lower level leads again to a poor behaviour, although self sustained by a more complex dynamical attractor, involving the green, yellow and black expert at the lower level, if contrasted to the previous case [Fig. 30].

Fig. 30: the trajectory performed in blind conditions displays a more complex dynamical attractor in the form of a limit cycle involving three experts, although resulting in a poor behaviour.
The activation dynamic of the hidden units at the lower and upper level complete the analysis of the present case [Fig. 31], showing the system does not lack internal dynamics.

**Fig. 31:** phase space plot of the activation of the hidden units during the last two laps in the room S (blue) and R (red), at the lower (top five rows) and higher level (bottom). Each column represents a different RNNs; each row a different combination of activation values for the hidden units.

**Other Cases: 5 experts; increasing the number of gate levels**

A wide set of further tested parameters led to little or no new outcomes with respect to the previously presented case. Therefore the cases from C to F, whose number of gate levels is sequentially incremented according to Table 1, will be briefly reported in order to focus solely on the most peculiar and surprising aspects.
Table 1: The set of the further tested conditions.

Although in cases C, E, F there are no qualitative differences in the sensory prediction error with respect to Fig. 26, in case D four of the experts present a diverging accuracy after the first 400 time steps, during which all of the experts receive the whole training according to the learning rate, and never recover a better performance (not shown). A similar effect is visible in the gate prediction error in case D. Although case C is qualitatively tuned to the curve shown in Fig. 28, cases E [Fig. 32] and F have in common a different development. Increasing in the number of the levels in the gating mechanism the system tends to segment the sensory flow on a lower number of concepts (not shown).

Fig. 32: Gate prediction error at the higher level of the architecture as a function of the time steps during the whole training phase in case E. The development is far more stable than the one previously shown in Fig. 29.
Fig. 33: The gate prediction error curve obtained using the best available local prediction plotted over the space trajectory during the last lap in room S in case E improves by a whole order of magnitude the higher errors produced in case B.

Whereas in case C the gate prediction error tends to perform quite similarly to Fig. 29- right, in the remaining cases the error decreases noticeably [Fig. 33]. The activity of the hidden units at the higher level tends to collapse to a single point attractor [Fig. 34- right] as their task becomes easier (increasing the number of the gate mechanism levels the expert's switch slows down at the lower levels). At the same time the hidden units at the lower level preserve a vivid dynamic [Fig. 34- left].

Fig. 34: phase space plot of the activation of the hidden units during the last two laps in the room S (blue) and R (red) in case D, at the lower (left) and higher level (right). In the latter the collapse of the internal dynamic is quite apparent.
In all cases the achieved navigation in blind conditions is qualitatively similar to the ones already shown in Fig. 30 and Fig. 24, passing from the extremely huge cycle involving three experts in case D to the tiny one involving two experts only in case C (not shown).
CHAPTER 7:

DISCUSSION
Even a brief, superficial exploration of the reported results, although their preliminary value, is enough to perceive the intrinsic complex nature of the described system. Most of the parameters tested during the implementation of the experiment determine a quite unpredictable, sometimes critical change in the system performances. In this section some of these problems will be discussed, together with possible developments of the system and, compatibly with the goals of the project, with the implicit ambition to broaden the discussion to more theoretical issues.

7.1 Accuracy and effectiveness of the prediction

The fact that the prediction error does not significantly increase during the transition between the two rooms is a sign, already reported by some authors [Nolfi and Tani, 1999], of the inaccuracy of the sensory prediction: obstacles, or new situations do not heavily affect the measure of the accuracy of the prediction. Only one out of the five single RNNs of case A, whose difference during the whole learning process lay on their random initialisation only, developed a qualitatively good sensory prediction, although mostly limited to the frontal beam and in room S. It never happened within the competing experts integrated in the monolithic architecture in the cases B to F. This negative statistics shows how the creation of a sensory prediction endowed with an effective model of the actual dynamic is not a trivial task: the networks learning the prediction task tend to filter off the high frequencies in the actual sensory flow, leaving only faint hints about its characteristic dynamics. Anyway the same fact also shows that during the training phase, after an initial transitory period, the sensory predictions become quite stable and rare events do not heavily influence their outcome. The inaccuracy of the sensory prediction constitutes a problem given the declared goal of blind navigation. In the experiment by Ziemke et al. [Ziemke et al., 2003], reported in section 4.1, accuracy seems less important than effectiveness with respect to the blind navigation task; the fact the prediction embeds some form of representation mirroring the environmental characteristics is nevertheless a necessary requirement to achieve the goal. In the architecture I developed, the capability to produce an effective prediction for the sensory dynamics is clearly missing, whereas the system, endowed with a pre-given navigation unit with no adaptive capability and with a supervised learning strategy, is only offered the possibility to learn how to produce a reliable copy of the sensory input. Therefore in my experimental set-up
accuracy and effectiveness necessarily equate. This is a first direction requiring further development, for example allowing some degree of plasticity in the navigation unit.

Another issue deserving investigation is relative to the choice of the supervised learning algorithm. Although Tani and Nolfi [Tani and Nolfi, 1999] did not report any result about the accuracy of the sensory predictions in their experiment, the use of the back-propagation through time algorithm (BPTT) could produce better results. In the case of BP the system is asked to generate an accurate prediction merely relying on its current input, correlated to the general input-sequence only through the activation of its own internal state. Conceptually BPTT broadens the amount of dynamical information locally provided to the system over the chosen input time window, thus potentially facilitating learning. It also tends more to bound the internal dynamics system to the ones forced during the learning phase [Floreano, 1996].

7.2 Which gating mechanism?

Beside the usual back-propagation parameters, the learning phase in the implemented system presents an open issue about the gating mechanism. How many gate levels should be allowed? Experts that fail the best prediction for a long time span should continue learning, although being updated by a little percentage of the learning rate only, or should remain unvaried?

Considering levels separated by uniform steps in the gate factor, the reported simulations show how an increment of the number of gate levels above some critical value tends to inhibit the opportunity of the experts to cross the gap from the lower to the higher factors, thus encouraging the creation of a single expert that is continuously updated by an high gate factor and tends to dominate all the others. This determines a situation empirically closer to the case A (independent RNN predictor), and consequently the higher level to a better prediction (the gate predictors can easily generate a single lower level gate configuration).

In order to answer the second question some qualitative tests, not reported in the previous chapter, showed how updating the experts which failed the best prediction over a long time window by 10% or 1% of the learning rate or no updating at all did not produce substantial differences.

Undoubtedly the training process does not fulfil the need for reliable and accurate
sensory predictions, capable of a dynamic tuning with the global specific context (that means: accurate sensory prediction in the specific environment and given the specific pre-given navigation).

The gating mechanism is probably the most noticeable difference from the implementation by Tani and Nolfi. The benefits their complex dynamical gating mechanism are not very apparent though. A simpler, far more intuitive gating mechanism was chosen here in order to facilitate the analysis. Nevertheless, relying on a more cooperative mechanism (for example, according to Tani and Nolfi, producing the global prediction as a linear combination of each expert's prediction weighted by the gate values) with respect to the strong competition between experts implemented in my set-up could be useful during the test phase. In fact, within the scenario in the presented implementation, a even a sub-optimal expert tends to take entirely charge for the global prediction until its gate value is overcome by a different expert. Therefore a cooperative gating mechanism could produce locally better global predictions during the transient phases.

7.3 Which experimental protocol? How many experts?

The proliferation in the number of the expert involved in the second level prediction, which suddenly rises in the cases B and C as soon as room S is presented for the second time, could be overcome by a different experimental protocol. Nevertheless, the effect of the room R, creating a richer switch at the lower level compatibly with the results reported by Tani and Nolfi, should be encouraged unless it gets out of control. A number of different protocols, starting from the simplest case of a single visited room should be anyway tested.

In the current implementation the system is a static entity. The number of the experts was arbitrarily pre-defined, thus limiting the maximum number of concepts into which the input flow can be segmented at any given level. A more interesting set-up could be implemented by using a dynamic creation or destruction of the competing experts available at run-time. For example, once all of the experts are engaged with some specific prediction a new expert could be generated. This could enrich the simulation with a more flexible mechanism to match unpredictable environments, tuning the number of the experts to the actual environmental complexity, thus reducing the designer's bias.
7.4 About distal interpretations and cultural biases

Commenting case B, with respect to Fig. 27- top, I distally interpreted the results claiming that the green expert takes charge for the predictions when the agent is moving along a corridor. Changing the point of view and looking at the same figure from a different perspective, we can also try a different guess [Fig. 35]: is the green expert sensitive to the concept 'corridor' or to the sensory-motor sequence 'slight turn right- slight correction left' we can observe in the trajectory when it becomes the most accurate predictor? Is there any relationship, in the expert's perspective, between the corridor and this sensory-motor sequence? Is there a cause? Is there an effect? At this stage of the analysis of the system we actually ignore the answer. Shouldn't we simply interpret the role of the green expert as the result of the specific coupling between the agent and its environment with no further guess? This naive example provides us with a message, which could be more generally extended to our attitude in the study of the mind; therefore a brief digression appears necessary at this point. The biological evolutionary history suggests a subtle coupling between behaviour and language. A mainly phylogenetic behavioural mapping (the model of the coupling between the organism and the environment), receiving further shape by phenotypical learning, precedes in evolutionary terms the creation of a linguistic map [Parisi, 1989]. Behaviour and language are strongly interwoven, although in an asymmetrical way: the latter relying on the former, irrationally inheriting some of its products.

Fig. 35: Is the green expert mapping the concept corridor or a mere sensory-motor sequence?
The study of intelligence has a peculiar reflexive nature, for the mind is engaged with its own description: the linguistic map tends to project itself on the non-linguistic one, following its own terms and transforming the whole mind into language [Parisi, 1989]. The cultural result is a rational-linguistic-anthropo-adult-centric concept of mind, often linearly descending from the classical vision of the human being as a fully rational goal seeker, biasing both the analysis and synthesis processes. Furthermore, the only reference criterion the 'western culture' developed in order to relate, interpret and study our environment has the same 'rational/linguistic/anthropo/adult-centric' nature. We tend to tacitly assume that this cultural tool fully fits and grasps all of the relevant properties and complex nature of the events under observation. Is this true? Or may it even become a misleading trap? Arguing the fundamental importance of a different epistemological attitude is far beyond the intentions of the present work.

Nevertheless, our common sense locks the whole dynamics of cognitive processes in the brain, unifying and giving it a fuzzy name (intelligence), rationally integrated purposes and a proper 'owner' (the agent). On the other hand, some recent experimental evidences in new synthetic approaches to the study of mind (for example, on the role of self-organization in the design of control systems, see [Nolfi, 1998]) and neuroscience (i.e.: noise as a useful way to increase the sensitivity of cray-fish tail sensors by stochastic resonance [Moss et al, 1995]; spatially and temporally uncorrelated spike trains underling globally reliable sensory-motor responses in the medicinal leech [Zoccolan et al, 2002]) suggest a critical analysis of the most basic tools we use in order to observe, describe and design cognitive processes.
CHAPTER 8:

CONCLUSIONS
Often the consequence of building systems governed by a number of non-linear interactions between its components is to obtain complex dynamical behaviours. This is the case with the present work: small changes in one of the many parameters, relative to both the architecture and the learning process, which characterize the tested system determined rather unpredictable, often critical effects on the global outcome.

The declared goal of achieving blind navigation was undoubtedly missed, for the system produced nothing but a collection of similar limit cycles with different geometrical characteristics. A genuine example of experimental failure then? The critical discussion of Chapter 7 implicitly showed how the system still needs a fine tuning set-up of its extremely rich set of parameters. A matter of time: it looks like the current level of investigation led to an only extremely superficial knowledge of the systems' behaviour. Similarly to ancient maps, few known territories still alternate with huge mysterious areas labelled 'hic sunt leones'. An exploration that should be extended: the richness of the dynamical behaviour discovered in the hidden layers at both hierarchical levels are a clear hint about the enormous dynamical potential of the implemented system. The fact my work did not bring sounding arguments to argue that such a potential is able to develop a proper inner world, thus fulfilling the needs for achieving my goal, cannot be considered an argument 'against'. Further architectural developments could even extend such a potential, although the price should be an increment in the already high level of the system complexity.

Should we consider such an emerging complexity a curse? A price too high? Or the nest where a whole class of interesting cognitive processes would finally emerge, thus offering the opportunity of being simulated and therefore analysed at a higher level of detail? In my opinion that part of the connectionist community genuinely interested in modelling cognition should actually meet the challenge of complexity. In the nonlinear dynamics of RNNs, with far more articulated architectures than the ones currently under study, could be hidden the very key to cognition. The biological nervous systems, with their extremely high degree of connectivity, do not look like implementations of engineering-like flow charts, where few modules can communicate with each other through well disambiguatable information pathways. Eventual borderlines between different functional modules are not bounded by the equivalent of sharp lines drawn over a map, but more likely modules tend to melt in each other. My guess is that connectionism is neglecting its holistic potential and its
opportunities to develop when it ignores such an even anatomically apparent biological evidence, referring to conceptual modules to approach more mature engineering fields. With some reasons of course: accepting the challenge of complexity (meaning the use of non-linear, highly connected, self organized huge architectures) requires new mathematical tools, especially under the point of view of the analysis of the resulting systems. Science, which was provided with the basic practical tools for dealing with complexity (namely the computers) only in relatively recent times, still lacks their effective theoretical complements. A lack that will be hopefully compensated in future. For example the development of methods like the statistical tools that the electro-physiologists started to implement in the last few years to analyse the enormous quantity of information achieved over hundreds or thousands of parallel channels by multi-electrode arrays could be useful to investigate the activity of huge, highly connected, self organized networks [Tam and Gross, 1994]. The commitment towards the development of new, powerful theoretical tools for the analysis of complex systems is the way towards a complete scientific maturity. There is still a wide scientific community oriented towards a synergy between the independent potentials of connectionism and more classical forms of AI [Serra and Zanarini, 1990]. The usual claim is that connectionism should take charge for the (detestably) so called 'low level functionalities' (sensory motor coordination, vision, etc.), leaving the 'high level' realm (logical reasoning, language, planning, etc.) to more traditional forms of AI. Other voices defend a different argument, claiming that constructivism should be considered a completely different and independent scientific paradigm that should not accept to be limited to a complementary role with respect to computationalism and should fully develop its own methods and implications relatively to the modelling of even the 'high level' cognitive processes [Stewart, 1996].

The work reported here goes in this direction, dealing with the costructivist flavour of representations. With limited success indeed (my whole project is still fighting with the mud of complexity) but still in that direction. If the final objective is to understand some of the secrets of the human-like intelligence, why deny the possibility that a symbolic vocabulary, even in the traditional sense of AI, could autonomously emerge from an unbiased design within the framework of radical connectionism as soon as it 'fits the particular experimental context'? Is not the self-organization of an impressive
map that we can (distally) interpret as topological representation of an environment, as reported by Floreano and Mondada in their experiment of 'homing' behaviour [Floreano and Mondada, 1996], an enlightening example in this direction?

What do I exactly mean by 'fit the context'? I suggest a proper approach, according to biological considerations, should reduce as much as possible the human bias, exploiting GA and rooting the design of the fitness functions in energetic terms. Elementary control systems would behave in order to achieve a local and immediate maximization of their energy. More complex system architectures might evolve more complex behaviours: global strategies aiming to spatially exploit the whole available environment with reference to a wider temporal planning (even 'unreasonable', in the short term contradictory behaviours might result in a huge future reward). The competition among agents relying on the same and limited resources could express its adaptive power. Populations of agents might finally discover useful forms of cooperation or symbiotic interactions and develop further useful tools, like a proper language and a system of beliefs. Therefore, fit the context should mean maintaining a viable current energetic level in order to survive the ongoing activities which are supposed to provide a satisfactory and in some way appealing future reward. The resulting dynamics should be explained at the light of 'economical principles', relating to the adaptation of means to the end of surviving in a synthetic world that the experimenter should explore with the eyes of an artificial ethologist provided with a powerful set of mathematical tools.

A challenge then, offering the hardest theoretical problems on many different fronts. Offering the highest reward in the faint promise of some light over our own deepest nature too. Nothing but the challenge and promise of science....
BIBLIOGRAPHY


APPENDIX A

Table A.1 illustrates the temporal sequence for the fundamental routine calls during the simulation of the training phase. Considering the first architectural level only, at the first time step (t= 0) the system will (in the exact following order):

1. updates its distance sensor;
2. sets all the 1st level expert's inputs equal to the global sensory-motor state;
3. activates all the 1st level experts.

At the next time step (t= 1, 2, ...), in the interval between the first and the second call it will also:

a. sets the 'example input vector' for the BP algorithm at the 1st level equal to the current 1st level expert's input (that is still up to be updated: therefore the BP example at the 1st level is equal to the sensory motor state at the previous time step);
b. sets the 'target vector input' for the BP algorithm at the 1st level equal to the current sensory state (that is already updated: therefore the BP target at the 1st level is equal to the current sensory state);
c. calls the BP algorithm for all of the experts at the 1st level, providing also an update for the gating mechanism state at the same level.

The 2nd level of the architecture depends on this last point, since it relies on the vector of the gate values at the 1st level as its own input. The (arbitrary) choice was to wait for the first meaningful gate vector available at the 1st level (t= 1) to feed all the experts at the 2nd level and to activate them. After an (arbitrary) time interval of 5 time steps (t= 6, 11, ...., (5*n)+ 1; where n= 1, 2, ...), these two calls are preceded by:

a. set the 'example input vector' for the BP algorithm at the 2nd level equal to the current 2nd level expert's input (that is still up to be updated: therefore the BP example at the 2nd level is equal to the gate state at the 1st level at the previous time step);
<table>
<thead>
<tr>
<th></th>
<th>1st level experts</th>
<th>2nd level experts</th>
</tr>
</thead>
</table>
| \( t = 0 \)             | distance sensor update                                                             | set ANNs input := distance sensor
|                         | set ANNs input := distance sensor                                                 | ANNs activation                                                                   |
|                         | set ANNs activation                                                              | new (null)                                                                        |
| \( t = 1 \)             | distance sensor update                                                             | set BP example := ANNs input
|                         | set BP target := distance sensor                                                  | set BP target := gate
|                         | \( BP = \text{-----------} \)                                                     | set ANNs input := gate
|                         | set ANNs inputs := distance sensor                                                | ANNs activation                                                                   |
|                         | ANNs activation                                                                   | new (null)                                                                        |
| \( t = \ldots \)        | null                                                                              | null                                                                              |
| \( t = 6 \)             | distance sensor update                                                             | set BP example := ANNs input
|                         | set BP target := distance sensor                                                  | set BP target := gate
|                         | \( BP = \text{-----------} \)                                                     | set ANNs input := gate
|                         | set ANNs inputs := distance sensor                                                | ANNs activation                                                                   |
|                         | ANNs activation                                                                   | new (null)                                                                        |
| \( t = \ldots \)        | null                                                                              | null                                                                              |

**TABLE A1:** temporal sequence for the fundamental routine calls during learning.

b. sets the 'target vector input' for the BP algorithm at the 2nd level equal to the current gate state at the 1st level (that is already updated: therefore the BP target at the 2nd level is equal to the current gate state at the 1st level);

c. calls the BP algorithm for all of the experts at the 2nd level, providing also an update for the gating mechanism state at the same level.
APPENDIX B

Experimental set-up

The described experiments, based on a custom development of the YAKS simulator for the Khepera robot [Carlsson and Ziemke, 2001], were simulated on a Sony VAIO Laptop computer, endowed with an AMD Athlon 1.3 Ghz microprocessor and 512 Mbytes RAM, Microsoft Windows XP Professional (Service pack 1a) operating system and compiled by Microsoft Visual C++ 6.0 IDE. All the results presented here were achieved by an initial seed for the pseudo-random number generator set to 225.