Robots with Self-Developing Brains

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Simon Igal Bovet

von
Fleurier (NE)

Promotionskomitee

Prof. Dr. Rolf Pfeifer
Prof. Dr. Yasuo Kuniyoshi
Prof. Dr. Luc Steels

Zürich, 2007

Die vorliegende Arbeit wurde von der Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich im Herbstsemester 2007 als Dissertation angenommen.

Promotionskomitee:	Prof. Dr. Rolf Pfeifer (Vorsitz und Leitung der Dissertation)
	Prof. Dr. Yasuo Kuniyoshi
	Prof. Dr. Luc Steels

Abstract

This thesis aims at contributing to a deeper understanding of natural organisms in which we observe an amazing behavioral diversity and a surprising adaptivity, as well as to the crystallization of principles underlying behaviors that are commonly regarded as intelligent.

The work of this thesis is based on two core assumptions. (a) Behavior is a process involving brain, body and environment as a whole. Indeed, brain and body are inseparable constituents that always have evolved together, and that interact in subtle harmony: intelligence cannot be reduced to properties of the brain in isolation. (b) The purpose of natural behaviors only exists in the eyes of the observer. Organisms are products of natural evolution, a process without goal or intention. The functions that we, as observers, may attribute to them have thus not been explicitly evolved, but are rather side effects that emerge out a complex, self-organizing dynamics resulting from the drift of natural evolution.

This thesis explores behaviors produced by embodied artificial systems – i.e., robots – endowed with a generic, self-developing neural architecture. By showing how the reciprocal interaction between physical and neural dynamics can lead to a variety of behaviors, ranging from insect navigation strategies to predictive reafference cancellation to solving a memory task in a maze, this conceptual approach sheds new, sometimes even unexpected light on the possible mechanisms underlying different natural behaviors, as well as on preconceptions that bias and restrain our view of the world around us.

Résumé

Cette thèse vise à la contribution d'une meilleure compréhension des organismes naturels, chez qui l'on observe une fabuleuse diversité comportementale et une surprenante adaptivité, ainsi qu'à la cristallisation des principes à la base des comportements que l'on qualifie communément d'intelligent.

Le travail de cette thèse est fondé sur deux hypothèses essentielles. (a) Tout comportement met à la fois en jeu le cerveau, le corps et l'environnement d'un organisme. En effet, cerveau et corps ont toujours évolué conjointement, constituant ainsi un tout indissociable qui interagit en une subtile harmonie. L'intelligence ne peut donc être réduite exclusivement à des propriétés du cerveau. (b) Le but de tout comportement naturel n'existe que dans les yeux de l'observateur. Les organismes sont des produits de l'évolution naturelle, un processus sans aucun but ni intention. Les fonctionnalités que nous, observateurs, pouvons leur attribuer n'ont pas été explicitement développées en tant que telles, mais au contraire émergent d'une complexe dynamique auto-organisationnelle produite par la dérive de l'évolution naturelle.

Cette thèse entame ainsi l'exploration de comportements produits par des robots dotés d'une architecture neuronale générique, dont la structure va pouvoir se développer spontanément. En illustrant la façon dont l'interaction réciproque entre la dynamique physique et la dynamique neuronale peut engendrer toute une série de comportements, allant de stratégies de navigation observées chez les insectes, à la suppression prédictive de réafference, en passant par la résolution d'une tâche mémoire dans un labyrinthe, cette approche conceptuelle permet de faire la lumière, parfois de façon inattendue, sur des mécanismes se trouvant potentiellement derrière différents comportements naturels, ainsi que sur les préconceptions qui biaisent et restreignent la vue du monde autour de nous.

Zusammenfassung

Diese Dissertation zielt sowohl auf ein tieferes Verständnis natürlicher Organismen, bei welchen eine erstaunliche Vielfalt von Verhalten und Anpassungsvermögen beobachtet werden, als auch auf die Kristallisierung derjenigen Prinzipien hin, welche den gemeinhin als intelligent angesehenen Verhaltensmustern zugrunde liegen.

Zwei Kernaussagen charakterisieren diese Arbeit: (a) Jedes Verhalten ist ein Prozess, der das Gehirn, den Körper und die Umgebung als Ganzes einschliesst. In der Tat haben sich Gehirn und Körper immer zusammen entwickelt, und bilden somit eine untrennbare Gesamtheit: Intelligenz kann nicht alleine auf Eigenschaften des Gehirns reduziert werden. (b) Der Zweck von natürlichen Verhalten existiert nur in den Augen des Beobachters. Organismen sind Produkte der natürlichen Evolution, ein Prozess ohne Ziel oder Absicht. Die Funktionen, welche wir ihnen aus einer Beobachterperspektive zuordnen können, sind nicht explizit entwickelt worden, sondern Nebeneffekte einer komplexen, selbstentwickelnden Dynamik, welche von der Drift der natürlichen Evolution stammt.

In dieser Dissertation werden Verhaltensmuster von Robotern untersucht, welche mit einer generischen, selbstentwickelnden neuronalen Architektur ausgerüstet sind. Es wird gezeigt, wie die Kopplung von physischer und neuronaler Dynamik eine Vielfalt von Verhaltensmustern erzeugen kann, wie z.B. Navigationsstrategien von Insekten, prädiktive Reafferenzkompensation oder das Lösen einer Gedächtnisaufgabe in einem Labyrinth.

Das Hauptergebnis ist die Erkenntnis, dass der gewöhnliche, auf elementaren Funktionsprinzipien (wie Reflexe, Regelung oder explizites Gedächtnis) basierte Ansatz zu konzeptionellen Schwierigkeiten führen kann, welche hingegen bei einem Ansatz, der auf der selbstorganisierten Entwicklung der Interaktion zwischen Agent und Umgebung basiert, nicht auftreten.

Acknowledgments

First of all, I would like to thank my supervisor, Rolf Pfeifer, for his enthusiasm, for his support, and for providing me with a unique environment where I could explore the scientific areas that fascinated me the most. I'd also like to thank all my colleagues of the AI Lab for the many productive discussions and collaborations, as well as for the great time we spent together. I am further indebted to Yasuo Kuniyoshi and Luc Steels for kindly reviewing this thesis and for providing valuable and insightful comments. Finally, I would like to express my gratitude to all members of my family for their unconditional love and support.

This research has been supported by the IST-2000-28127 European project (AMOUSE) and by the IST-2004-004370 European project (ROBOTCUB).

Contents

Chapter 1

Introduction

The richness, the diversity and the adaptivity of behaviors observed in natural organisms constitute an unceasing source of admiration, fascination and surprise. Understanding how these phenomena come about in the world around us has become a major and challenging topic in many scientific areas.

When the digital computer was invented more than half a century ago, the focus was on the brain, commonly regarded as the "seat of intelligence." The analogy between brain and computer rapidly emerged, and was soon accepted unanimously: computers were called "electronic brains".

This comparison, which still persists nowadays, has at least two most unfortunate consequences. The first one lies at the core of the traditional *computational* view of behavior: the brain is considered as a central system, distinct from the body, that processes information retrieved from the sensors, and controls the organism in which it is embedded. In the mid-1980s, researchers from various areas began to realize that this computational view was not only inappropriate, but also quite misleading. An exciting new scientific field has grown around this realization, known as *embodied* cognitive science or *embodied* artificial intelligence.

The second unfortunate consequence caused by the analogy between brain and computer, which we will refer to as the *purposive* view of behavior, is somehow more subtle and has therefore received much less attention. In this view, natural organisms are seen, even implicitly, as systems *designed* to produce particular behaviors – similar to computers programmed to perform particular operations. However, natural organisms are products of natural evolution, a process without goal or intention. They are thus not purposively designed to produce the behaviors we observe. On the contrary, behaviors emerge – in the eye of the observer – out of a complex, self-organizing dynamics resulting from the drift of natural evolution.

By deepening the realization of these consequences, this thesis aims at increasing our comprehension of adaptive, seemingly intelligent behavior. On the one hand, the work of this thesis is firmly rooted in the field of embodied artificial intelligence: using different robots as embodied autonomous agents, and following the synthetic methodology of "understanding by building," we acknowledge the fact that intelligence needs a body to manifest itself through behavior. On the other hand, this thesis explores a novel and complementary approach, which investigates how interesting behaviors can be observed with artificial systems not explicitly designed for any particular task, and how concrete principles can be drawn from the obtained behaviors. Inspired by canonical findings about the brain, we propose a self-developing neural architecture, and show that the interaction between neural and physical dynamics can produce various coherent, seemingly goal-directed behaviors. In particular, the results obtained illustrate how the proposed conceptual approach can shed new, sometimes even surprising light on the possible underlying mechanisms of different natural behaviors.

1.1 Embodied Artificial Intelligence

Over the last decades, the research field of artificial intelligence (AI) has undergone a profound paradigm shift. The discipline of AI was born at the Dartmouth conference in the summer of 1956, soon after the advent of digital computers. Drawing an analogy between human thinking and processes taking place in a computer, researchers started building artificial systems with the professed goal of emulating, equalling or even surpassing human mental capabilities. It became natural to think of the human brain as an information processing device that receives input from the environment (perception), processes that information (thinking), and acts upon the decision reached (action). The hope was that intelligence could be achieved by sufficiently complex computation and rule-based manipulation of abstract symbols.

Three decades later, it became clear that the approach of the first generation of AI visionaries was intrinsically flawed. Since then, researchers from various areas – including artificial intelligence, computer science, brain and cognitive science, and psychology – have begun to realize that this traditional *computational* approach to intelligence is based on dubious assumptions and has lead to many severe misconceptions. The brain does not run "programs". It does something entirely different.

It is known from evolutionary theory that the body and the nervous system of organisms have evolved together. They are inseparable constituents that interact, in the environment, in complex and subtle harmony. Intelligence thus cannot be reduced to properties of the brain in isolation, but is something that manifests itself in behavior. What we must therefore understand is behavior, a process involving body, nervous system and environment as a whole.

Embodied artificial intelligence – or embodied cognitive science – is an interdisciplinary field that has been developing around the central role of embodiment in understanding natural intelligence and building artificial intelligence. This approach views intelligent behavior as a dynamic and reciprocal interaction between an agent's body, brain and environment. In this sense, the term embodiment not only carries the trivial meaning that "intelligence requires a body," but also places emphasis on the importance of various factors alien to the computational perspective. These factors include the morphology and the material properties of the body, the physical and social interaction with the surrounding environment, or various time scales at which phenomena are observed, such as the developmental (ontogenetic) or the evolutionary (phylogenetic) time scales. Interestingly, this approach has lead, surprisingly rapidly, to a radical rethinking of many of the old and comfortable ideas about the nature of intelligence (Braitenberg, 1984; Varela et al., 1991; Brooks, 1991, 1999; Edelman, 1992; Thompson, 1996; Pfeifer and Scheier, 1999a; Dickinson, 2000; Lipson and Pollack, 2000; Lakoff and Núñez, 2000; Nolfi and Floreano, 2000; Iida et al., 2004; Collins et al., 2005; Dautenhahn, 2007).

Embodied AI is also characterized by its synthetic methodology of "understanding by building". The synthetic methodology consists in creating artificial systems – robots, typically – that reproduce certain aspects of natural systems. This approach not only allows verifying unwarranted assumptions, but also capturing general principles. This point can be metaphorically illustrated by the history of understanding how birds can fly. The essence of flight was long thought to be the mere presence of wings. This incomplete explanation lasted until people started constructing artificial flying systems – thus realizing that other factors, such as the particular profile of the wings or the weight distribution, are of equal importance. It was only after artificial systems were built that the general principles could be identified and abstracted into a theory of aerodynamics.

The work of this thesis is firmly rooted in embodied AI. It aims at understanding intelligent-like behaviors by building artificial systems – i.e. robots – interacting in the real, physical world. In particular, the series of experiments presented in this dissertation illustrate many of the insights developed in recent years within the field of embodied cognitive science (Pfeifer, 1996; Brooks, 1999; Asada et al., 2001; Webb, 2002; Kuniyoshi et al., 2004; Lungarella, 2004; Lichtensteiger, 2005; Pfeifer et al., 2005; Ishiguro et al., 2006; Cruse et al., 2007), and thus contribute to an increased understanding of the general principles underlying intelligent behavior.

1.2 Self-Organization without Purpose

Trying to recognize and get rid of the preconceptions or idées fixes that bias our view of the world around us, and that restrain our understanding of it, is a challenging but also potentially very rewarding endeavor.

The computer metaphor of the brain, or of human thinking, has been for a long time the prevailing view of intelligence. Retrospectively, this analogy looks like an aberrant paradox: wasn't the computer precisely invented to do what the human *cannot* do – numerous, fast and accurate arithmetical operations?

The previous section described the most significant preconceptions that resulted from this analogy, namely the computational view of intelligence. Yet, the computer metaphor carries another set of preconceptions of which we are less aware, which we will refer to as the *purposive* view of behavior, and discuss in the following paragraphs.

The purpose of behavior

When we observe the behavior of natural organisms, it is difficult not to see a goal or a purpose to it – such as maintaining critical internal variables within viable limits, avoiding potential dangers or foraging for food. This tendency reflects itself in artificial models under various, more or less explicit forms, ranging from specific behavioral primitives (e.g. reflexes, or motor programs) to general regulatory or motivational principles.

However, couldn't the behaviors we observe be the result of a dynamical process without any purpose – such as self-organizing phenomena – and to which we attribute a goal as we observe them? After all, isn't even the fact that natural organisms behave adaptively a *side effect* of natural evolution, rather than its reason?

The purposive view of behavior, when we think about it, is surprisingly similar to the concept of computer. Indeed, a computer cannot simply "be there": it must be programmed to do something, to work in a particular, desired way. Isn't then our difficulty not to see an intrinsic purpose in the behaviors we observe yet another consequence of the computer metaphor?

Let us be clear: it is self-evident that organisms display various regulatory, aversive or exploratory behaviors so that they can survive; the point is whether they have been *expressly* created for that. The somehow subtle difference between having a purpose and *appearing* to have a purpose can be illustrated as follows. Reflexes are commonly accepted as behavioral primitives with clearly identifiable functionality – similar to subroutines of a computer program – that constitute basic elements for more complex behaviors. However, when artificial systems are built with this view, the combination of several reflexes turns out, often relatively soon, to be problematic: reflexes start to interfere with each other, and other additional, potentially much more complex means are required to solve the conflictual situation. In contrast – i.e. from a non-purposive approach – a natural or artificial organism is seen as being governed by an arbitrary, self-organizing dynamics, out of which reflex-like behaviors can be observed. The difference is that no additional, complicated regulatory means are required anymore. In other words, this means that a preconceived approach can potentially lead to artificial, unnecessary complications – and thus to ungrounded conceptual difficulties in understanding natural behaviors. (This point will be discussed in detail in Chapter 5.)

Achieving arbitrary behaviors

Another point is the tacit assumption, when artificial systems are developed as models of natural organisms, that it is possible to build artificial systems achieving any desired behavior; that by finding appropriate conditions for artificial agents (including their morphology, neural system, environment, or rules for learning, development or evolution), it is possible to achieve, in a stepwise fashion, behaviors with arbitrarily increasing complexity – and thus any behavior.

We must however be careful not to confuse the possibility of achieving behaviors of arbitrary complexity with the possibility of achieving *any* behavior. For instance, we know the importance of several critical factors underlying biological behaviors (e.g. developmental processes or social interactions, such as offspring rearing), which simply cannot be artificially reproduced with today's technology.

Isn't this illusion of being able (with sufficient engineering efforts) to reproduce any desired behavior oddly reminiscent of the common view of computers as machines that can perform (with sufficient programming efforts) any arbitrary computation? Couldn't this illusion be a consequence of the circular and often over-interpreted "universality" attributed to the computational power of Turing machines – the theoretical concept of a computer? Acquainted with the ever-growing number of computer applications in our everyday life, we tend to easily forget that even Turing machines, despite being "universal," have definite limits: there are even simple and well-defined computations that simply cannot be performed by any Turing machine.

A fertile historical background

It is probably pertinent to note that this purposive view of behavior is not only due to the computer metaphor, but has also found a fertile historical background. For a long time, people from many cultures have been indeed fascinated by the idea of creating intelligent machines – robots – that could replicate aspects of behaviors displayed by humans and animals in order to accomplish all sorts of useful, desired tasks.

This idea of intelligent machines as useful artifacts dates at least as far back as the ancient Greek mythology, which includes the legend of Kadmos, who sowed dragon teeth that turned into soldiers, or the telling of Talos, a bronze automaton guarding the shore of Crete. The 8th century Persian alchemist Jabir ibn Hayyan, inventor of many basic processes still used in chemistry today, included in his coded "Book of Stones" several recipes for creating artificial creatures subject to the control of their creator. Traditional Jewish mysticism contains narratives about golems, animated beings created from inanimate matter, such as the famous legend involving the Maharal of Prague said to have created a golem with clay from the banks of the Moldau to defend the ghetto of Prague.

The first historical attempts to create artificial "intelligent" machines also clearly reveal this tight association made between intelligent machines and purposeful artifacts. One of the first designs on record of a humanoid robot was made by Leonardo da Vinci in around 1495. Da Vinci's notebooks, rediscovered in the 1950s, contain detailed drawings of a mechanical knight able to make some basic gestures (Rosheim, 1996). The first true robots are usually credited – at least in the Western culture – to the French engineer Jacques de Vaucanson, who built in the 1730s various musician androids and a digesting duck. Other examples include the writing, drawing, and organ-playing doll automata built around 1770 by the Swiss watchmakers Pierre and Henri-Louis Jaquet-Droz, as well as the Japanese karakuri ningyô, the traditional craft of building mechanised entertainment puppets in the 18th and 19th century.

Even today, the same idea can still be noticeably felt. For a layman, an intelligent robot is typically conceived as a machine that does something useful, or at least something that makes sense for us. Who hasn't heard the usual question whether an intelligent robot couldn't be built to clean someone's kitchen, or to perform some similar, precisely defined tasks? Clearly, intelligent behavior is still commonly seen from a purposive view: intelligent behavior must have a purpose, and this purpose can be any arbitrary desired task. In contrast to this view, however, is the following question: doesn't an "intelligent" machine, specifically built for a particular task, only reflect the intelligence of the designer?[1]

---

[1] Even in various approaches where agents are only told *what* to do but not *how* to do it (such as self-supervised or reinforcement learning, evolutionary robotics, etc.), it is well known that the outcome critically

Self-organization without purpose

Self-organization – a phenomenon which the research of this thesis is based on – as well as its importance in the world around us has been recognized already since long ago (Ashby, 1947; Nicolis and Prigogine, 1977; Deneubourg et al., 1991; Kauffman, 1993; Langton, 1995). Nowadays, the term "self-organization" is ubiquitous in the scientific literature, especially in research fields such as dynamical systems and artificial life. Self-organization is a property of systems composed of many interacting sub-systems where ordered patterns emerge at the macroscopic scale that can not be directly deduced from the microscopic properties of the sub-systems, i.e. where order is produced without centralized activity.

Even though the view of natural behaviors as self-organizing phenomena has established itself during the last decades (Maturana and Varela, 1992; Thelen and Smith, 1994; Kelso, 1995; Kaneko and Tsuda, 2001; Camazine et al., 2001), it is interesting to observe that, in contrast, the synthetic approach to the understanding of intelligent behaviors has only rarely been able to depart from the purposive view of behavior described so far.

On the one hand, some of the earliest steps in synthetic modeling have been investigating behaviors produced with artificial systems governed by arbitrary dynamics. Gray Walter's *Machina speculatrix* (1950), or the first series of agents introduced by the neuroscientist Valentino Braitenberg, the well-known "Braitenberg vehicles" (1984), are examples that illustrate how surprisingly complex, seemingly purposeful behaviors can emerge from systems composed of a few, arbitrarily interacting elements – such as a few wires randomly connecting the sensors to the motors in the case of Braitenberg's vehicles[2]. However, these early approaches were rapidly abandoned in favor of more purposive ones, where the arbitrary dynamics of the artificial system is gradually replaced by elements with precisely defined functionality: Gray Walter called his second-generation machine *docilis* (1951) – a name that reveals the underlying intention of having a machine that can learn any behavior the designer want it to; and Braitenberg's agents, starting with vehicle 5, began to incorporate elements with explicit purposes, such mechanisms for logic operations, shape detection, or prediction.

More recently, in the research field of developmental robotics, there has been increasing efforts to explore artificial systems without explicit behaviors (Almassy et al., 1998; Weng et al., 2001; Andry et al., 2002; Lungarella, 2004; Steels, 2004). However, developmental processes are still generally driven by motivational principles – e.g. imitation (Dautenhahn

---

depends on pertinent choices of the human designer (see e.g. Matarić and Cliff, 1996).

[2] Another example to the point is the experiment of Maris and te Boekhorst (1996) that shows how the interaction of agents with an arbitrary neural dynamics – as a matter of fact, Braitenberg vehicles – can lead to a collective, self-organized heap building process.

and Nehaniv, 2002), or maximization of some internal reward (Kaplan and Oudeyer, 2004) – that confer a purpose to the self-organization process, and so an explicit raison d'être to the artificial system.

There seems thus to be an apparent paradox in the synthetic methodology. The understanding of general principles underlying natural behaviors manifest itself – as discussed earlier in this chapter – in the ability to apply these principles for building useful, functional artifacts. Yet it is precisely this desire to build functional artifacts that biases and potentially constrains the understanding of these very same principles.

There are however other lines of research that explore the self-organization of behaviors produced without motivational drive, and which thus constitute interesting exceptions to the somehow simplified landscape described so far. For instance, Ishiguro et al. (2004, 2006) have been investigating how "emergent phenomena" stemming from the interaction between control and mechanical dynamics can be exploited to control the morphology of a modular robot and to generate locomotion. To this end, they focus on functional material, mutual entrainment, and non-linear oscillators.

Also, Kuniyoshi and colleagues (Kuniyoshi et al., 2004; Kuniyoshi and Suzuki, 2004; Kuniyoshi and Sangawa, 2006) have been exploring and studying the emergence of meaningful behaviors without predefined coordinated control – i.e. without any motivational drive such as an evaluation function. Their work is based on simulated musculo-skeletal systems driven by a number of chaotic elements interacting with each other through the physical body and the environment of the system. They showed how the rich intrinsic dynamics and entrainment properties of chaotic elements can be exploited to produce autonomous exploration and rapid adaptation of motor behaviors without motor primitives.

1.3 A Complementary Approach

This thesis intends to explore artificial systems endowed with a self-developing dynamics, yet without any particular task or motivation, and to study the general principles that can be drawn from the observed behaviors.

It is important to understand that this methodology aspires first and foremost to offer a complementary approach to the currently existing study of intelligent behaviors. It is undeniable that enormous progress has been achieved with artificial models designed to reproduce specific behaviors; that focusing on a particular phenomenon allows better appreciating its intricacy; that building an artificial system, even though the principles used to construct it are not necessarily valid for natural systems, provides a extremely valuable source of insight.

Nevertheless, it can also be beneficial – as this thesis will demonstrate – to take a step back and to try better understanding the preconceptions that bias and restrain our view of the world around us. By exploring embodied systems from a slightly different perspective, it becomes possible to discover new – and often simpler – principles underlying familiar behaviors, or to discern common properties shared by a priori unrelated phenomena.

The analogy to the understanding of flight can again illustrate our point. By focusing too much on a behavior observed in particular natural organisms (how birds can fly), and by trying to reproduce only a given, clearly useful aspect of it (making a human fly), it is difficult – if not impossible – to make progress. Yet, by stepping back and looking at things from a somewhat different angle (exploring what can fly – even if it is a priori useless objects), it is possible to realize enormous progress, both conceptually (understanding the principles of aerodynamics) and technologically (building airplanes, with all the engineering, economical and social consequences that follow). What is particularly interesting in this process is the fact that the phenomenon in question can be understood in great details even though the original goals (making someone fly autonomously) are not, and might never be achieved – as if the understanding of a phenomenon is tightly coupled with the appreciation and acceptance of its limits.

This step back towards a broader perspective on our preconceptions, this interrogation whether things can be different from what we think they are: these are the contributions the work of this thesis is striving for.

1.4 Outline and Contributions

The following Chapter 2 introduces a minimal neural architecture for embodied agents, based on systematic sensorimotor coupling with Hebbian plasticity, which has the potential of generating non-trivial behaviors. The potential of this arbitrary generic neural architecture is then explored during the rest of the thesis, where the various behaviors produced with five different robotic setups are consecutively investigated in each of the successive chapters that follow.[3]

Chapter 3 describes a first series of experiments showing that the interaction between the neural dynamics of the network and the physical dynamics of an embodied agent can produce various seemingly goal-directed behaviors.

The next Chapter 4 presents a series of experiments, still using the same neural architecture, inspired by studies on insect navigation. We show that the behaviors observed on

---

[3] Chapters 2, 3 and 7 are based on the work published in (Bovet and Pfeifer, 2005a,b). The main results of Chapter 4 were published in (Bovet, 2006).

a robot situated in a desert-like environment closely match two different homing strategies observed in natural desert ants – path integration and visual landmark homing.

Chapter 5 investigates the role played by the homogeneous connectivity of the neural network. Through a series of experiments inspired by research on cricket phonotaxis and by the theoretical framework of forward models, we show that neural activity flowing through the various synaptic coupling of the neural network produces both reflex-like behaviors as well as predictive reafference cancellation. The results thus suggest that reflexes and forward models do not necessarily require distinct mechanisms, as they can potentially be produced by the same underlying neural structures.

A fourth series of experiment with the proposed neural architecture is then conduced in Chapter 6, showing that it can successfully be applied on a robot with a more complex body dynamics – a four-legged running robotic dog – to generate coherent behavior. In particular, we observe that the robot spontaneously discovers a motor strategy for turning that is more robust than the approach usually taken to control the rotation of legged robots.

Chapter 7 describes a last series of experiments that further explores the potential of the neural architecture with a robot endowed with more sensorimotor modalities and situated in a more dynamic environment. Inspired by maze experiments used to study navigation and learning in rodents, we choose to engage the robot in a T-maze memory task. We not only show that the robot solves this delayed reward learning task, but also that the agent does not possess any explicit working memory of previous events and actions. By illustrating how the memory can be off-loaded into the interaction with task-irrelevant structures of the environment, this chapter provides a case study shedding new light on the neural basis of memory and supporting recent concepts in memory research developed as alternatives to the common but problematic view of memory as stored information.

Finally, a summarizing discussion of the general principles drawn from the results obtained throughout the thesis is provided in the concluding Chapter 8.

Chapter 2

Neural Architecture

This chapter introduces a model of self-developing neural architecture for embodied agents, which will be used throughout the thesis. The basic idea is to systematically and reciprocally couple all sensors and motors the agent is equipped with through synaptic coupling with Hebbian-like plasticity, and then observe the behaviors that are produced as the agent interacts with the environment.

2.1 Introduction

With this chapter, we enter the core of this thesis, and begin our exploration of noteworthy behaviors displayed by "robots with self-developing brains."

2.1.1 Motivation

As outlined in the preceding introduction, the work presented here is about better understanding, through the synthetic methodology, the behaviors of natural organisms – i.e. products of natural evolution, a process without goal or intention.

Natural evolution does not use genes to explicitly program organisms to produce particular behaviors. Rather, natural evolution is a process whose influence on certain parameters *within a whole self-organizing dynamics* produces organisms displaying various, seemingly adapted or even "intelligent" behaviors.

This motivates us to explore artificial systems endowed with an arbitrary, self-organizing and simplest possible dynamics[1] – yet at the same time without any particular goal, pur-

---

[1] The term *dynamics* is used in this thesis to refer to the behavior of a particular sub-system of the agent, such as in "body dynamics" – how the different parts of the body interact with each other – or in "neural dynamics" – how neurons and synapses influence each other. This allows us to dedicate the term *behavior* to the aspects of the whole system visible to an external observer.

pose or motivation. We decide, inevitably constrained by the technology currently available, to conduct experiments with robotic platforms, and to equip those with a generic, self-developing neural architecture.

The essence of this neural architecture, which will be described in detail in this chapter, is the following. 1) All signals of the sensors and motors the robot is equipped with are represented through the activity of artificial neurons. 2) All populations of artificial neurons are homogeneously coupled to each other through artificial synapses, whose plasticity follows a simple rule well-known to biologists: "neurons that fire together wire together".

Coupling the neural architecture to an embodied system will inevitably produce correlated neural activity, which will shape – through the plasticity of the synapses – the connectivity of the network. Neural activity can thus propagate between the different populations of neurons. Whenever motor neurons get activated, the corresponding physical motors of the robot are set in motion, therefore producing an observable behavior of the robot. It is these behaviors, produced by the self-developing neural architecture and observed under various conditions, which this thesis is going to explore.

From this starting point, one of the simplest possible architectures consists in each sensory or motor signal being represented by the activity of one neuron, and in all neurons being fully connected to each other (Section 2.2.1). Unfortunately, it turns out – maybe not so surprisingly – that this first implementation does not lead to any interesting behavior. This chapter will discuss how a more stable dynamics – and thus more interesting behaviors – can be achieved, without abandoning the essence described so far, by modifying as little as possible 1) the way sensory and motor signals are represented internally, and 2) the way neural populations are coupled to each other (Sections 2.2.2 and 2.2.3).

The use of a Hebbian-like synaptic plasticity is not quite fortuitous. First, one ultimate goal of this work is – as mentioned before – to learn something about natural organisms. It is thus not surprising to see some definite biological inspiration. Second, this work stems from the following question: what is Hebbian learning good for? In contrast to other learning rules for neural networks whose functionality is clearly defined, but whose biological plausibility is fairly doubtful, what is the role of this Hebbian synaptic plasticity observed in natural nervous systems? The experiments that we will describe in the following chapters will show that Hebbian plasticity is not good *for* something, but rather allows – by picking up correlations of neural activity – a rich and interesting dynamics to take place between the sensors and motors of the robot, its interaction with the environment, as well as its own behavior.

Before concluding the introduction of this chapter, we have to be aware of an arbitrary bias that characterizes our approach: the very existence of an artificial nervous system distinct from the physical components of the robot. We do not want to debate on what distinguishes the brain from the body – this would exceed the scope of this thesis – but just mention that the artificial neural system is nothing else than a means – such as the set of biological cells with distinctive characteristics – allowing distant elements of an agent to interact more directly with each other (such as receptors sensitive to physical properties at certain locations, and actuators modifying the interaction with the environment at other, perhaps partially overlapping locations).

Finally, being interested in the interaction of agents – both natural and artificial – with the real, physical world, we have chosen to study most of all systems that are physically instantiated, rather than virtual systems in simulations. As we will see in the following chapters, the richness in the physics of the real world will turn out to be an essential factor for the generation of interesting behaviors, a factor whose role could have probably not been replaced by simulations.

2.1.2 Comparison with Similar Approaches

The neural architecture introduced in this chapter is based on the early work of Braitenberg (1984), who initiated the exploration of behaviors produced by agents endowed with very simple "brains." As in his early vehicles, sensors and motors of the agent are represented by artificial neurons that are directly coupled to each other through artificial synapses. But instead of using a fixed network connectivity, which after all partly relies on some arbitrary assumptions such as connecting the sensors to the motors rather than vice-versa, we want to explore how the network connectivity can develop spontaneously (using what Braitenberg called "Mnemotrix wires," i.e. artificial synapses with Hebbian plasticity).

By exploring embodied systems with some arbitrary neural dynamics – i.e. without explicit control or motivational drive – our approach coincides with the line of research followed by Ishiguro et al. (2006) and Kuniyoshi and Sangawa (2006) introduced in the previous chapter (see page 8). Despite some differences at the implementation level (we will use linear neurons with Hebbian plasticity, as opposed to non-linear or chaotic elements; also, the richness in the neural dynamics will originate in our case mostly from the coupling of multimodal activity, rather than from the intrinsic dynamics and entrainment properties of non-linear coupled oscillators), the work of our thesis bears much similarity with these approaches, all exploring how the embodied interaction between physical and neural systems can spontaneously produce interesting behavior.

On the other hand, something that distinguishes our approach is the conceptual goal it is aiming at. Less trying to explore – using a particular model, such as coupled non-linear oscillators – how emergent phenomena stemming from the interaction between a complex neural dynamics and the physical dynamics of an embodied system can be exploited for the autonomous generation of behavioral primitives, our approaches strives first and foremost to investigate – with systems endowed with an arbitrary, self-developing dynamics – the biases and preconceptions that restrain our view of intelligence.

2.2 The Model

2.2.1 Coupling All Sensors and Motors to Each Other

Let us consider an autonomous agent – a robot – equipped with various sensors and motors. The first property of the proposed neural architecture is that the sensory and motor signals are all internally represented by artificial neurons[2], i.e. simple units whose activities reflect the physical characteristics either measured by the sensors or produced by the motors.

In the first version of the model, each single sensory or motor signal is simply represented by one neuron (see Figure 2-1). For instance, a robot equipped with a camera will have a population of visual neurons representing the brightness of each pixel in the camera image. The same schema applies for the motors as well.

The second property of the neural architecture that needs to specify is its connectivity. In order to reduce the designer commitments as much as possible, we opt for a homogeneous connectivity in the network, where all sensory and motor modalities are systematically coupled to each other, irrespective of the physical devices they represent (see Figure 2-2). There is no self-recurrent connections (the reason is discussed below). Each connection is a simple artificial synapse, whose connection strength – initially set to zero – is modified by a simple Hebbian (i.e. correlation-based) learning rule: whenever two neurons connected by a synapse are active at the same time, the connection strength of the synapse is increased.

The working principle of the neural network is as follows. Initially, no behavior is observed. Indeed, the different population of neurons do not interact with each other since all synaptic strengths are set to zero: there are, in other words, no built-in behavioral primitives. Consequently, the agent is first let to explore its interaction with the environment. This is achieved, during an initial exploration phase, by randomly activating the motor neurons, and thus producing random motions of the agent in its environment. As the robot moves and interacts with its environment, sensory and motor neurons will be activated, some of them at the same time. Because of the Hebbian learning rule, the neural network

---

[2] In this thesis, all artificial neurons are real-valued and have a linear activation function.

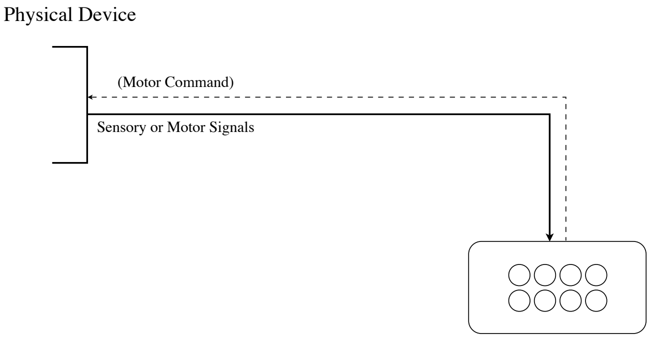
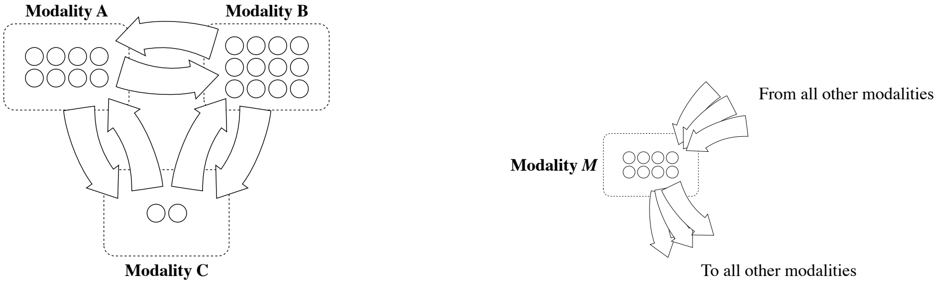


Figure 2-1: In the first version of the neural architecture, each sensory or motor modality is represented by a population of artificial neurons whose activities correspond to the individual signals read from the physical device. In motor modalities, the neural activity also generates the motor command activating the physical motor.



(a) (b)

Figure 2-2: Homogeneous connectivity in the first version of the neural architecture. The synaptic couplings are represented by arrows that indicate full connectivity between populations of neurons. (a) All modalities are systematically and homogeneously coupled to each other, irrespective of the physical device (sensor or motor) they represent. (b) All neurons in one modality are fully connected from and to all neurons in all other modalities.

will with time spontaneously develop structures in its connectivity, thus allowing neural activity to propagate from one population of neurons to the other.

Unfortunately, testing this neural architecture on an agent will soon reveal that this model in fact does not lead to any interesting phenomenon. There are at least two main reasons for that. First, sensory and motor activity is in general not directly correlated. Consider for instance a mobile robot equipped with a camera: there is clearly no relation between the motion of the robot and the brightness of the perceived image. The synaptic strengths thus do not change significantly. Second, two neurons recurrently coupled to each other generally produce unavoidable deadlocks: if one neuron is active enough to stimulate the second, the positive feedback created by the bidirectional synaptic coupling will keep both neurons indefinitely active.

In the next section, we describe how these problems can be overcome by slightly modifying this initial model of neural architecture.

2.2.2 States and State Changes

If motor activity is in general not directly correlated to sensory signals, it is often correlated to *temporal variations* of sensory signals. For instance, in the example mentioned previously, we pointed out that there is no correlation between motion of the robot and the brightness of the perceived image. However, any motion of the robot will elicit a corresponding correlated optical flow in the camera image. There is thus a clear correlation between the motion of the robot (i.e. motor activity) and the perceived optical *flow* (i.e. variation of visual sensory activity).

We thus decide, in the second version of the neural architecture, to augment each modality – i.e. each collection of neurons representing the signals of a particular sensor or motor – with a population of neurons representing *changes* in the sensory or motor signals. This is achieved by adding in fact two populations of neurons to the original population representing the current sensory or motor state: one extra population representing again the state, but delayed by a small amount of time – the “delayed state” population; the other extra population representing the state change simply as a *difference* between the delayed and current sensory or motor state – the “state change” population (see Figure 2-3).

Let us turn now to the connectivity of the network. Instead of coupling each pair of sensory or motor modalities by fully connecting all neurons of one modality to all neurons of the other modality – we previously discussed that recurrent connections between two neurons lead to inopportune deadlocks – the different populations of neurons are connected according to the following rule: each population of neurons representing the state change

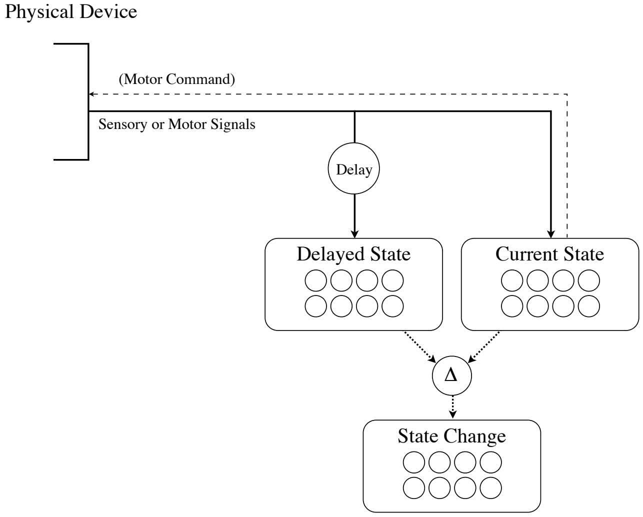
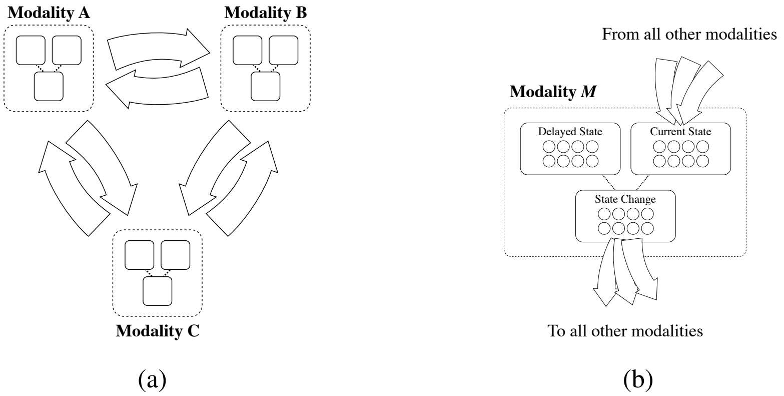


Figure 2-3: In the second version of the neural architecture, each sensory or motor modality is represented by three populations of artificial neurons: the activity of the "current state" population represents the sensorimotor signals from the physical device (as in Figure 2-1); the activity of the "delayed state" population also represents the sensorimotor signals, but with a small time delay; the "state change" population represents variations in the sensorimotor signals as a difference between the activities in the "delayed state" and "current state" populations.



(a) (b)

Figure 2-4: Homogeneous connectivity in the second version of the neural architecture. (a) As previously (see Figure 2-2), all modalities are systematically and homogeneously coupled to each other, irrespective of the physical device they represent. (b) The synaptic couplings from all other modalities are fed into the "current state" population; the synaptic couplings to all other modalities leave the "state change" population. Thus, each "state change" population in the network if fully connected to the "current state" populations of all other modalities.

of one modality is fully connected to the populations of neurons representing the current state of all other modalities (see Figure 2-4). As before, the synaptic strengths – initially set to zero – are modified by a simple Hebbian learning rule, thus capturing cross-modal correlations of neural activity (this time between the state change of one modality and the state of another modality).

The rationale underlying this particular connection scheme is to have a single rule that can be applied indiscriminately to connect each and all pairs of modalities. However, there are a couple of reasons why we decide to stick to this particular choice, rather than explore other possibilities. First of all, this coupling rule is one of the simplest ways to connect each pair of modalities to each other such that neural activity can spread throughout the whole network (see Figure 2-5):[3] activity propagating from a first modality will activate the current state population of a second modality, therefore generating a difference of activity between delayed and current states, and thus inducing activity in the state change population of the second modality, activity which will in turn propagate further into a third modality, and so on. Second, the advantage of connecting neurons indicating changes of sensorimotor signals with neurons indicating steady sensorimotor signals is that it allows a simple Hebbian learning rule to be used without the synapses capturing spurious correlations[4]. Finally, it turns out that this particular choice leads to various interesting observable behaviors when applied on different agents, as will be shown in the next subsection and in

---

[3] Since we want to avoid direct recurrent coupling, we are obviously left with either state → change, or change → state couplings (any other coupling scheme leads to direct recurrent connections). The first possibility will however not allow neural activity to propagate throughout the network, whereas the latter possibility does (see Fig. 2-5). This eventually leaves us with either change → delayed state or change → current state couplings (again, both connections at the same time would obviously prevent activity from propagating in the network). We arbitrarily choose to investigate the latter possibility.

[4] To illustrate this point, let us consider two sensorimotor signals $x(t)$ and $y(t)$ with non-zero mean values, the first one being constant and positive $x(t) = c > 0$ and the second one begin cyclic and positive $y(t) > 0$, $y(t) = y(t + \tau)$. Clearly, these two sensorimotor signals are independent, and thus not correlated.

Let us further consider a simple Hebbian learning rule $\frac{d}{dt}w(t) = \eta \cdot a(t) \cdot b(t)$ modifying the synaptic strength w between two neurons whose activities are denoted by $a(t)$ and $b(t)$, respectively. If the activity of both neurons reflects the above sensorimotor signals, i.e. $a(t) = x(t)$ and $b(t) = y(t)$, then the synaptic connection will capture a positive correlation:

$$w(\tau) = \int_0^\tau \eta \cdot x(t) \cdot y(t) dt > 0$$

However, if the activity of one neuron reflects the *variation* of the corresponding signal, i.e. either $a(t) = \dot{x}(t)$ or $b(t) = \dot{y}(t)$, then the synaptic connection will capture no spurious correlation anymore:

$$w(\tau) = \int_0^\tau \eta \cdot \dot{x}(t) \cdot y(t) dt = \int_0^\tau \eta \cdot x(t) \cdot \dot{y}(t) dt = 0$$

Note that this particular form of correlation-based learning is often referred to as differential Hebbian learning (e.g. Kosko, 1986).

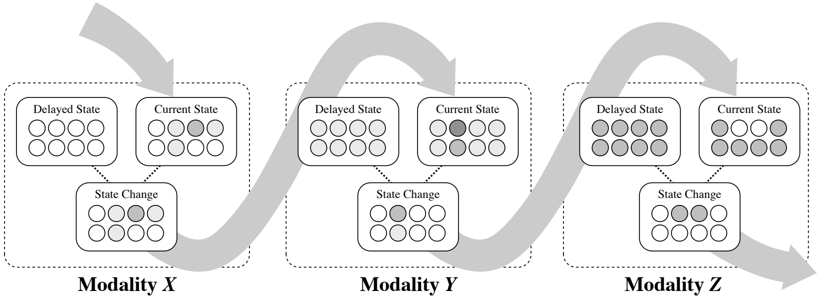


Figure 2-5: Illustration of propagation of neural activity across multiple modalities. Activity is generated in the current state population of the first modality. The difference of activity between the delayed and the current state populations induces activity in the state change population of the first modality, which then propagates into the current state population of the next modality. Neural activity can thus propagates across several modalities. For the sake of clarity, this figure illustrates only a sequential propagation of activity, whereas in the proposed neural architecture, activity can potentially spread throughout the whole network in any direction.

the next chapters of the thesis.

The network connectivity in this version of the neural architecture displays similar properties as the one introduced previously. First, the connectivity is homogeneous: all pairs of modalities are systematically coupled to each other, regardless of the particular sensors or motors they represent. Second, there is no hierarchical structure in the network: all modalities both receive and send synaptic connections from and to all other modalities. Finally, neural activity can potentially spread throughout the whole network, as discussed previously.

Illustrating the Working Principle

To illustrate the working principle of this neural architecture, let us consider the fictive scenario sketched in Figures 2-6 and 2-7. The agent possesses one eye and one arm, which it can move back or forth. The agent is also equipped with a sensor measuring its thirst.

Each time the agent contracts its arm muscle, the object the agent is holding in its hand moves closer to its eye. The agent then perceives simultaneously motor activity and an expanding optical flow corresponding to the growing image of the object on the retina (Figure 2-6(a)). The neural network therefore learns a correlation between expanding optical flow (the state change in one modality) and motor activity (the current state in another modality). Also, whenever the agent is “drinking,” it perceives a decrease of thirst together with the perceived image of an object – say a cup – close to its eye (Figure 2-6(b)). The neural network therefore also learns a correlation between decrease of thirst (the state change in one modality) and a large image of the cup (the current state in another modality).

Let us study what happens if the activity of the neuron representing the current thirst level is artificially decreased (Figure 2-7). The neuron corresponding to the state change in this modality will display an activity corresponding to a decrease of thirst. Because of the correlation learned between decrease of thirst and the image of a close cup, this activity will propagate into the visual modality, projecting a large image of a cup in the current state population (Figure 2-11(a)). Metaphorically speaking, the agent, when "feeling" a decrease of thirst, "sees" a cup close to its eye.

Now, if the agent is holding a cup in its hand, the difference between the delayed visual state (a small image of a cup) and the current visual state (a large image of a cup) will elicit a pattern of activity corresponding to an expanding optical flow. Finally, the latter activity will propagate further into the motor modality – because of the correlation learned between expanding optical flow and motor activity – inducing a contraction of the arm muscle (Figure 2-11(b)).

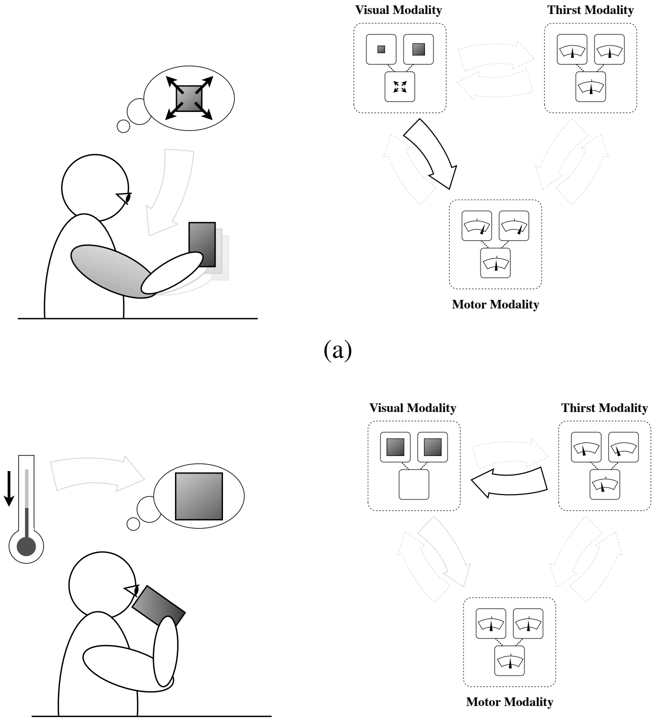
The observed behavior can be summarized as follows. As the agent explore its interaction with the environment, the artificial neural network learns correlations between muscle activity and expanding optical flow, as well as between decrease of thirst and the image of an object close to its eye. Then, whenever the agent "feels" a decrease of thirst and holds a cup in its hand, it will contract is arm muscle to move the cup close to its mouth (Figure 2-11(c)).

This thought experiment illustrates how the proposed neural architecture, despite being homogeneous in its connectivity, can produce a seemingly coherent behavior by picking up different cross-modal correlations. In particular, this experiment illustrates how activity generated in one modality (in our case, making the agent "feel" less thirsty than it actually is) can potentially lead to motor actions that can be described as some kind of "procedural," or seemingly goal-directed behavior (bringing a cup to the mouth as for reducing the thirst).

2.2.3 Parallel Pathway for Propagation of Activity

The neural neural network proposed so far seems to have the potential of generating coherent behaviors. However, a closer inspection of the model reveals some further inherent difficulties.

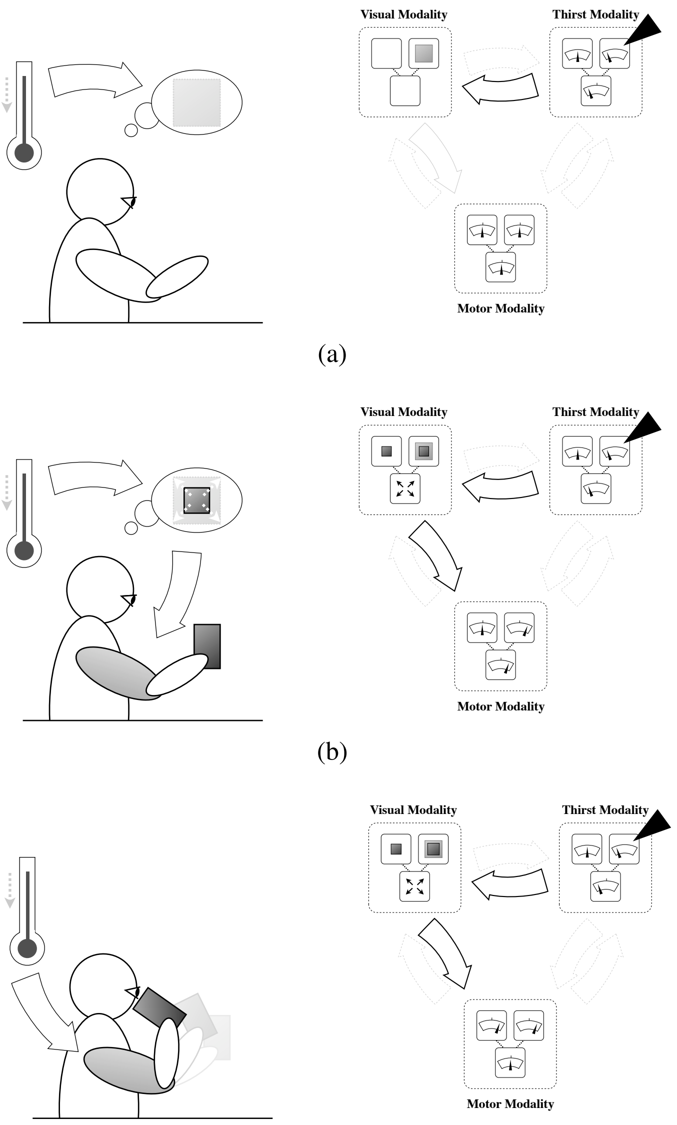
Let us for instance consider the mobile robot with camera that was mentioned earlier. We previously discussed how rotation of the robot (motor activity) is correlated to a uniform lateral optical flow (change of visual activity). If the robot starts turning, the rotation of the body will obviously produce an optical flow. This optical flow will be detected by



(a)

(b)

Figure 2-6: Fictive scenario illustrating the working principle of the proposed neural architecture. The left side illustrates the agent interacting with its environment. The balloon represents the visual modality of the agent, and the thermometer its thirst modality. The right side illustrates the dynamics of the neural architecture. The neural activity in each sensory and motor modality is represented graphically. The outlined arrows indicate the cross-modal correlations learned by the neural architecture. See text for details.

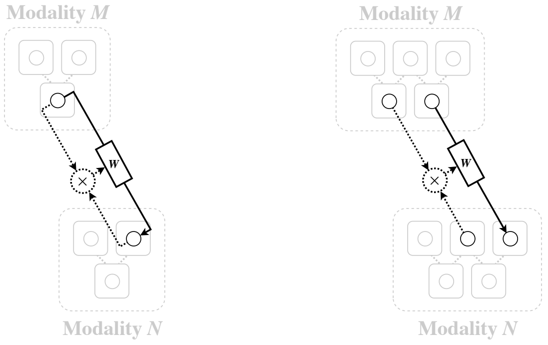


(a)

(b)

(c)

Figure 2-7: Fictive scenario illustrating the working principle of the proposed neural architecture (continued). The black triangle indicates that the activity of neuron representing the current thirst level is artificially decreased. The outlined arrows indicate the propagation of neural activity across the different modalities. See text for details.



(a) (b)

Figure 2-8: The two roles of an artificial synapse. On the one hand, it captures the correlation between the activities of two neurons: the synaptic weight W is modified proportionally to the product of the two activities (Hebbian learning rule). On the other hand, it allows activity to propagate from one neuron into another. (a) Traditionally, an artificial synapse connects two neurons. For instance, in the previous version of the neural architecture, each synapse connected two neurons – one neuron in the state change population with one neuron in the current state population. (b) In the final version of the neural architecture, each synapse connects four neurons, as illustrated in the figure.

the visual population responding to changes in visual input. Now, because of the correlation mentioned just before, this neural activity will propagate from the visual modality back into the motor neurons, inducing an ever-increasing turning speed of the robot. In other words, the coupling between the visual and motor modalities leads to an undesirable positive feedback – this time through the physical interaction of the robot with the environment.

To solve this problem, we propose to modify the neural architecture as follows. The idea is based on the fact that a synapse, coupling two neurons, serves in fact two distinct functions. On the one hand, the synapse captures the correlation of neural activity (through the Hebbian learning rule). On the other hand, the synapse allows neural activity to propagate from one neuron to the other.

So far in the neural architecture, each single synapse connected two neurons from different modalities, simultaneously learning the correlation in their respective activities, as well as allowing one neuron (in the state change population) to activate the other (in the current state population), as illustrated in Figure 2-8(a). Here, the idea is basically to separate those two functions into two distinct parallel pathways, as illustrated by Figure 2-8(b). The first function – namely capturing the correlation of neural activity between the state change population of one modality, and the current state population of another modality – is left unchanged. In contrast, propagation of neural activity now takes place along a parallel pathway between two new populations of neurons.

These two additional populations of neurons, which mirror in each modality the “current state” and “state change” populations, will be referred to as “desired state” and “desired change,” respectively. Note that the adjective “desired” is purely arbitrary (and maybe a bit unfortunate – remember that there is nothing such as desire, intention or motivation!), chosen so only to help the reader better understand the model – albeit from an observer perspective (the reason for this is discussed below).

The way how the five populations of artificial neurons are organized within each sensory and motor modality is given in Figure 2-9. For symmetry, the “desired state” population contains as many neurons as the “current state,” and the “desired change” population reflects the same difference as the “state change” population.

In summary, the sensory and motor signals of each modality are represented in the final version of the neural architecture by the following five populations of neurons:

1. Delayed state population: reflecting the sensorimotor signals with a small delay.
2. Current state population: reflecting the sensorimotor signals with no delay.
3. “Desired” state population: receiving synaptic input from all other modalities.
4. State change population: reflecting the difference between the delayed state and the current state.
5. “Desired” change population: reflecting the difference between the current state and the desired state, and projecting activity through the synaptic connections into all other modalities.

As in the previous versions of the neural architecture, the signals from each sensory and motor modality are represented through the same multi-population structure of artificial neurons. All modalities are systematically and reciprocally coupled to each other, irrespective of the physical devices (motors or sensors) they represent. Also, the weights of the synaptic couplings, initially set to zero, capture the correlations of activity between the state change population in one modality, and the current state population in another modality.

The only difference is that neural activity now propagates between the modalities along a parallel pathway, namely from the desired change populations to the desired state populations (see Figure 2-10). Also, each artificial synapse now involves four neurons – learning the correlation of activity between the first two, and allowing activity to propagate between the last two neurons. Interestingly, it can be shown that this abstract synaptic coupling can

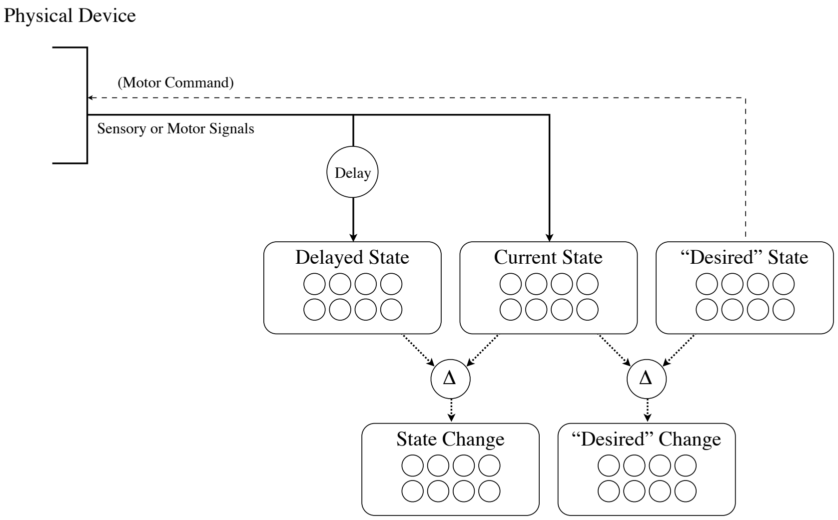
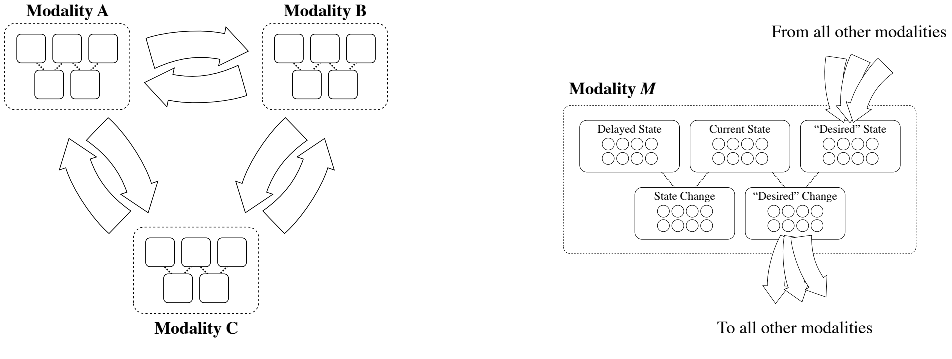


Figure 2-9: In the final version of the neural architecture, each sensory or motor modality is represented by five populations of artificial neurons (compare with Figure 2-3). As before, the "state change" population indicates the difference between the activities in the "delayed state" and "current state" populations. Similarly, the "desired change" population indicates the same difference, but between the activities in the "current state" and "desired state" populations.



(a) (b)

Figure 2-10: Homogeneous connectivity in the final version of the neural network. (a) As previously (see Figure 2-4), all modalities are systematically and homogeneously coupled to each other, irrespective of the physical device they represent. (b) The neural activity coming from all other modalities activates the "desired state" population; the neural activity of the "desired change" population propagates into all other modalities. Note that the synaptic weights still capture as before the cross-modal correlations between activity in the "state change" and "current state" populations (see Figure 2-8(b)).

be approximated by a local network of traditional artificial synapses (see Appendix 2.A on page 35).

Illustrating the Working Principle

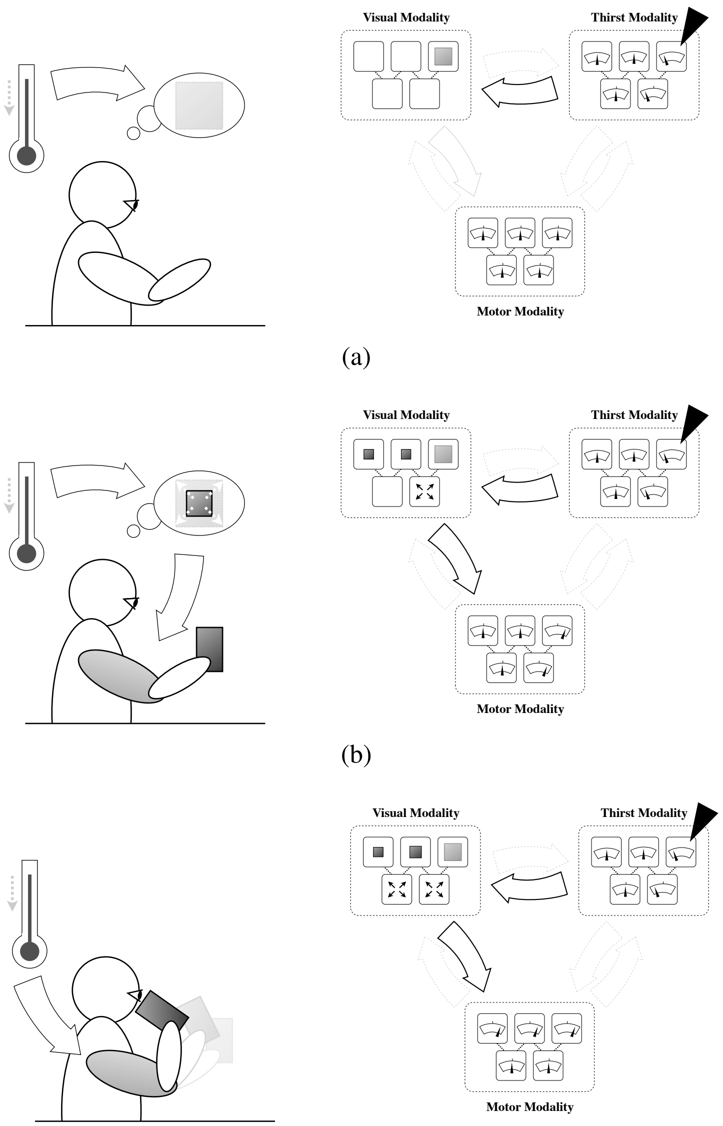
Before giving a formal description of the model, we briefly show how its working principle, despite the modifications, remains nevertheless very similar to the one discussed in the previous section. Let us therefore consider again the fictive agent of Figure 2-7. As argued in the preceding paragraphs, the neural network will capture the same cross-modal correlations as the one described previously. Activity is generated this time by artificially decreasing the desired state in the thirst modality. Figure 2-11 describes how the same coherent, seemingly goal-directed behavior will be observed with the agent.

This example also provides a first informal justification for using the adjective “desired” to describe – from an external observer perspective – the additional populations of neurons. Indeed, the observed behavior of the agent, bringing the cup to its mouth, can be seen as an action performed to adjust its current thirst level to the “desired” thirst level. Similarly, the “desired” visual state (i.e. the image of the cup close to the eyes) can be interpreted as an intermediate, seemingly desired goal making the agent contract its arm when it sees the cup it is holding in its hand. For the sake of readability, this adjective will from now on be used without quotation marks.

2.3 Formal Description

Let $\mathbf{X}^M(t) = (x_1^M(t), x_2^M(t), \ldots)^\mathrm{T}$ be the state vector of modality M, i.e. the signals corresponding to the state of the physical device – sensor or motor – at time t. Each modality M consists of five populations of linear real-valued neurons (i.e. artificial neurons with a linear activation function), whose activity is represented by the following five vectors:

1. Delayed state $\mathbf{A}^M(t)$
2. Current state $\mathbf{B}^M(t)$
3. Desired state $\mathbf{C}^M(t)$
4. State change $\mathbf{D}^M(t)$
5. Desired change $\mathbf{E}^M(t)$



(a)

(b)

(c)

Figure 2-11: Fictive scenario illustrating the working principle of the proposed neural architecture (final version). (a) The activity of the neuron representing the desired thirst level is artificially decreased (indicated by the black triangle in the neural network), mimicking the idea that the agent "wants to feel" a low level of thirst. Activity propagates into the visual modality, projecting the image of a large object in the desired visual population. (b) If the agent holds a cup in its hand, neural activity propagates further into the motor modality, generating a contraction of the arm muscle. (c) This produces the same "procedural" behavior as discussed before (compare with Figure 2-7).

The neural activity of the delayed and current state populations is defined as:

$$\mathbf{A}^M(t) \quad := \quad \mathbf{X}^M(t-\tau) \tag{2.1}$$

$$\mathbf{B}^M(t) \quad := \quad \mathbf{X}^M(t) \tag{2.2}$$

where $\tau > 0$ is a small time delay, typically one time step. The neural activity of the state change and desired change populations is defined as:

$$\mathbf{D}^M(t) \quad := \quad \Delta^M(\mathbf{A}^M(t), \mathbf{B}^M(t)) \tag{2.3}$$

$$\mathbf{E}^M(t) \quad := \quad \Delta^M(\mathbf{B}^M(t), \mathbf{C}^M(t)) \tag{2.4}$$

where $\Delta^M(\cdot, \cdot)$ is a function computing a variation between two states. Often, Δ^M is chosen simply as the component-wise difference:

$$\Delta^M(\mathbf{a}, \mathbf{b}) \quad := \quad \mathbf{b} - \mathbf{a} \tag{2.5}$$

The synaptic weight matrix $\mathbf{W}^{MN}(t)$ connecting any two different modalities $M \neq N$ is initially set to zero

$$\mathbf{W}^{MN}(0) \quad := \quad 0 \tag{2.6}$$

and is updated with a simple Hebbian learning rule capturing the correlation between the current state of one modality $\mathbf{B}^M$ and the state change of the other modality $\mathbf{D}^N$:

$$\Delta\mathbf{W}^{MN}(t) \quad := \quad \eta \cdot \mathbf{B}^M(t) \cdot \mathbf{D}^N(t)^{\mathrm{T}} \tag{2.7}$$

$$\mathbf{W}^{MN}(t+1) \quad := \quad \mathbf{W}^{MN}(t) + \Delta\mathbf{W}^{MN}(t) \tag{2.8}$$

where $\eta > 0$ is the learning rate, typically set to $\eta = 0.01$. In order to prevent the synaptic weights from increasing infinitely, Equation (2.7) could also include a forgetting term and be rewritten as:

$$\begin{aligned} \Delta\mathbf{W}^{MN}(t) \quad := \quad & \eta \cdot \mathbf{B}^M(t) \cdot \mathbf{D}^N(t)^{\mathrm{T}} \\ & - \varepsilon \cdot \|\mathbf{B}^M(t)\| \cdot \mathbf{W}^{MN}(t) \end{aligned} \tag{2.9}$$

However, in most practical cases, this active forgetting is not necessary. The forgetting rate $\varepsilon \geq 0$ is thus usually set to $\varepsilon = 0$.

Finally, the “desired” state $\mathbf{C}^M$ is defined as the summed weighted input from all other modalities:

$$\mathbf{C}^M(t+1) \quad := \quad \sum_{N \neq M} \mathbf{W}^{MN}(t) \cdot \mathbf{E}^N(t) \tag{2.10}$$

If M is a motor modality, the “desired” state $\mathbf{C}^M(t)$ is the command sent at time t to the physical device.

The implementation of the proposed neural architecture on a particular robotic agent thus only requires defining the value for the learning and forgetting rates η and ε, and for each modality M, the sensorimotor signals X^M (e.g. n signals in the proximity modality whose values between 0 and 1 represent the readings of the n physical infrared sensors), as well as the state variation function Δ^M (most usually the component-wise difference between two successive states, except for the visual modality – see Appendix 3.A).

2.4 Discussion

2.4.1 Properties of the Neural Architecture

The main characteristic of the neural architecture is the absence of explicit control or regulation mechanism. The network is indeed designed to produce some stable neural dynamics using Hebbian synaptic plasticity. At the same time, we showed how the particular coupling scheme, initially chosen to allow neural activity to propagate throughout the network, displayed an interesting additional property: by capturing certain correlations of sensory-motor activity, the network learns to associate the previously experienced sensory-motor patterns that lead the system to given states. This property allowed us to refer – from an observer perspective – to one population of neuron in each modality as the “desired” state.

Another important aspect of the proposed neural architecture is its structural homogeneity. Indeed, the signals of every sensor and motor are represented – for each and every modality – with the same neural structure (five populations of artificial neurons with a particular arrangement shown in Figure 2-9). Also, all modalities are systematically and reciprocally coupled to each other with the same coupling structure (see Figure 2-10), regardless of the particular nature of the sensors or motors that each modality represents. The consequences of this homogeneity will constitute the subject of Chapter 5.

The neural architecture is also minimal in the sense that there are no additional populations of neurons beyond the structures defining every modality. The network does not contain any so-called “hidden” layers of inter-neurons. Similarly, there are no specific structures – such as “contextual” neurons with self-recurrent connections – that maintain neural activity over extended periods of time, thus providing some kind of “working memory” to the system. This issue, together with some remarkable outcome, will be discussed in Chapter 7.

Finally, the architecture developed in this chapter has been shown to be one of the simplest solutions we could think of, based on Hebbian synaptic plasticity, to let sensory and motor signals interact with each other and potentially produce some stable dynamics when coupled to an embodied system. This implies that, even though the work of this thesis only explores this particular solution, they are potentially many other neural architectures – for instance based on different models of artificial neurons or synaptic plasticity, or possibly with simpler solutions – that are also worth exploring (see discussion in Chapter 8).

2.4.2 No Built-In Reflexes and No Goal

Before the agent starts interacting with its environment, all synaptic weights of the network are initialized to zero. In other words, sensors and motors initially do not interact: the agent has no built-in reflexes, nor any similar behavioral primitives[5] provided by an external designer. On the contrary, the neural network will spontaneously develop – through its Hebbian-like synaptic plasticity – structures in its connectivity when the agent moves and interacts with its surrounding world.

Similarly, the system has no goal. The Hebbian learning rule, which modifies the synaptic weights of the network, is not modulated by any value system that would define a particular goal. It is worth noting at this point that the term *learning* can be slightly misleading: the Hebbian learning rule is not a learning strategy allowing the agent to achieve a given task or optimizing a given fitness function; rather, it is an arbitrary rule that defines the synaptic plasticity of the network, allowing the internal neural dynamics to evolve and to be influenced by the external physical dynamics of the agent.

2.4.3 Comparison with Neurobiology

The neural architecture introduced in this chapter consists of an artificial neural network. As such, it obviously inherits the biological inspiration that underlies the now well-known and widely applied field of neural networks, such as the parallel nature of a large number of simple and highly connected elements. Similarly, the Hebbian learning rule used in our

---

[5] These primitives refer to basic behaviors designed to achieve incrementally complex tasks, such as those found in the subsumption architecture (Brooks, 1986), rather than to the intrinsic behavior of the system produced with no neural control.

model is a particular adaptation of the principle suggested by the physiologist Donald Hebb about the synaptic plasticity of biological brains (Hebb, 1949).

Apart from this, the proposed architecture does not aim at modeling any further features of biological systems. Nevertheless, it is possible to discover, at various levels of abstraction, more aspects related to canonical neurobiological findings (Kandel et al., 1991).

For instance, the proposed neural architecture contains basically two types of neurons: one type of neurons representing the sensorimotor signals (the states), the other representing only variations of these signals (the state changes). Interestingly, a similar distinction is found in the classification of biological neurons: slowly adapting (tonic) neurons respond continually to enduring stimuli, whereas rapidly adapting (phasic) neurons respond only to changes in stimulus intensity.

Also, the global organization of the artificial network evokes the coarse structure found in the cortex: both systems consist – at least to some extent – of local collections of neurons (sensory and motor areas) sharing similar structures, responding mainly to one particular sensory or motor modality, and being highly interconnected through synaptic projections.

In addition, the more local structure of the artificial network shares further similarities with its biological counterpart. The internal structure of every modality in our model consists of distinct populations of neurons, some receiving input and others providing output to other modalities. Interestingly, viewing these populations as *layers* allows a surprising similitude to be drawn with the layered structure characteristic of the cerebral cortex. Indeed, every area of the cerebral cortex is also divided into several well-defined layers, some layer (4) receiving most of the afferent input, and some other layers (2 and 3) providing much of the output to other cortical regions.

On the other hand, it is important not to overstate the analogy, or expect a too close comparison between our model and neurobiological findings. Our approach has its roots in robotics and tries above all to better understand some of the preconceptions that prevent artificial models from further shedding light on principles of biological systems. At a certain level of abstraction, the work of this thesis will raise several questions regarding natural systems, and we have tried our best to discuss and document connections with our model (see the introductions and discussions of Chapters 4, 5 and 7, as well as Appendix 5.A and Section 7.5). Yet, the motivation underlying our model (such as aiming for a minimal complexity in the neural architecture) is obviously very different from the constraints any million-year-old biological system is dealing with. Thus, even though our approach can demonstrate possible alternative principles for different behaviors (such as the artificial need for dedicated neural structures), we should not be surprised that no concrete neurobiological evidence have (yet) been found.

2.4.4 Relative Complexity

Complex – i.e. non-linear – dynamic systems are in general hard to understand. For instance, it is often difficult to determine how neural networks function, why they produce certain output, which connections play a key role, or which parts of the network can be discarded. This is especially true for recurrent networks or networks that learn. The reason is that single components of the system often tightly interact with each other, making difficult to systematically analyze them in isolation.

In contrast, an interesting property of the proposed neural architecture – which is possibly already suggested by the fictive scenario discussed in this chapter, but which will become more evident during the next chapters – is that, even though the observed behavior of the agent results from the global interaction of the different components of the system, the dynamics of the neural network can nevertheless be fully accounted for by an analysis of the individual parts. Indeed, each cross-modal coupling will have a clearly defined role in the neural network. By representing the synaptic strengths graphically, it will be relatively straightforward to get an intuitive understanding of the cross-modal correlations – i.e. the associations – learned by the network. Moreover, these associations will often be quite obvious (e.g. motor activity generating self-rotation in one direction associated with optical flow in the opposite direction). Finally, the particular structure of the proposed network – which could at first sight be a little confusing – will make possible to follow relatively easily the propagation of neural activity from one sensorimotor modality to the next, and to identify the neural pathways across the neural network that are involved in the generation of the observed behaviors.

The next chapters will show that the complexity of the observed behaviors does not stem from complex structures within the neural substrate. Rather, the observed complexity will be shown to be produced by a rich interaction between the agent's body, sensorimotor system and environment.

2.4.5 Terminology, Notation and Practice

This chapter not only introduces the neural architecture, but also the accompanying terminology, notation and practice, which will be used in the remaining chapters of this thesis.

For instance, Figure 2-10 on page 25 illustrates both the nomenclature as well as the graphical arrangement describing the five different populations of neurons constituting a modality. Similarly, Figure 2-11 on page 27 illustrates how cross-modal couplings are symbolized by outlined arrows: dotted arrows represent synaptic couplings with negligible weights, continuous arrows represent synaptic couplings having captured some significant correlations, and bold arrows indicate neural activity propagating between two modalities. This figure also shows how neural activity within different populations of neuron can be illustrated graphically, and how an artificial increase or decrease of neural activity is indicated by a black triangle.

The fictive scenario used to illustrate the working principle of the proposed architecture also introduces the typical steps followed when an experiment is conducted. First, the chosen agent is described, including the sensors and motors it is equipped with and the environment in which it is engaged. What follows next is a preliminary exploration phase where the motors of the agent are randomly stimulated for a while, allowing the neutal network to capture some initial cross-modal correlations. We then observe the behavior of the agent produced when the motor neurons get activated by neural activity propagating across the network. This neural activity can be spontaneously generated in any sensory or motor modality, namely as soon as the difference between activity in the current and desired state populations induces activity in the desired change population. An example of such purely "spontaneous" behavior will be discussed in Chapter 3. In addition, neural activity can also be artificially produced in an arbitrary modality by externally increasing or decreasing the activity of one neuron in the desired state population. Metaphorically speaking, this external stimulation can be seen as a generating a "drive" in the system, mimicking the idea that the agent "wants to feel" a particular sensory signal, such as a low level of thirst (as in the thought experiments described previously in this chapter), or a high battery level (if the agent has a sensor detecting its battery level).

The proposed architecture has the interesting property of being, formally, invariant to the calibration of the sensorimotor signals: the zero-point of the signals in any modality can be shifted without changing the behaviors observed with the agents[6]. Consequently, this

---

[6] Let us consider the case where the calibration of modality M is shifted by some arbitrary value θ, i.e. where $\mathbf{X}^M(t) = (x_1^M(t), x_2^M(t), \ldots)$ is replaced by: $\tilde{\mathbf{X}}^M(t) = \mathbf{X}^M(t) + \theta = (x_1^M(t) + \theta, x_2^M(t) + \theta, \ldots)$. The difference function $\Delta^M(\cdot,\cdot)$ is in general invariant to calibration (see Eq. 2.5): $\Delta^M(\mathbf{a}+\theta, \mathbf{b}+\theta) = \Delta^M(\mathbf{a}, \mathbf{b})$. As a consequence, the state change activity is invariant to calibration (2.1 – 2.3):

$$\begin{aligned}\tilde{\mathbf{D}}^M(t) &= \Delta^M(\tilde{\mathbf{A}}^M(t), \tilde{\mathbf{B}}^M(t)) \;=\; \Delta^M(\mathbf{A}^M(t)+\theta, \mathbf{B}^M(t)+\theta) \\ &= \Delta^M(\mathbf{A}^M(t), \mathbf{B}^M(t)) \;=\; \mathbf{D}^M(t)\end{aligned}$$

Also, the weight matrix $\mathbf{W}^{MN}$ of the synapses connecting this modality to any other modality N becomes, using (2.7):

$$\begin{aligned}\tilde{\mathbf{W}}^{MN} &= \sum \eta \cdot \tilde{\mathbf{B}}^M \cdot \mathbf{D}^N \;=\; \sum \eta \cdot (\mathbf{B}^M + \theta) \cdot \mathbf{D}^N \\ &= \sum \eta \cdot \mathbf{B}^M \cdot \mathbf{D}^N + \sum \eta \cdot \theta \cdot \mathbf{D}^N \\ &= \mathbf{W}^{MN} + \eta \cdot \theta \cdot \sum \mathbf{D}^N\end{aligned}$$

In general, change populations have zero mean activity: any modality N displays, on average, as much external stimulation of neurons in the desired populations could thus be simply replaced – and can thus be viewed – as some "renormalization" of the signal values in the corresponding modality. Nonetheless, we think it is more convenient to stick to the "explicit notation" where neural activity is explicitly generated in some modality, in order to better appreciate the propagation of neural activity in the neural network, rather than using some special, implicitly chosen calibration.

2.5 Summary

This chapter introduced a generic neural architecture for embodied agent which constitutes the basis for the remaining chapters of this thesis. We discussed the underlying motivation, described the detailed structure of the network, and illustrated the working principle with a thought experiment. The aim of this model is to provide a minimal neural architecture based on Hebbian synaptic plasticity with the potential of generating various coherent behaviors when coupled to an embodied system.

In the following chapters, we will not only show that a variety of behaviors can be observed with different robots using the proposed architecture, but that this novel approach can shed new light on several issues related to embodied artificial intelligence and cognitive science.

---

increase of sensorimotor signals as decrease thereof. Thus, $\sum \mathbf{D}^N = 0$. Consequently, the weight matrix does not change either under calibration: $\tilde{\mathbf{W}}^{MN} = \mathbf{W}^{MN}$.

Appendix 2.A
Synaptic Coupling as a Local Network

This appendix shows that the abstract synaptic coupling used in the proposed neural architecture, where the strength is modified according to the correlation of activity in two neurons, and where the neural activity is transmitted between two other neurons, can be approximated by a simple local network of traditional artificial synapses.

Figure 2-12(a) illustrates a simple network of standard artificial synapses that approximates the behavior of the abstract synapse, shown in Figure 2-12(b), which learns the correlation of activity between two neurons (a and b) and allows neural activity to propagate between two other neurons (a' and b').

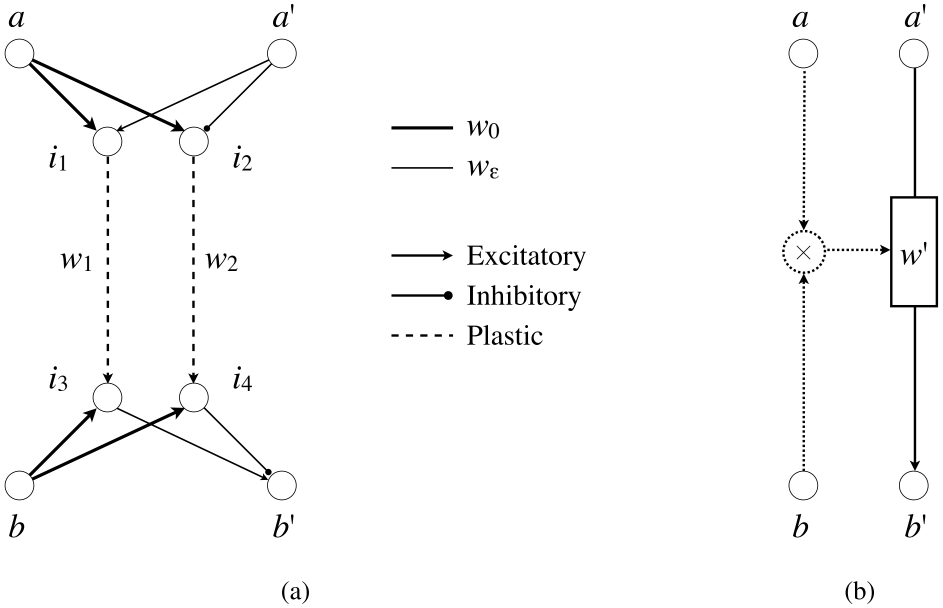
The basic idea is that the two pairs of interneurons (i_1, i_2) and (i_3, i_4) are essentially activated by the neurons a and b, respectively. Therefore, the plastic weights w_1 and w_2 both capture at first approximation the correlation between the activity of a and b. On the other hand, the synaptic input to neuron b' is the difference between the activity of i_3 and i_4, and therefore reflects the synaptic input from a' only. The following paragraphs provide a formal description of the system.

The activity of the interneurons $i_1, \ldots, i_4$ is given by:

$$\begin{aligned} i_1 &= w_0 \cdot a + w_\varepsilon \cdot a' & (2.11)\\ i_2 &= w_0 \cdot a - w_\varepsilon \cdot a' & (2.12)\\ i_3 &= w_0 \cdot b + w_1 \cdot i_1 & (2.13)\\ i_4 &= w_0 \cdot b + w_2 \cdot i_2 & (2.14) \end{aligned}$$

The weights of the plastic synpases w_1 and w_2, both initially set to zero, are modified by a simple Hebbian learning rule:

$$\begin{aligned} \dot{w}_1 &= \eta \cdot i_1 \cdot i_3 & (2.15)\\ \dot{w}_2 &= \eta \cdot i_2 \cdot i_4 & (2.16) \end{aligned}$$



(a) (b)

Figure 2-12: (a) Local network of traditional artificial synapses approximating the behavior of (b) the abstract synaptic coupling used in the proposed neural architecture.

with learning rate $\eta > 0$. Combining the above equations gives:

$$\begin{aligned} \dot{w}_1 &= \eta \cdot (w_0 \cdot a + w_\varepsilon \cdot a') \cdot (w_0 \cdot b + w_1 \cdot (w_0 \cdot a + w_\varepsilon \cdot a')) && (2.17) \\ &= \eta \cdot (w_0 \cdot a + w_\varepsilon \cdot a') \cdot (w_0 \cdot (b + w_1 \cdot a) + w_1 \cdot w_\varepsilon \cdot a') \\ \dot{w}_2 &= \eta \cdot (w_0 \cdot a - w_\varepsilon \cdot a') \cdot (w_0 \cdot b + w_2 \cdot (w_0 \cdot a - w_\varepsilon \cdot a')) && (2.18) \\ &= \eta \cdot (w_0 \cdot a - w_\varepsilon \cdot a') \cdot (w_0 \cdot (b + w_2 \cdot a) - w_2 \cdot w_\varepsilon \cdot a') \end{aligned}$$

Assuming $w_0 \gg w_\varepsilon$ and $w_1, w_2 \ll 1$, the last two equations can be approximated by:

$$\begin{aligned} \dot{w}_1 &\approx \eta \cdot (w_0 \cdot a) \cdot (w_0 \cdot b + w_1 \cdot w_\varepsilon \cdot a') && (2.19) \\ &\approx \eta \cdot w_0^2 \cdot a \cdot b \\ \dot{w}_2 &\approx \eta \cdot (w_0 \cdot a) \cdot (w_0 \cdot b - w_2 \cdot w_\varepsilon \cdot a') && (2.20) \\ &\approx \eta \cdot w_0^2 \cdot a \cdot b \end{aligned}$$

The activity of the neuron b' is:

$$\begin{aligned} b' &= w_\varepsilon \cdot i_3 - w_\varepsilon \cdot i_4 = w_\varepsilon \cdot (i_3 - i_4) \\ &= w_\varepsilon \cdot ((w_0 \cdot b + w_1 \cdot i_1) - (w_0 \cdot b + w_2 \cdot i_2)) \\ &= w_\varepsilon \cdot (w_1 \cdot i_1 - w_2 \cdot i_2) \\ &= w_\varepsilon \cdot (w_1 \cdot (w_0 \cdot a + w_\varepsilon \cdot a') - w_2 \cdot (w_0 \cdot a - w_\varepsilon \cdot a')) \\ &= w_\varepsilon \cdot ((w_1 - w_2) \cdot w_0 \cdot a + (w_1 + w_2) \cdot w_\varepsilon \cdot a') \end{aligned} \tag{2.21}$$

Since $\dot{w}_1 \approx \dot{w}_2$, we can further assume that at first approximation $(w_1 - w_2) \approx 0$, which gives us:

$$b' \approx w_\varepsilon^2 \cdot (w_1 + w_2) \cdot a' \tag{2.22}$$

Finally, substituting $w' := w_\varepsilon^2 \cdot (w_1 + w_2)$ and $\eta' := 2 \cdot \eta \cdot (w_\varepsilon \cdot w_0)^2$, the system can be rewritten as:

$$b' \approx w' \cdot a' \tag{2.23}$$

$$\dot{w}' \approx \eta' \cdot a \cdot b \tag{2.24}$$

Thus, at first approximation, the synaptic input to neuron b' is proportional to the activity of neuron a', with a synaptic weight w' that is modified with a simple Hebbian rule learning the correlation of activity between neurons a and b.

Chapter 3

Seemingly Goal-Directed Behaviors

This chapter describes a first series of experiments where the neural architecture introduced previously is implemented on a physical mobile robot. The aim of this study is twofold. On the one hand, it shows that coherent and seemingly goal-directed behaviors are observed as the robot interacts with a particular environment. On the other hand, it highlights some interesting and innovative insights that can be gained from the approach proposed in this thesis.

3.1 Introduction

In Chapter 2, we introduced a new neural architecture for autonomous agents, and briefly sketched its working principle. We now study how this neural model can successfully be used with a real robot – a physically embodied agent – to generate interesting observable behaviors. It is indeed essential to show that the model is robust enough to cope with the intrinsic complexity of a real-world situation. Embodiment implies that the agent is continuously influenced by its environment. For instance, sensors are subject to inherent noise, and motors to friction and energy dissipation. Nevertheless, we will show that interaction with the environment creates various correlations between sensory and motor activities, which are stable enough to be captured and exploited by the neural architecture to generate coherent behaviors.

The results of this initial series of experiments, even though performed in a relatively simplified environment, are interesting in many respects. First, they show that contact avoidance and object following behaviors can be achieved without any built-in reflex, behavioral primitives, reinforcement mechanism, or dedicated structures in the neural architecture. This point is closely related to issues of the frame of reference problem (Clancey, 1989, 1991a). In particular, it shows that there is not necessarily a direct relation between observed behaviors and underlying mechanisms: one cannot be deduced from the other since the behavior of the agent is always the result of a system-environment interaction.

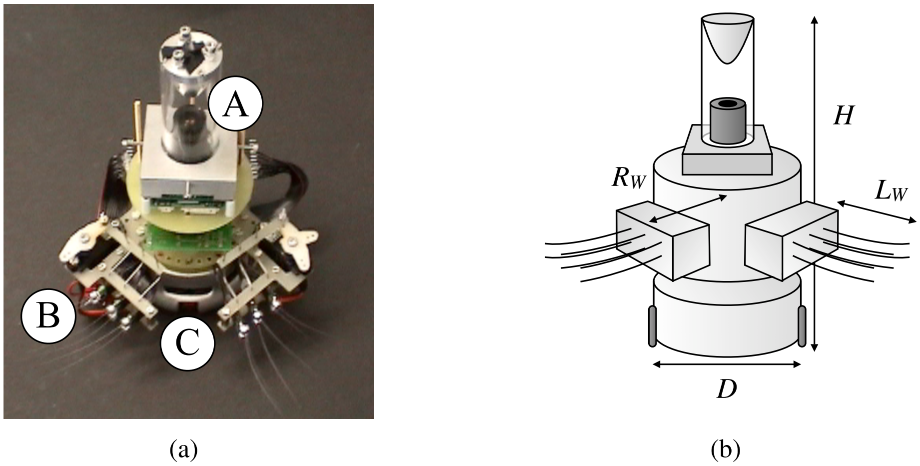
Second, this study highlights the role that material properties can play in generating observed behaviors (Hara and Pfeifer, 2000; Lungarella et al., 2002a; Pfeifer et al., 2005; Hosoda et al., 2006). More specifically, the tactile sensor of the robot used in this chapter is composed of compliant whiskers and signals whether contact with an external object is detected or not. One might therefore intuitively assume that this sensor, being used as a binary contact detector, could be substituted with any other kind of contact sensor – for instance, an infrared proximity sensor. However, we will show that, on the contrary, the observed behaviors are only produced if the robot is equipped with a *flexible* contact sensor. In other words, the particular material properties of the tactile sensor generate cross-modal correlations in sensory activity – correlations that will shape the dynamics of the neural network.

Finally, a consideration of the agent-environment interaction at different time scales unveils some interesting characteristics of the system. When activity is generated in the tactile modality – mimicking the idea that the robot "feels" tactile stimulation – the robot moves towards any object placed in the center of its visual field, i.e. displays a behavior that expresses the causal relation between approaching a distant object and (only later) detecting contact with it. It is interesting to note that this seemingly goal-directed behavior is produced even though the neural network only learns instantaneous correlations, i.e. correlations between sensory and motor events that happen simultaneously. This suggests that, at least to some extent, learning immediate correlations is sufficient to capture some temporal relation between motor actions and their delayed effects.

The chapter is structured as follows. The next section describes the agent – a mobile robot equipped with tactile and visual sensors – and the environment. We then analyze in Section 3.3 the various cross-modal correlations captured by the neural network as the robot is randomly driven across the environment. Section 3.4 describes behaviors observed under different conditions, and explains how they are generated by propagation of neural activity throughout the neural network. Section 3.5 provides some discussion and analysis about the implications gained from the presented study. The final section ends the chapter with conclusions.

3.2 Agent and Environment

This experiment is conducted with the AMouse robot (illustrated in Figure 3-1), a Khepera-based mobile platform equipped with an omnidirectional camera and artificial whisker ar-



(a) (b)

Figure 3-1: (a) The “AMouse” robot, consisting of an omnidirectional camera (A) and artificial whisker sensors (B) mounted on a mobile Khepera platform (C). (b) Schematic of the robot. The specifications of the robot are given in Table 3.1.

Parameter	Description	Approx. Value
H	Height	16 cm
D	Diameter	6 cm
R_W	Axial distance of whisker basis	5 cm
L_W	Average whisker length	3 cm
m	Mass of the robot	300 g

Table 3.1: Mechanical specifications of the robot.

rays used as a tactile sensor. The neural network contains the three following modalities:

Tactile Modality

The current state in the tactile modality is represented by one single neuron whose binary activity indicates whether contact is detected by the whisker sensor. The tactile sensor is composed of two arrays of active[1] whisker sensors mounted in the front of the robot. Each whisker sensor consists of a natural rat hair glued onto a microphone. The recorded signal being proportional to the deflection of the microphone membrane, it provides an indication of how strong the whisker is being stimulated (Fend et al., 2003). Contact is detected as soon as the low-pass filtered input signal from any whisker sensor is above a given threshold (see Figure 3-2).

Note that for the sake of simplicity, and because of the very noisy signals, we deliberately choose for this initial series of experiments a single, binary contact representation, despite the fact that the agent is equipped with multiple whiskers.

---

[1] The whiskers are constantly and rapidly swept slightly back and forth.

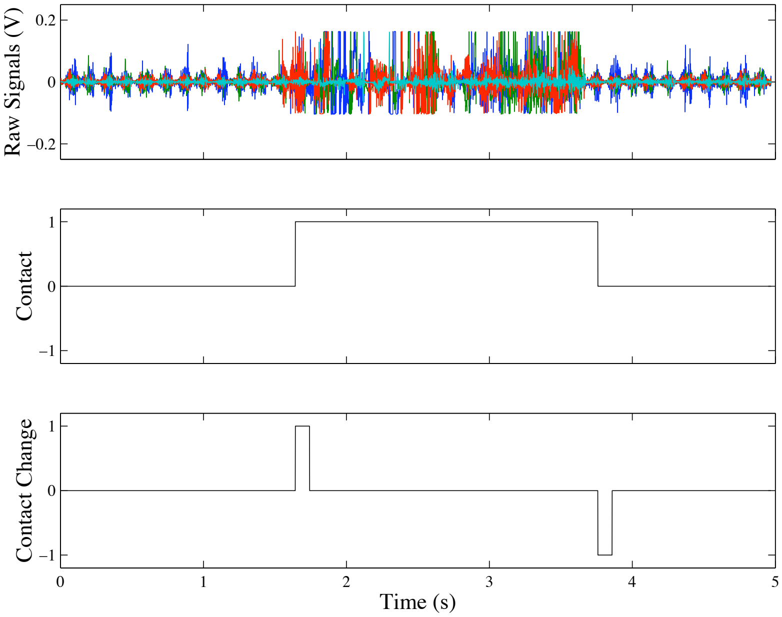


Figure 3-2: Tactile sensory input. Top: Raw signals from the whisker sensors. Middle: Binary contact detection obtained by thresholding the low-pass filtered raw signals. Bottom: Change of contact, indicating the on-set and off-set of contact detection.

Also, since onsets and offsets of tactile contact are relatively sparse, we set the time delay (see Equation 2.1) for the tactile modality to $\tau = 5$.

Visual Modality

The current visual state is represented by an array of 200×100 neurons, whose activities correspond to the grayscale pixel values of the panoramic image of the surrounding scene extracted from the omnidirectional camera (see Figure 3-3).

In the tactile and motor modalities, the state change is simply defined as the component-wise difference between two successive states. Yet in the visual modality, the state change is defined slightly differently, such that it provides an approximation of the horizontal and vertical components of the optical flow computed at each pixel. Basically, the visual state change, rather than being defined only as $\mathbf{I}(t) - \mathbf{I}(t-1)$, is defined as $\{\mathbf{\Phi}_x(t), \mathbf{\Phi}_y(t)\}$, where $\mathbf{\Phi}_x$ and $\mathbf{\Phi}_y$ are the following estimations for the horizontal and vertical components

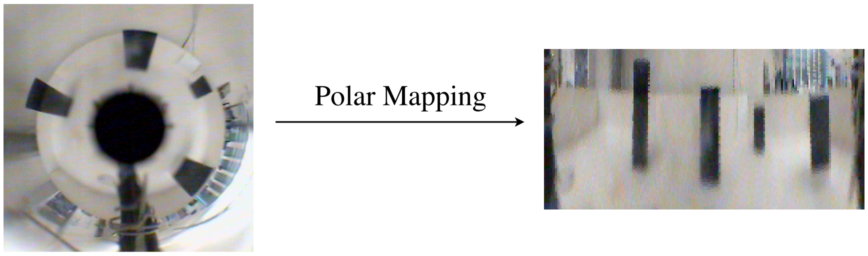
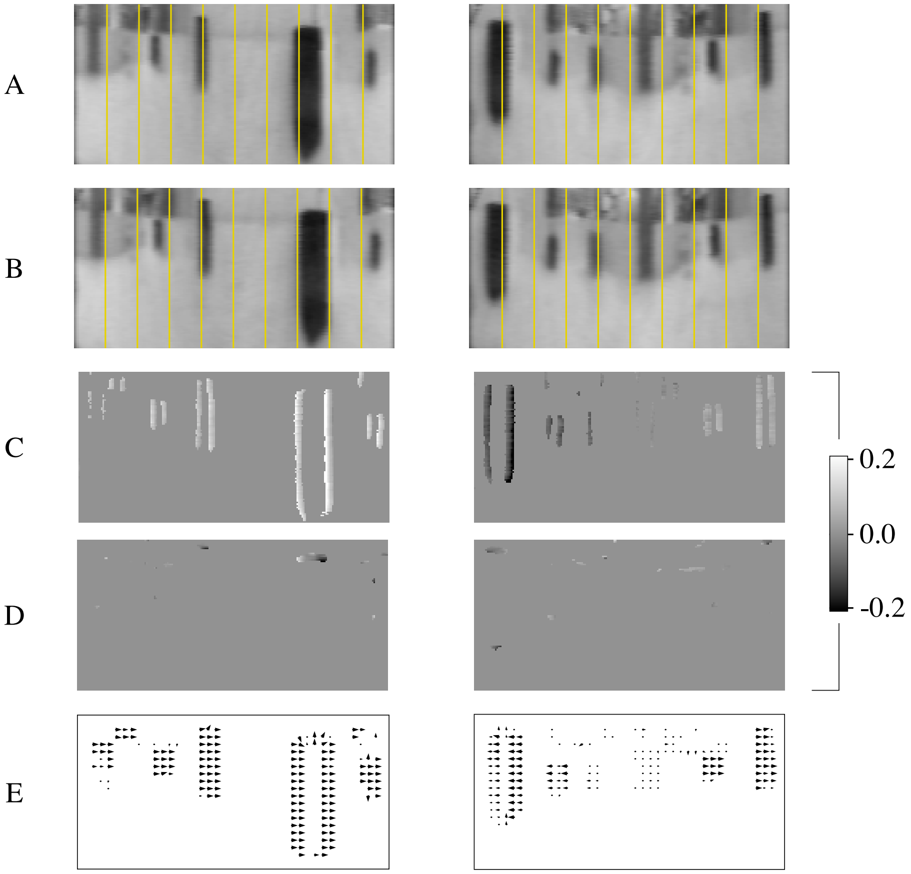


Figure 3-3: Omnidirectional vision system. Left: Original image taken from the camera pointing on the hyperbolic mirror placed on the top of the robot. Right: Panoramic view obtained by a polar transform of the original camera image.



(a) (b)

Figure 3-4: Approximation of optical flow between two successive visual states. A: Previous visual state, $\mathbf{I}(t-1)$. B: Current visual state, $\mathbf{I}(t)$. C, D: Horizontal and vertical components of the optical flow, respectively $\mathbf{\Phi}_x(t)$ and $\mathbf{\Phi}_y(t)$. E: Optical flow field, $(\mathbf{\Phi}_x(t), \mathbf{\Phi}_y(t))$. The two samples were recorded as the robot was either (a) turning to the left or (b) moving forward.

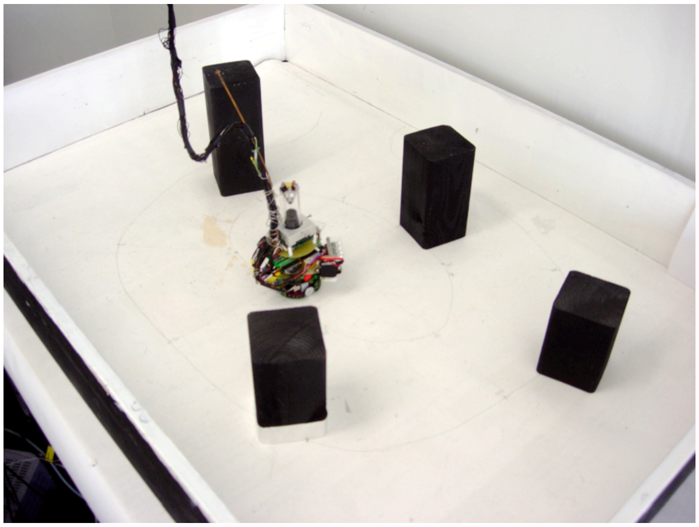


Figure 3-5: White-painted arena containing a few black blocks, used as environment for the series of experiments discussed in this chapter.

of the optical flow:

$$\begin{aligned} \mathbf{\Phi}_x(t) &:= -\frac{\mathbf{I}(t) - \mathbf{I}(t-1)}{\nabla_x \mathbf{I}(t)} \\ \mathbf{\Phi}_y(t) &:= -\frac{\mathbf{I}(t) - \mathbf{I}(t-1)}{\nabla_y \mathbf{I}(t)} \end{aligned}$$

where $\mathbf{I}(t)$ is the visual state at time step t. A detailed account for the optical flow estimation is given in Appendix 3.A. Figure 3-4 illustrates how the visual state change defined in this way indeed provides an approximation of the perceived optical flow. In this chapter, the gradients are estimated over a radius of $h = 3$ pixels (see Equation 3.5 on page 60).

Motor Modality

The current state consists of two neurons, indicating respectively the forward and turning speeds of the robot. Forward and turning speeds are mapped to the individual rotation speed of the left and right wheels of the robot as follows:

$$\begin{pmatrix} v_{\text{left}} \\ v_{\text{right}} \end{pmatrix} := \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_{\text{forward}} \\ v_{\text{turn}} \end{pmatrix}$$

The translational speed of the robot for $v_{\text{forward}} = 1$ is approximately 1.5 cm/s, and the rotational speed for $v_{\text{turn}} = 1$ is approximately 30 °/s.

Environment

The robot can move freely in a flat, white-painted arena containing a few objects consisting of black blocks (see Figure 3-5).

3.3 Cross-Modal Correlations

This section describes the cross-modal correlations learned during the initial phase where the robot is randomly driven across the environment. As the agent "explores" its interaction with the environment, the Hebbian plasticity of the neural network continuously modifies the weights of the synapses coupling the different modalities.

It turns out that most synaptic weights do not change significantly and remain essentially to zero (Figure 3-6). However, the neural network will capture two significant cross-modal correlations (Figure 3-7), which are described in the following subsections.

Visual Change → Motor State

Each time the robot turns on the right, the panoramic image of the surroundings is translated to the left, and vice versa. In other words, there is a correlation between neural activity in the visual modality corresponding to a uniform lateral optical flow and activity of the motor neuron corresponding to the turning speed. Similarly, forward motion is correlated to an optical flow pattern expanding from the front direction.

These obvious correlations can be seen in Figure 3-8, which graphically represents the weights of the synapses coupling the visual modality to the motor modality.

Tactile Change → Visual State

A relatively intuitive account can be given for the correlation between change of tactile stimulation and visual state. Indeed, whenever the robot gets close enough to an object in front of it, there is an onset of tactile stimulation and at the same time, a black shape corresponding to the object is perceived more or less in the center of the visual field. Therefore, roughly speaking, onset of tactile stimulation is correlated to black pixels in the center of the camera image.

A more detailed description is provided in Figure 3-9, for the case where the whiskers are swept against an object while the robot is turning on the spot in successively two opposite directions. The correlation learned between tactile change and visual state represents an outlined object placed in front of the robot.

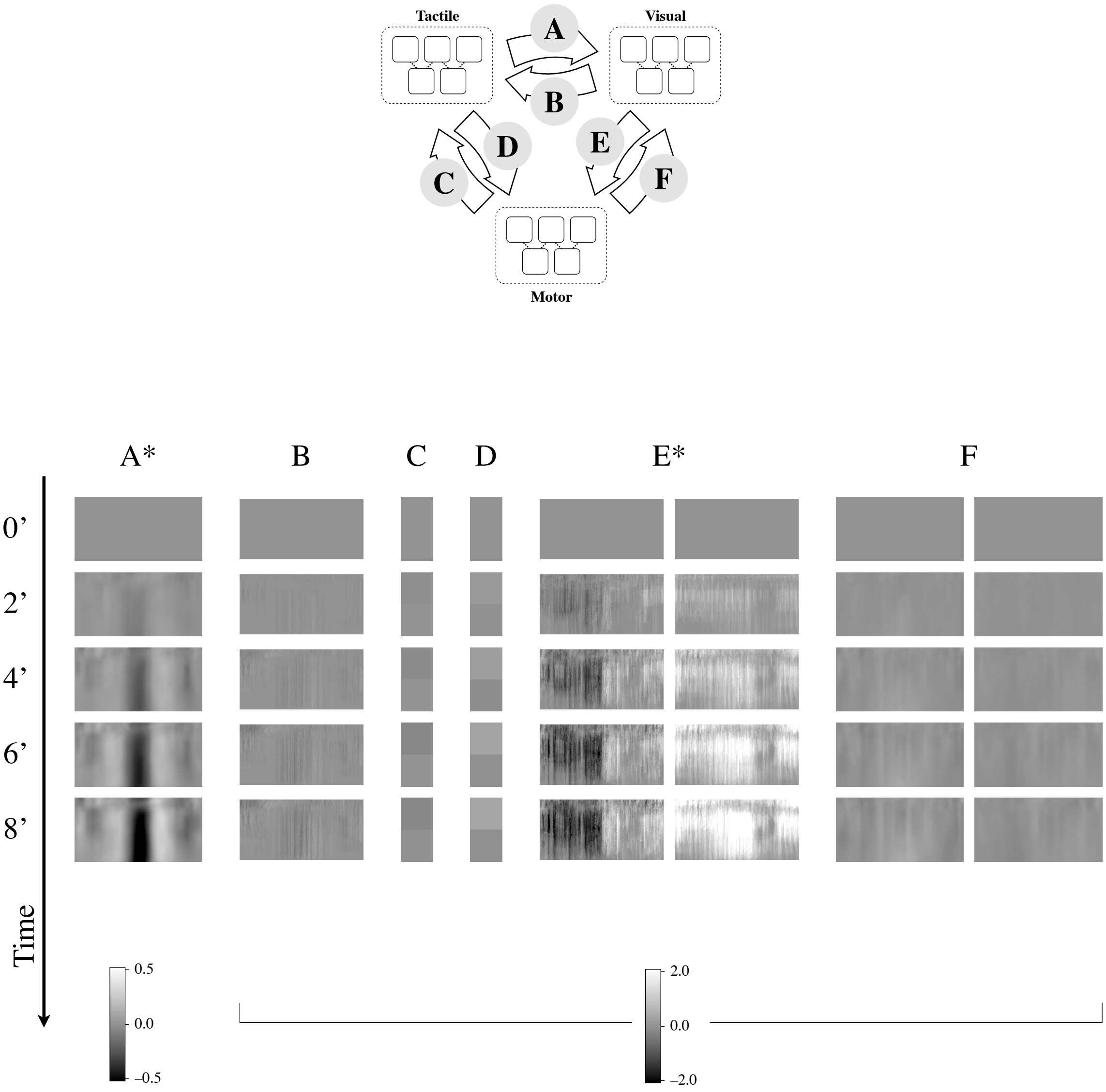


Figure 3-6: Time evolution of the synaptic weights. Top: Schematic illustration of the neural network. Bottom: Weight matrices corresponding to the cross-modal synaptic couplings indicated above. The synaptic couplings that capture a significant correlation are indicated with asterisks. A: visual activity correlated to change of tactile activity. E: optical flow correlated to forward motion (left column) and turning motion (right column); a graphical representation of these weights is provided in Figure 3-8. For the sake of clarity, the figure only shows the synaptic weights corresponding to the horizontal components of the optical flow. The learning and forgetting rates of the network (see Equation 2.9) are $\eta = 0.01$ and $\varepsilon = 0$, respectively.

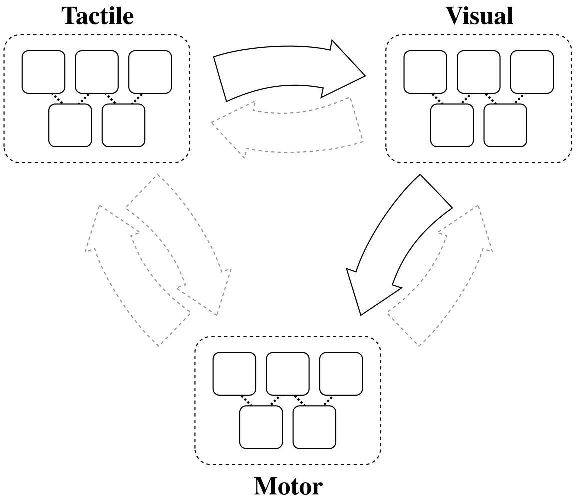
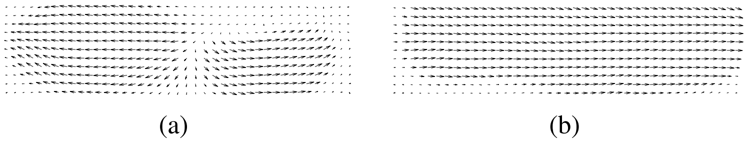


Figure 3-7: Schematic illustration of the neural network. The arrows represent the synaptic couplings between all modalities. The network captures only two significant cross-modal correlations, indicated by plain arrows (the corresponding synaptic weights are illustrated in Figures 3-8 and 3-10). The other cross-modal correlations are negligible, and are indicated by dotted arrows.



(a) (b)

Figure 3-8: Graphical representation of the synaptic weights coupling the visual modality to the motor modality, showing the optical flow correlated to activity of the motor neuron indicating (a) forward motion and (b) turning motion.

The actual correlation captured by the robot after various encounters with objects from different angles and positions is shown in Figure 3-10. Because of the multiple conditions under which the whiskers are stimulated, there is no significant correlation with pixels in the background. In summary, change of tactile stimulation is correlated with a pattern in the visual state representing an isolated object in the center of the visual field.

3.4 Experiments

Let us now study the behaviors displayed by the robot when it is let to move on its own. We investigate the behaviors observed under two conditions. First, no neural activity is artificially generated in the network; this corresponds to the "intrinsic" behavior of the robot. Second, we observe what happens when one neuron in the tactile modality is externally

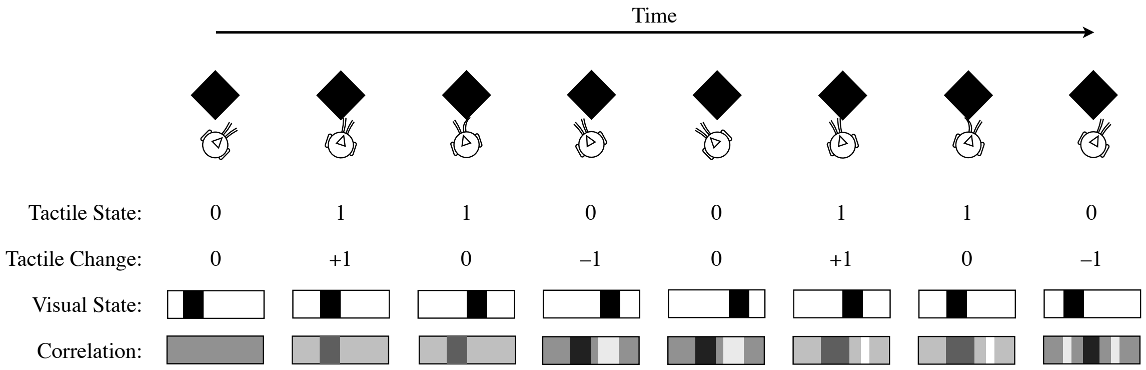


Figure 3-9: Correlation learned between tactile change and visual state as the whiskers as swept against an object (black box) in two different directions. The tactile change indicates the difference between the current and previous tactile state. The bottom row represents graphically the synaptic weights as they progressively learn a stable correlation between tactile change and visual input.

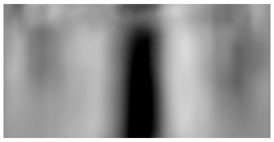


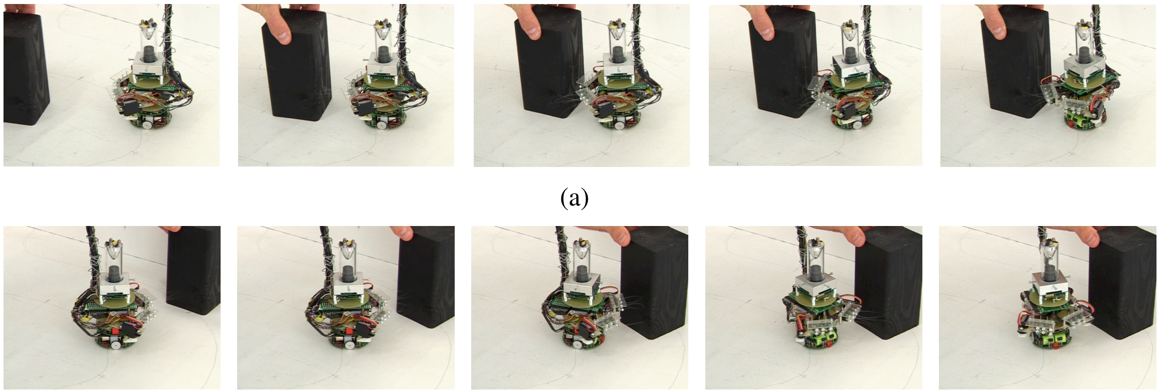
Figure 3-10: Graphical representation of the synaptic weights coupling the tactile modality to the visual modality, showing the visual image correlated to change in tactile activity.

activated, mimicking the idea that the robot "feels" some tactile stimulation.

3.4.1 Observed Behaviors

Initially, we observe that the robot does not move at all. Even if different objects are placed in, removed from or moved around the arena, the robot stays put. However, a first reaction is observed when an object is brought close enough to the robot to stimulate one of the whiskers. As soon as the robot detects contact with the object, it responds by turning away from it. If the object is more on the left of the visual field, the robot turns away to the right. If the object is more on the right, the robot turns away to the left (see Figure 3-11). This first behavior can be described as "contact avoidance".

Let us now observe what behavior is observed when neural activity is externally generated in the tactile modality. More specifically, we artificially increase the activity of the neuron corresponding to the "desired" state in the tactile modality. As soon as an object is placed in front of the robot, the latter starts moving towards the object, adjusting its orientation and moving towards the object until the tactile sensor detects contact with the object. If the object is displaced, the robot continues to follow it, approaching if it is moved away,



(a)

(b)

Figure 3-11: Contact avoidance. As soon as the robot detects contact with an object, it turns away either to the left (a) or to the right (b) until no contact is detected anymore.

backing away if it is moved closer and turning in the appropriate direction if the object is moved to the side. As long as the activity is maintained in the tactile modality, the robot never gives up, as if it were mysteriously attracted towards the object. We call this second behavior, illustrated in Figure 3-12, "object following".

3.4.2 Network Dynamics

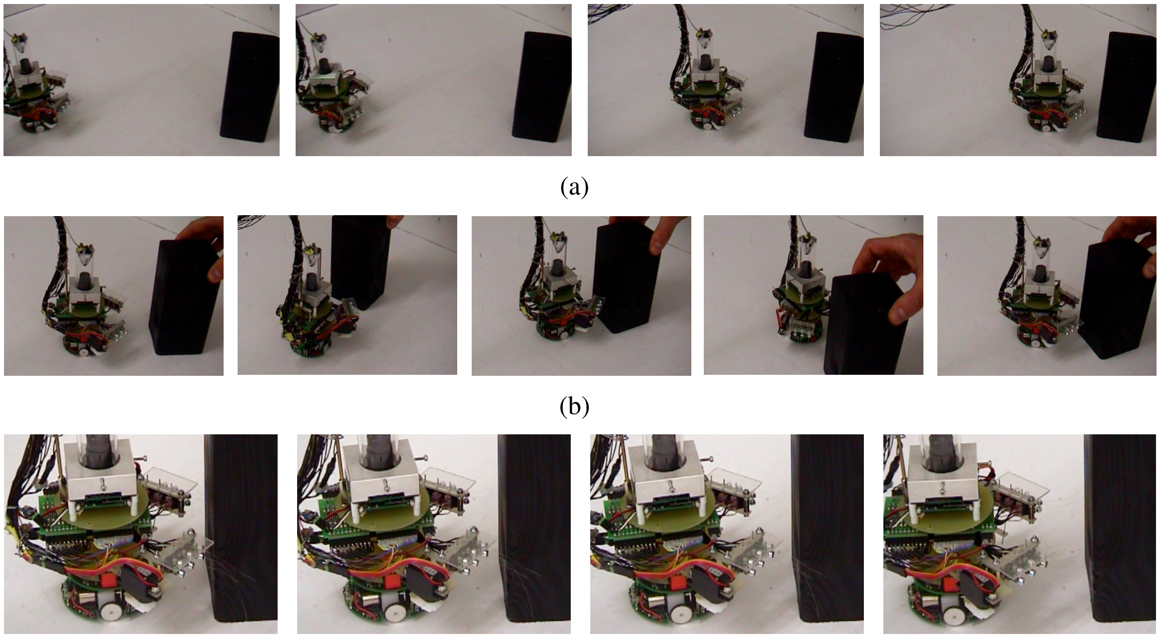
How do those two behaviors come about? This section analyses how neural activity generated in the neural network spreads through the different modalities, how the propagation of internal activity is modulated by the interaction with the environment, and how activity thus produced in the motor neurons creates the observable behaviors of the robot.

Let us first focus on the second behavior, namely object following. The activity of the neuron corresponding to the "desired"[2] tactile state is externally set to a positive value corresponding to tactile stimulation. If no input is detected from the whiskers, the difference of activity in the tactile modality between current and desired states will lead to positive activity in the desired change neuron. Because of the correlation between tactile change and visual state described previously, neural activity will propagate into the visual modality, projecting black pixels in the center of the visual field of the robot. If the robot perceives no real object, activity will not propagate any further and the robot will stay put (see Figure 3-13(a)).

Now, if an object is placed in front of the robot, the difference between current and desired visual states will elicit some optical flow in the corresponding visual population.

---

[2] As announced in the previous chapter, the quotation marks will not be used anymore to refer to this particular population of neurons.



(a)

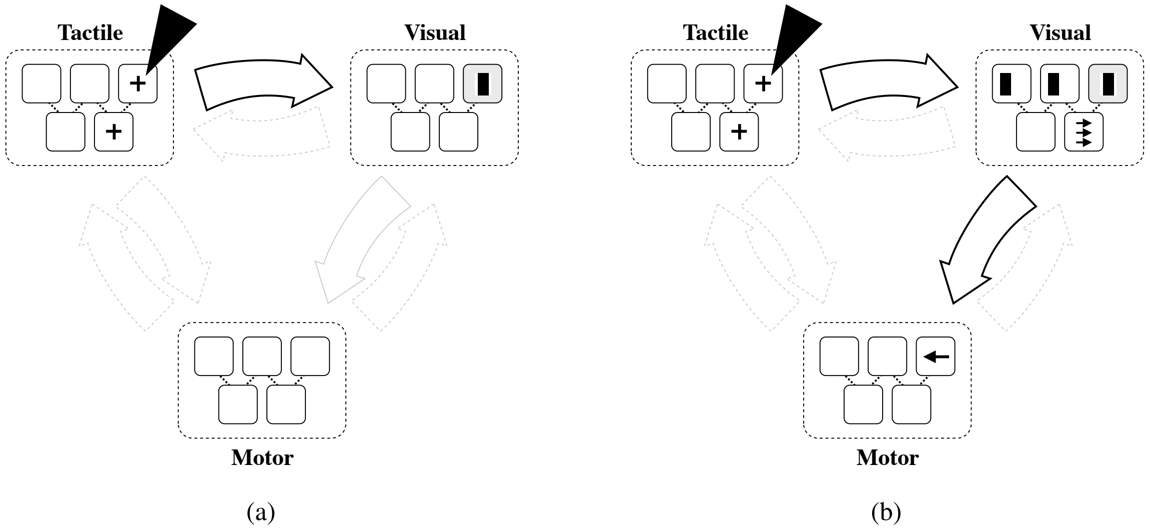
(b)

(c)

Figure 3-12: Object following. (a) The robot moves towards the object, adjusting both its relative position and orientation. (b) If the object is moved laterally, the robot turns in the corresponding direction. (c) The robot moves back if the object comes too close.

This neural activity will in turn propagate to the motor modality, thus producing a motion of the robot, as illustrated in Figure 3-13(b). If for instance the object is visible on the left of the visual field, the activity in the current visual population will reflect the dark pixels perceived on the left. The difference of activity between the current visual population (dark pixels on the left) and the desired visual population (dark pixels in the center) will elicit activity in the desired change population corresponding to an optical flow to the right – a pattern of visual change associated to motor activity producing a turning motion to the left. Interestingly, we see that the optical flow patterns generate exactly the appropriate activity in the motor neurons for the robot to approach and follow the perceived object. This explains the observed object following behavior.

The original contact avoidance behavior can be now easily explained. If no contact is detected, both the current and the desired state neurons have the same level of activity. No further activity is generated in the tactile modality, and the robot stays still. However, as soon as an object stimulates one whisker, the difference of activity between current and desired tactile states will generate activity in the desired change neuron that propagates exactly as described previously, but with the opposite sign, as illustrated by Figure 3-14. This time, the robot moves away from the object until no contact is detected anymore. This explains the observed contact avoidance behavior.



(a) (b)

Figure 3-13: Propagation of neural activity in the network leading to the observed "object following" behavior. The black arrow indicates activity generated externally in the "desired" population of the tactile modality. (a) Activity propagates into the visual modality. (b) If an object is visible, activity propagates further into the motor modality, generating the object following behavior. In this example, the object is perceived slightly on the left of the visual field. The activity generated in the motor modality produces a turning motion of the robot to the left.

3.5 Discussion

In this section, several aspects highlighted by this series of experiments are discussed in more details.

3.5.1 Observed Behaviors

First, the study conducted in this chapter shows that the robot can display various coherent behaviors even though the neural architecture is not specifically designed for these particular behaviors. There is no built-in reflex or any analogous predefined structure in the connectivity of the network: all modalities are homogeneously coupled together. The strength of the synaptic connections is only modified by correlation of neural activity captured by Hebbian learning as the robot interacts with the environment. In other words, the connectivity of the neural network, i.e. the way how the different sensors and motors interact internally with each other, arises only from the robot's own history of the perceived interaction with the real world.

The observed behaviors are global phenomena that are produced by the parallel interaction of several simple processes without any explicit coordination. This is an example of

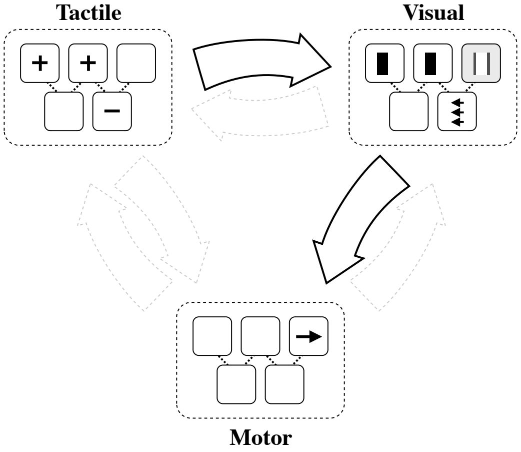


Figure 3-14: Propagation of neural activity in the network leading to the observed "contact avoidance" behavior. When contact is detected, neurons in the tactile modality are activated. The neural activity propagates into the visual modality, and then into the motor modality, generating the observed avoidance behavior. In this example, the object is perceived slightly on the left of the visual field, as in Figure 3-13. This time however, the activity generated in the motor modality produces a turning motion of the robot to the right.

self-organization process, an important concept encountered with complex systems (Thelen and Smith, 1994; Kelso, 1995; Kaneko and Tsuda, 2001). It is interesting to notice that in robotics, self-organization is mostly studied in situations involving multiple agents (collective or swarm intelligence). Our findings show that effects of self-organization can also be exploited to generate coherent behavior with a single agent.

3.5.2 Role of Embodiment

The results discussed in this chapter also illustrate the role of embodiment for the generation of the observed behaviors.

First, the robot has a particular morphology that shapes the dynamics of the agent-environment interaction. As the agent moves around in the arena, it produces various sensory stimulations (optical flow, tactile input). The particular choices for the motor representation (forward and turning speeds rather than rotation speed of the individual wheels) and for the definition of visual state change (optical flow rather than variation of pixel brightness) turn out to generate multiple cross-modal correlations that can be captured by simple Hebbian learning. This shows that an appropriate sensory motor system automatically reveals stable structures between different sensorimotor channels that shape the connectivity of the neural network. In other words, the particular morphology of the robot generates "good" sensorimotor data (i.e. correlated neural activity) that can be exploited by the neural architecture. This corroborates the recent concept of "information self-structuring" (Kuniyoshi et al., 2004; Lungarella et al., 2005; Iida and Pfeifer, 2005; Lungarella and Sporns, 2005, 2006; Olsson et al., 2006), which acknowledges that embodied agents do not passively absorb information from the surrounding environment, but rather actively shape and structure their sensory experience by dynamically interacting with the real world.

Second, the environment is a necessary component that allows the learned cross-modal correlations to be exploited for generating behavior. Let us consider as an example the contact avoidance behavior, where the robot turns away from an object in contact with its whiskers. In fact, this reflex-like behavior is not produced by a direct internal coupling between tactile and motor neurons: if only the whiskers are stimulated (e.g. with a transparent object), the robot does not display any avoidance behavior. The behavior results from the propagation of neural activity from the tactile modality to the motor modality through the visual modality, where the propagation is modulated by the visual perception of the environment. If no object is visible when the whiskers are stimulated, neural activity does not propagate further than the visual modality, and the robot does not move; activity only propagates further into the motor modality if the robot perceives an object visually (see Figure 3-13). In other words, this shows that the internal network dynamics is intrinsically coupled with the agent-environment interaction, and that the observed behaviors can only be accounted for by taking the whole system – including the environment – into account.

3.5.3 Material Properties

A closer inspection of the visuo-tactile correlation described previously unveils interesting effects resulting from the specific material properties of the tactile sensor.

In Section 3.3, we explained the correlation learned between change of tactile stimulation and visual input (see Figure 3-9). Let us now investigate what happens if we consider a robot equipped with a rigid tactile sensor (e.g. an infrared proximity sensor) instead of flexible whiskers. Figure 3-15 shows that on average, the position of the object is the same when an onset or an offset of tactile stimulation is detected: as a result of this symmetry, there is no significant correlation between tactile and visual activity anymore. As a consequence, if the robot is equipped with a rigid tactile sensor, it will not display any of the behaviors observed otherwise.

This shows how the flexibility of the whiskers used as a tactile sensor in fact breaks the symmetry in the sensorimotor interaction, and thus creates a cross-modal sensory correlation that is captured and exploited by the neural network to produce the observed behaviors.

This fact highlights the importance of the role played by the material properties of

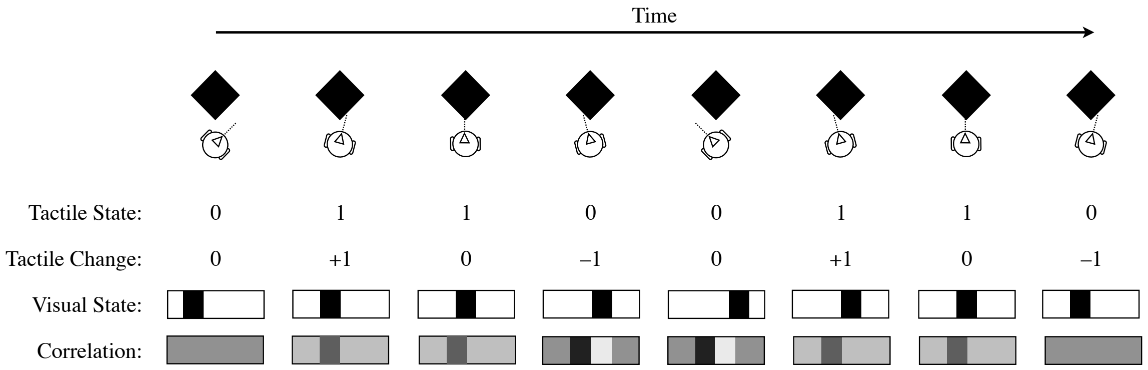


Figure 3-15: Correlation learned between tactile change and visual state. Compared to Figure 3-9, the only difference is that the flexible tactile sensor is now replaced by an undeformable proximity sensor, whose range is indicated by the dotted line in the illustrations of the top row. As a consequence, the synaptic weights do not learn any stable correlation between tactile change and visual state, as illustrated in the bottom row.

whiskers (Lungarella et al., 2002b; Bovet et al., 2004). For instance, a simulation of the experiments discussed in this chapter would simply fail to reproduce the observed behaviors if it would not include the fact that the tactile sensors bend while in contact with an object.

3.5.4 Generation of New Behavior

It is worth pointing out the fact that the agent is actually able to generate new behaviors, i.e. behaviors that has never been displayed previously.

This argument can better be understood if we consider the following variant of the experiment (see Figure 3-16). In a first phase, the robot is not driven but stays put, and objects are randomly moved around the robot, close enough to arbitrarily stimulate the tactile sensor of the robot. Obviously, only the visuo-tactile correlation described previously will be learned by the neural network. In a second phase, the objects are moved aside, and the robot is driven randomly across the empty arena. Consequently, the visuo-motor correlation described previously will now be learned by the neural network (the agent always detect some optical flow when it is moving, since the visual background is not perfectly uniform). Since the connectivity of the neural network is then qualitatively the same as in the original variant of the experiment, the same behaviors will be observed as soon as the robot is let to move on its own.

This demonstrates that the agent is able to spontaneously generate the contact avoidance and object following behaviors, without having experienced these behaviors previously, i.e. without having ever before moved towards or away from any object. Metaphorically speaking, the robot is not reinforcing or optimizing behaviors that occurred during the random

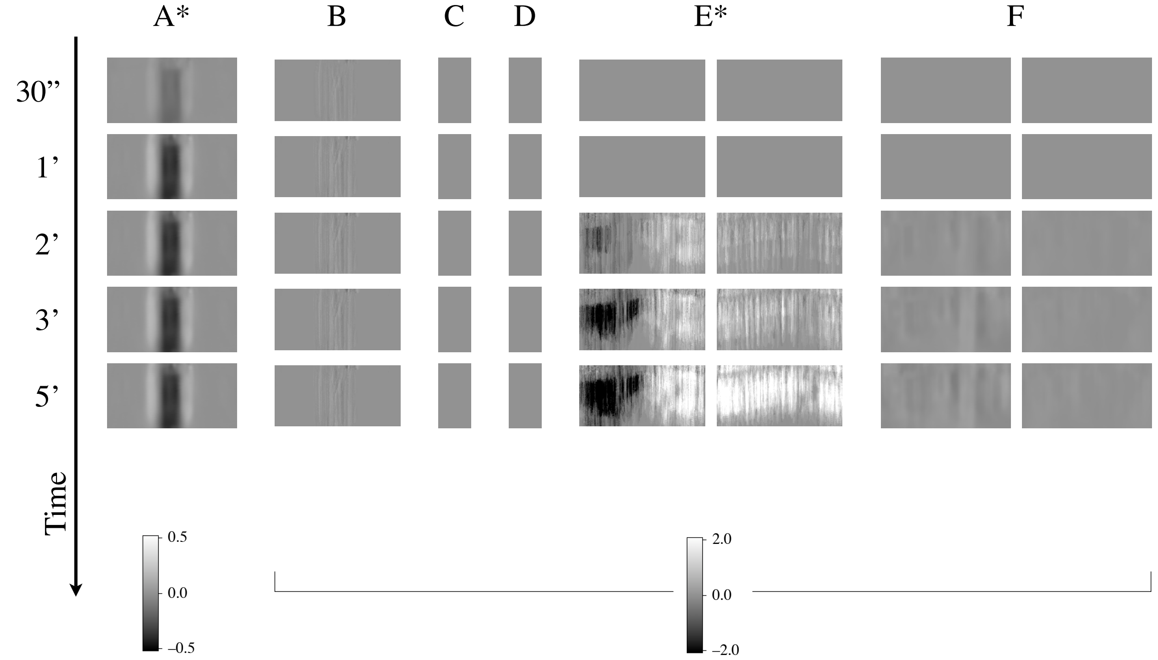


Figure 3-16: Time evolution of the synaptic weights during a variant of the initial exploration phase (compare with Figure 3-6). During the first minute (two first rows), the whiskers are stimulated by randomly moving objects, and the neural network only captures a correlation between tactile and visual activity (A). Thereafter, the robot is randomly driven across the empty arena; the neural network captures then a correlation between visual and motor activity (E). The captured correlations are qualitatively the same as in the previous condition (Figure 3-6).

exploration phase. Rather, it is able to somehow combine basic multi-modal associations – such as between touching and seeing a close object, and between moving forward and perceiving an expanding image – to create new, seemingly intentional behaviors.

3.5.5 Time Perspectives

This leads us to a consideration of the obtained results from different time perspectives. The results show that depending on the internal state of the desired tactile neuron, the robot displays behaviors that can be designated as either contact avoidance or object following.

At a time scale of several seconds, these behaviors look somehow intentional. Indeed, the motions of the robot generated by the neural network can be interpreted as “actions” performed to eventually reach the “goal state” indicated by the desired tactile neuron: if its activity is low, the robot turns away from objects to reduce tactile stimulation; if it is high, the robot moves towards objects to finally increase tactile stimulation. Note that the interval between the onset of the motion and the final state is typically a few seconds, and can be arbitrarily longer – for instance, if the object is continuously moved away from the robot following it.

However, the internal dynamics of the neural network takes place at a different, much shorter time scale. The cross-modal couplings only capture simultaneous correlations, i.e. sensory motor events occurring at the same time, and neural activity propagates through the network in less that a second[3].

How are these two different time scales related to each other? There are two main factors. On the one hand, the synaptic weights, even though they are continuously modified by the Hebbian plasticity, can learn and maintain significant correlations over time. Obviously, information acquired by the synaptic coupling remains available on a much longer time scale, thus coupling effects observable from both instantaneous and development (ontogenetic) perspectives. On the other hand, most physical processes governing the interaction of the robot with the environment are of continuous nature. For instance, the image of an object in front of the robot is steadily growing as the robot moves forward. As the robot approaches the object, a qualitatively stable sensory structure – namely, the expanding optical flow – is maintained over an extended period of time. This suggests that learning associations between sensory motor states and temporal fluctuations thereof (i.e. change of state between two successive time steps) is sufficient to grab some long-term causal relations between motor actions and sensory perception. This issue will be further developed in Chapter 7.

---

[3] The neural network of the robot is simulated with a cycle rate of approximately 10 Hz.

3.6 Conclusion

This chapter describes an initial series of experiments performed with a mobile robot using the homogeneous neural architecture introduced in the previous chapter. It illustrates how different coherent behaviors – contact avoidance and object following – can be generated by a neural model only learning cross-modal correlations, i.e. without any built-in reflexes or reinforcement mechanisms. It investigates the role played by the embodiment of the system, by the material properties of its tactile sensors, and by the interaction with the environment at different time scales. Finally, it discusses how an observer perspective reveals several interesting insights on how the agent is able to produce new, seemingly intentional behaviors.

Appendix 3.A
Estimation of Optical Flow

For most sensory or motor modalities of the proposed neural architecture, the state change is simply defined as the component-wise difference between two successive states, and thus indicates the temporal variation of the state. By using a slightly modified definition for the state change of the visual modality, we can obtain an estimation of the perceived optical flow, as explained in this appendix.

Consider first a one-dimensional, continuous visual state $I(x,t)$. Suppose that the visual input is a constant pattern given by a function $f(x)$, steadily moving with velocity v. The visual state at time t is thus:

$$I(x,t) = f(x - vt) \tag{3.1}$$

Using the chain rule, the temporal variation of the visual state at position x is given by:

$$\begin{aligned} \dot{I}(x,t) &= \frac{\partial}{\partial t} f(x - vt) \\ &= \frac{\partial}{\partial x} f(x - vt) \cdot (-v) \\ &= \nabla I(x,t) \cdot (-v) \end{aligned} \tag{3.2}$$

Assuming a non-zero gradient $\nabla I(x,t)$, the speed at which the pattern is moving – i.e. the optical flow – can be obtained as:

$$v = -\frac{\dot{I}(x,t)}{\nabla I(x,t)} \tag{3.3}$$

This leads to the following approximation in the discrete case. The estimated optical

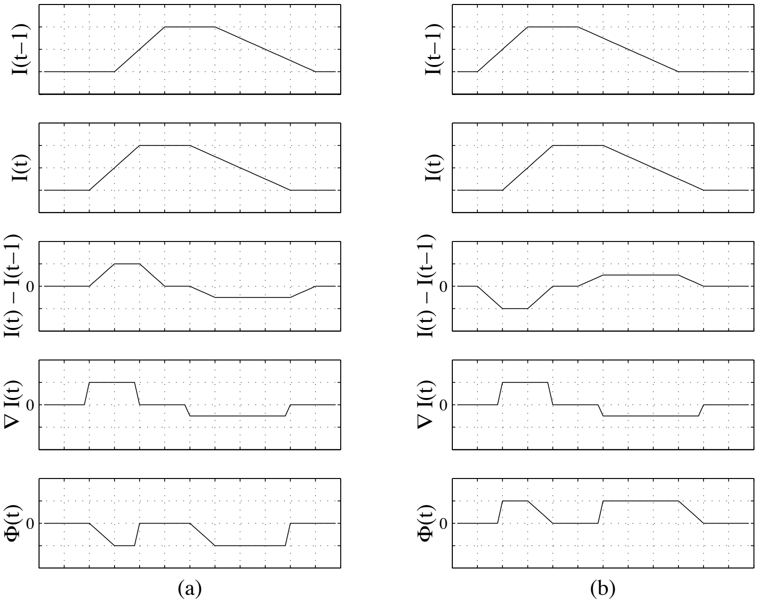


Figure 3-17: Approximation of optical flow between two successive visual states $I(t-1)$ and $I(t)$, shown in the top rows. The bottom row shows the estimated optical flow $\Phi(t)$. The visual input is either moving (a) to the left or (b) to the right.

flow $\Phi(x,t)$ at position x and time step t is defined as:

$$\Phi(x,t) := -\frac{I(x,t) - I(x,t-1)}{\nabla I(x,t)} \tag{3.4}$$

with the gradient

$$\nabla I(x,t) := \frac{1}{2h} \cdot (I(x+h,t) - I(x-h,t)), \quad h > 0 \tag{3.5}$$

The distance h over which the gradient is estimated is typically $h = 1$, but can be increased if the image is noisy. To avoid numerical instabilities, we set $\Phi(x,t) := 0$ if $|\nabla I(x,t)| < \epsilon$. The threshold ϵ is typically set, for visual states in the range $-1 \leq I(x,t) \leq 1$, to $\epsilon = 0.1$. Figure 3-17 shows that this definition indeed provides an acceptable approximation of the perceived optical flow for moving one-dimensional visual inputs.

The one-dimensional estimation of the optical flow (3.4) is generalized in a straightforward way to two-dimensional visual input. Let $I(x,y,t)$ be the value of the pixel at position (x,y) and time step t. The optical flow field $\Phi(x,y,t) = \{\Phi_x(x,y,t), \Phi_y(x,y,t)\}$ is given by the following estimations for the horizontal and vertical components of the optical flow:

$$\Phi_x(x,y,t) \;:=\; -\frac{I(x,y,t)-I(x,y,t-1)}{\nabla_x I(x,y,t)} \tag{3.6}$$

$$\Phi_y(x,y,t) \;:=\; -\frac{I(x,y,t)-I(x,y,t-1)}{\nabla_y I(x,y,t)} \tag{3.7}$$

with the gradients in the horizontal and vertical directions

$$\nabla_x I(x,y,t) \;:=\; \frac{1}{2h}\cdot\left(I(x+h,y,t)-I(x-h,y,t)\right) \tag{3.8}$$

$$\nabla_y I(x,y,t) \;:=\; \frac{1}{2h}\cdot\left(I(x,y+h,t)-I(x,y-h,t)\right) \tag{3.9}$$

Similar to above, we set $\Phi_{x/y}(t) := 0$ if $\left|\nabla_{x/y} I(t)\right| < \epsilon$ to avoid numerical instabilities. An illustration of the optical flow field estimated with this definition is provided in Figure 3-4 on page 43. It shows that, at least for slowing varying visual inputs, this estimation indeed provides an acceptable approximation of the perceived optical flow.

Note that the present definition of optical flow field is a straightforward generalization of a trivial solution for the one-dimensional problem. It is based on the assumption that the brightness of the moving pattern does not change:

$$\begin{aligned} 0 &= \frac{d}{dt} I(\vec{x}(t)) \\ &= \frac{\partial}{\partial t} I(\vec{x},t) + \frac{\partial \vec{x}}{\partial t}\cdot \nabla I(\vec{x},t) \\ &= \frac{\partial}{\partial t} I(\vec{x},t) + \vec{v}\cdot \nabla I(\vec{x},t) \end{aligned} \tag{3.10}$$

Even though this equation can be solved in the one-dimensional case (3.3), it becomes an ill-posed problem in two dimensions, since there are then two unknowns (the components of $\vec{v}$) for only one equation. We thus cannot expect our estimation to provide an accurate solution.

There is a large body of literature dealing with the estimation of two-dimensional optical flow (see e.g. Horn and Schunck, 1994; Barron et al., 1994; Lappe, 2000; Fermüller et al., 2001). Typically, gradient-based methods solve (3.10) under additional constraints, such as smoothness of the optical flow (Horn and Schunck, 1981). As a consequence, the solutions are generally iterative, and thus computationally much more expensive than our estimation.

In contrast, our method provides a computationally very inexpensive way to assess local apparent displacements between two consecutive images. Like with elementary motion detectors (Reichardt, 1969; Borst and Egelhaaf, 1993), the local estimated velocities do not necessarily have to always match the actual velocity of the moving pattern.[4] However, the *combination* of several such local estimations generally lead to an acceptable global estimation.

---

[4] It can easily be shown that $\vec{v} = -\frac{1}{2}(\frac{\partial_t I}{\partial_x I}, \frac{\partial_t I}{\partial_y I})$, which is up to the factor $\frac{1}{2}$ the continuous version of our estimated flow, is indeed, for non-vanishing components of the gradient, a solution of (3.10).

Chapter 4

Insect Navigation Strategies

This chapter studies the behaviors that are observed when the homogeneous neural architecture described in the previous chapters is employed on a different mobile robot. We show that the behaviors observed with the robot situated in a desert-like environment closely match distinct natural behaviors observed in desert ants. Finally, we discuss how this novel approach reveals original and promising insights in embodied artificial intelligence and biologically inspired robotics.

4.1 Introduction

Insects, despite their tiny brains, can exhibit remarkable navigation capabilities. For example, bees and ants regularly depart from and return to fixed positions in their environment, covering distances up to several hundreds of meters. It has been shown that these insects can resort to two fundamentally different navigation strategies. The first one is referred to as path integration (for reviews, see e.g. Wehner and Srinivasan, 2003; Etienne and Jeffrey, 2004): all angles steered and all distances covered are integrated into a mean home vector that monitors the current position of the animal relative to the nest. The second navigation strategy relies on visual landmarks, and is referred to as visual homing or landmark navigation (Collett, 1992): it is assumed that the insect stores a rather unprocessed visual snapshot of the scene around the nest; by matching this snapshot to the current retinal image, the insect can derive the direction it has to move in order to relocate the target position where the snapshot was taken.

In recent years, several models were successfully employed on mobile robots to implement and test hypotheses about mechanisms underlying the navigation capabilities of insects (Franz and Mallot, 2000). As a matter of fact, all models suggest that distinct processes, or specialized neural structures are required for each observed navigation strategy.

However, the existence of such processes in natural systems is still an open question, especially since up to now only very little is known about the neural circuits in the animals' nervous systems.

In this chapter, we show that both path integration and landmark homing strategies can actually be obtained with the same neural architecture as the one used in the previous experiments, using a mobile robot equipped with a particular sensory-motor system and engaged in a desert-like environment.

The outline of the chapter is as follows. First, Section 4.2 briefly reviews different models of navigation strategies observed in insects. In Section 4.3, the agent and the environment used for our experiments are described, followed in Section 4.4 by a detailed investigation of the various observed behaviors. Section 4.5 discusses various insights gained from the obtained results, which are concluded in Section 4.6.

4.2 Models of Insect Navigation

This section introduces different navigation strategies observed in insects, and briefly outlines the essential features of the models proposed in the literature.

An outstanding and well-studied example is the Saharan desert ant, *Cataglyphis fortis* (Wehner, 2003). This insect goes for foraging tours of several hundreds of meters and finds is way back to the nest entrance, which is an inconspicuous hole in the desert ground. Since Cataglyphis lives in the desert, it cannot resort to strategies that involve following pheromone trails deposited on the ground. Those would evaporate immediately in the desert heat.

4.2.1 Path Integration

The predominant homing strategy observed with Cataglyphis is a mode of dead-reckoning known as path integration (Mittelstaedt, 1983; Müller and Wehner, 1988; Collett and Collett, 2000; Andel and Wehner, 2004). During the tortuous outbound journey performed when searching for food in the flat desert habitat, which is often completely devoid of landmarks, the ant monitors its current position relative to the nest by a path-integration process. The resulting global vector that continuously connects the ant with its starting point enables – reversed in sign – the insect to find its way home and to return to the nest along an amazingly straight path. This navigation strategy is also referred to as egocentric strategy (Wehner et al., 1996).

The task is accomplished by integrating information about both the angular and linear components of movement, information provided by a compass and an odometer, respectively. Information about directions steered is obtained via the ant's celestial compass, which exploits polarization and spectral gradients in the sky (Wehner, 1994). On the other hand, attempts to unravel the ant's odometer have so far met with limited success. There are several hypothesis about the cues by which the ants measure the travel distance. The "energy hypothesis," of long standing in arthropod research (Heran and Wanke, 1952), posits that travel distance is estimated from energy consumption. However, ants have been shown to assess their walking distance with great accuracy, irrespective of the load they carry (Wehner, 1992). The "optic flow hypothesis" has been proven in honeybees (Esch and Burns, 1995; Srinivasan et al., 2000), but plays only a minor role, if any, in Cataglyphis ants (Ronacher and Wehner, 1995; Ronacher et al., 2000; Thiélin-Bescond and Beugnon, 2005). Thus, ants appear to rely primarily on proprioceptive cues, most probably derived from the movements of their legs (Wittlinger et al., 2006), rather than on external cues. However, ants must use more sophisticated means than merely counting the number of steps, since they have been shown to gauge ground distance with a high precision when walking over hilly terrain (Wohlgemuth et al., 2001).

In the past decade, several biologically inspired models of path integration have been implemented with autonomous robots. Visual odometers for mobile robots were proposed based on image interpolation (Srinivasan et al., 1999) or optic flow (Weber et al., 1996; Iida and Lambrinos, 2000). Lambrinos et al. (2000) built a series of robots to model navigation behaviors of the Saharan desert ant Cataglyphis, and compared their performance by conducting the experiments in the same natural habitat as that of the desert ant. Travel distance was estimated from the wheel encoder signals, and the polarization pattern of the sky was used as a compass. The path integration accuracy of the robot was surprisingly high, in a range comparable to that of real desert ants.

4.2.2 Landmark Navigation

In bees and ants, path integration employing a skylight compass is the predominant mechanism of navigation (Wehner et al., 1996). However, the path-integration process is error prone, and slight deviations result in a wrong estimate of the nest position (Wehner and Wehner, 1986). Since the nest entrance of the Cataglyphis ant is an inconspicuous hole in the desert ground, which is invisible to the insect even from small distance, alternative strategies have to be employed in order to finally locate the entrance. In the absence of visual landmarks, Cataglyphis ants start a systematic search around the position where the nest is expected after having reset their path-integration system (Wehner and Srinivasan, 1981; Müller and Wehner, 1994). However, when landmark information is available, both bees and ants exploit it (Collett, 1992; Wehner, 2003), and relocate the target position directly with a remarkable precision. This kind of navigation strategy is also referred to as geocentric strategy (Wehner et al., 1996).

Several experiments with bees (Cartwright and Collett, 1983; Brünnert et al., 1994) and ants (Wehner and Räber, 1979; Wehner et al., 1996) suggest that the animal stores a rather unprocessed visual snapshot of the scene around the nest, and by matching this snapshot with the current retinal image, derives the direction it has to move in order to relocate the target position where the snapshot was taken.

Different models for the matching of snapshot and current view can be found in the literature (for a review, see Franz et al., 1998). We mention here but a few. Cartwright and Collett (1983) suggested the so-called snapshot model, which reproduces some aspects of the search behavior of bees. This model assumes that the views are omnidirectional and aligned with an external reference direction. Landmarks appear as black sectors in the views, the background being white. Each landmark in the current view is matched to the closest black sector in the snapshot, and a movement direction is computed to reduce the perceived difference. The movement directions are summed up for all landmarks, and give the resulting movement direction of the model insect. The proportional vector model (Möller et al., 1998) is a variation of the snapshot model, where the difference in apparent size and bearing between the paired sectors is taken into account.

Lambrinos (1999) suggested an equivalent, but computationally more parsimonious model, the average landmark vector model. In this model, each visual landmark feature (e.g. the centers or the edges of the sectors) is associated with a unit vector pointing in the corresponding direction. For each view, all landmark vectors are averaged to produce the average landmark vector. The movement direction is given by the difference vector obtained by subtracting the stored average landmark vector corresponding to the home snapshot from the average landmark vector computed from the current view. The elegant simplicity of this model allowed Möller (2000) to implement it on an entirely analogue robot – i.e. without any digital electronic component – that could perform visual homing in an arena with white walls and black landmarks.

4.3 Agent and Environment

The experiments described in this chapter are performed both in simulation and in the real world with the “Samurai” mobile robot shown in Figure 4-1, using the same homogeneous neural architecture as in the previous chapters.

Parameter	Description	Approx. Value
H	Height	42 cm
D	Diameter	30 cm
m	Mass of the robot	6 kg

Figure 4-1: Left: The mobile robot "Samurai" used for the real world experiments, equipped with an omnidirectional camera and differential steering. Right: Mechanical specifications of the robot.

The agent is situated in a flat environment similar to the ecological niche of a desert ant. The home location – corresponding to the "nest" – consists of a dark hole in the ground. An external light and heat source – similar to the sun – shines upon the whole surroundings. The agent is equipped with the following sensory motor modalities:

Omnidirectional Camera

The one-dimensional 360° view is taken by thresholding, averaging and low-pass filtering a portion of the polar transformed image from the camera, as shown in Figure 4-2. The polar image is aligned to a geocentric coordinate system by using a compass[1]. As in the previous chapter, the state change in the visual modality is defined as the estimated perceived optical flow.

Temperature Sensor

We assume that the agent out of its nest is constantly exposed to the external heat source. Its temperature thus constantly increases with time. We choose an increasing rate of 0.1 s^{-1}.

Ambient Light Sensor

This binary sensor only detects light when the agent is outside the nest hole[2]. The state change in this modality is defined as the absolute value of the difference between two

---

[1] Due to the high level of electromagnetic perturbation in the robotic arena, which prevents the use of a magnetic compass, the compass value is obtained by triangulation, using the omnidirectional camera image and three salient landmarks of known positions.

[2] Obviously – no hole were drilled in the floor of the indoor office arena! – temperature and ambient light sensors are simulated on the real robot.

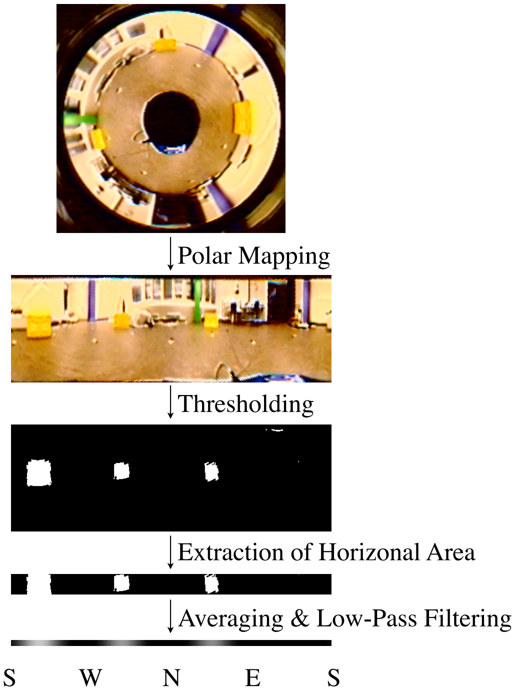


Figure 4-2: Image processing of the visual input. The image from the 360° camera (top) is used to extract the visual input for the system (300×1 pixel, bottom).

states, thus indicating with a positive value both the onset and offset of detected ambient light.[3]

Motors

The movement of the mobile robot is represented in the motor modality using a neural population coding, assumed to be more robust under noisy conditions and also more biologically plausible than direct encoding (Georgopoulos et al., 1986; Levi and Camhi, 2000). The population coding consists of several neurons (24 in this series of experiments) representing the magnitude of vectors in regularly spaced and uniformly distributed directions. The movement of the robot is determined by summing up all vectors (see Figure 4-3), the azimuth and length of the resulting vector indicating the heading direction and speed of the robot, respectively. A unit vector corresponds to a robot speed of approx. 20 cm/s, and the angular speed of the robot is defined as $\omega = \dot{\phi}_{\text{robot}} = -c_1 \cdot \tanh(c_2 \cdot (\phi_{\text{vector}} - \phi_{\text{robot}}))$ with $c_1 = 15\,^{\circ}/\text{s}$ and $c_2 = 10$.

---

[3] The reason for choosing an absolute difference value will be discussed in the next section.

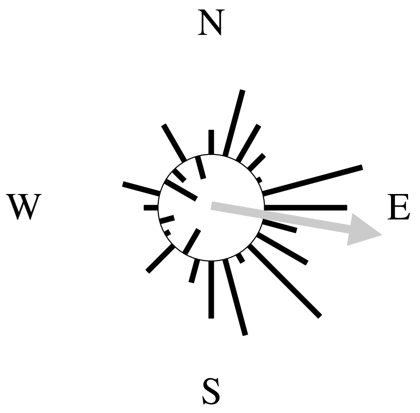


Figure 4-3: Population coding in the motor modality. Each component is represented by a radial bar whose length indicates the magnitude, and whose position indicates the direction. The gray arrow represents the resulting direction of movement of the robot, obtained by summing up all component vectors.

4.4 Observed Behaviors

As discussed in the previous chapters, the robot has to be randomly stimulated during the initial exploration phase in order to learn basic cross-modal correlations that will shape the connectivity of the network. The robot is thus first randomly driven across the arena, producing some kind of "foraging excursion". The robot is then let to move on its own.

In what follows, we study the behaviors observed when the activity of the desired state neuron in either the temperature or ambient light sensory modality is arbitrarily decreased – mimicking the idea that the robot "feels" like in a cooler or darker place, respectively. The striking result is that in both cases, the robot spontaneously returns to the starting position. In the first case, the robot relocates the home position following a straight trajectory. In the second case, the robot uses the visual landmarks of the environment and goes back home from almost any release position.

The time evolution of the weight matrices corresponding to all cross-modal synaptic coupling in the neural architecture are given, for the first robotic experiment, in Figure 4-4.

4.4.1 Path-Integration Homing

This section studies the behavior displayed when the activity of the desired temperature neuron is decreased. The resulting trajectories – obtained either in simulation or with the real robot – are shown in Figure 4-5 (a), (b), (d) and (e): the agent returns in an almost straight course to the starting location and stops there. If the agent is first displaced and

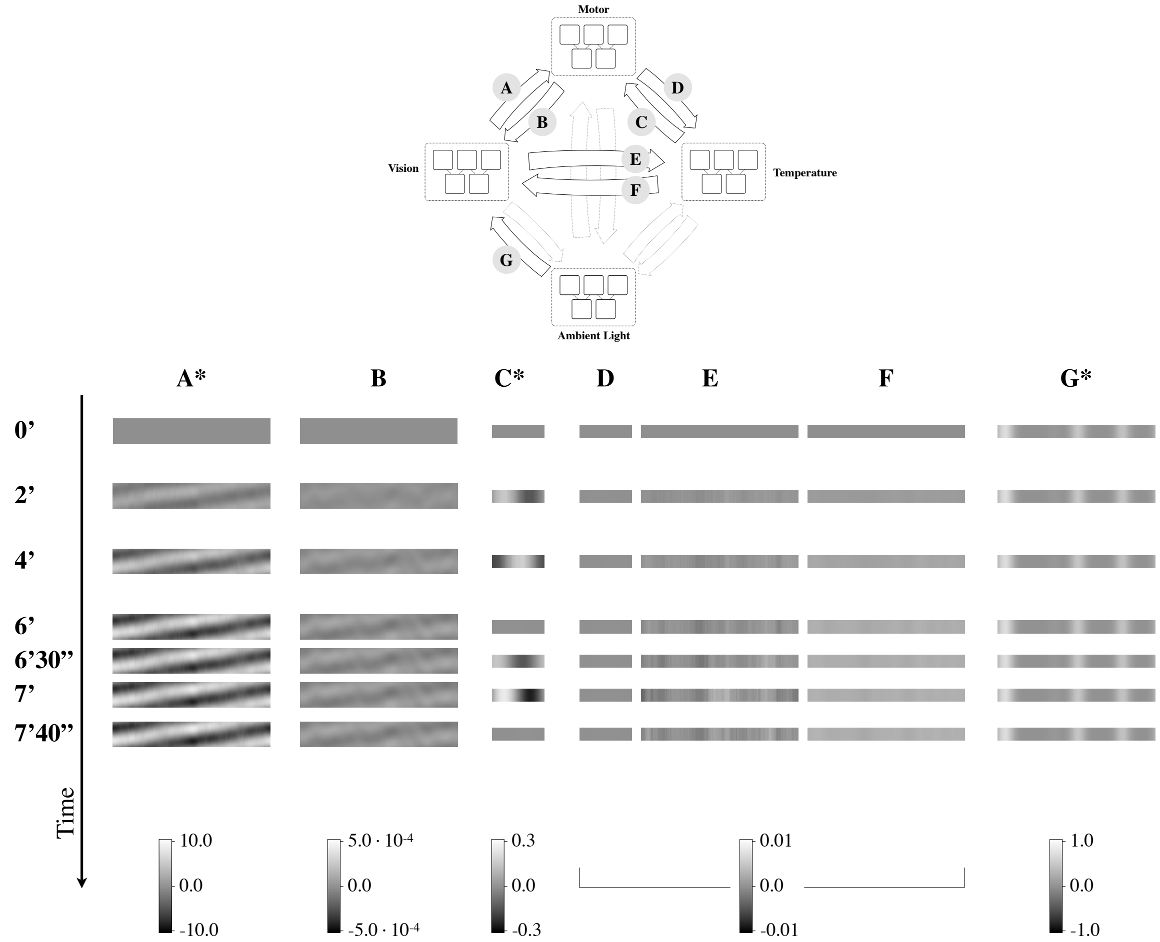
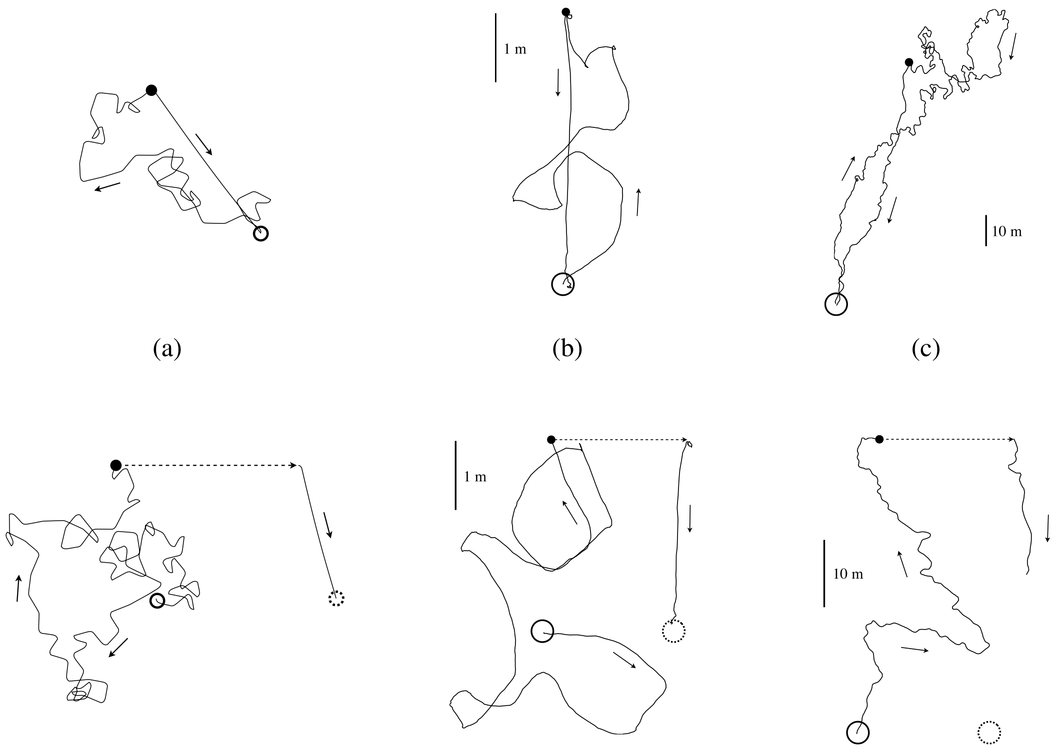


Figure 4-4: Time evolution of the synaptic weights. Top: Schematic illustration of the neural network. Bottom: Weight matrices corresponding to the cross-modal synaptic couplings indicated above. The synaptic couplings that capture a significant correlation are indicated with asterisks. The weight matrices of the remaining cross-modal coupling, which for the sake of clarity are not displayed graphically, have zero values during the whole experiment. The robot is randomly driven across the arena during the first 7 minutes. At this point, the activity of the desired temperature neuron is reduced, and the robot is let to move on its own. The trajectory of the robot between 6’ and 7’40” is shown in Figure 4-5(b). The learning and forgetting rates of the network are $\eta = 0.01$ and $\varepsilon = 0$ respectively, except for the synaptic coupling G, where the learning rate is increased to $\eta_G = 1.0$ (as there is only one time step in the ambient light modality – the first one – where the state change neuron displays some activity).



(a) (b) (c)

(d) (e) (f)

Figure 4-5: Trajectories of the agent corresponding to a path integration homing behavior. The solid circle indicates the home location. The dot specifies the location where the "homing behavior" is triggered by reducing the activity of the desired state neuron in the temperature modality. In (d), (e) and (f), the agent was displaced to a different location (dotted arrow); when released, the agent returns to a fictive home location corresponding to the integrated path (dotted circle). (a) and (d) show trajectories obtained in simulation, (b) and (e) plot trajectories of the real robot. (c) and (f) show for comparison trajectories of a real desert ant (adapted from Müller and Wehner, 1988).

released from a different location, it returns to a fictive home location corresponding to the integrated path. This behavior thus matches quite exactly what is described in the literature as path integration homing strategy (see Figure 4-5 (c) and (f)).

A consideration of the coupling between the motor and temperature modalities provides an explanation for this behavior. In the described environment, the value of the temperature sensor constantly increases, meaning that the state change neuron in the temperature modality displays a constant activity. This leads to the fact that the Hebbian learning rule modifying the synaptic weights of the connections from the temperature modality acts in fact like an integrator. Consequently, the values of the weights connecting the temperature modality to the motor modality are the time-integrated values of the motor commands, i.e.

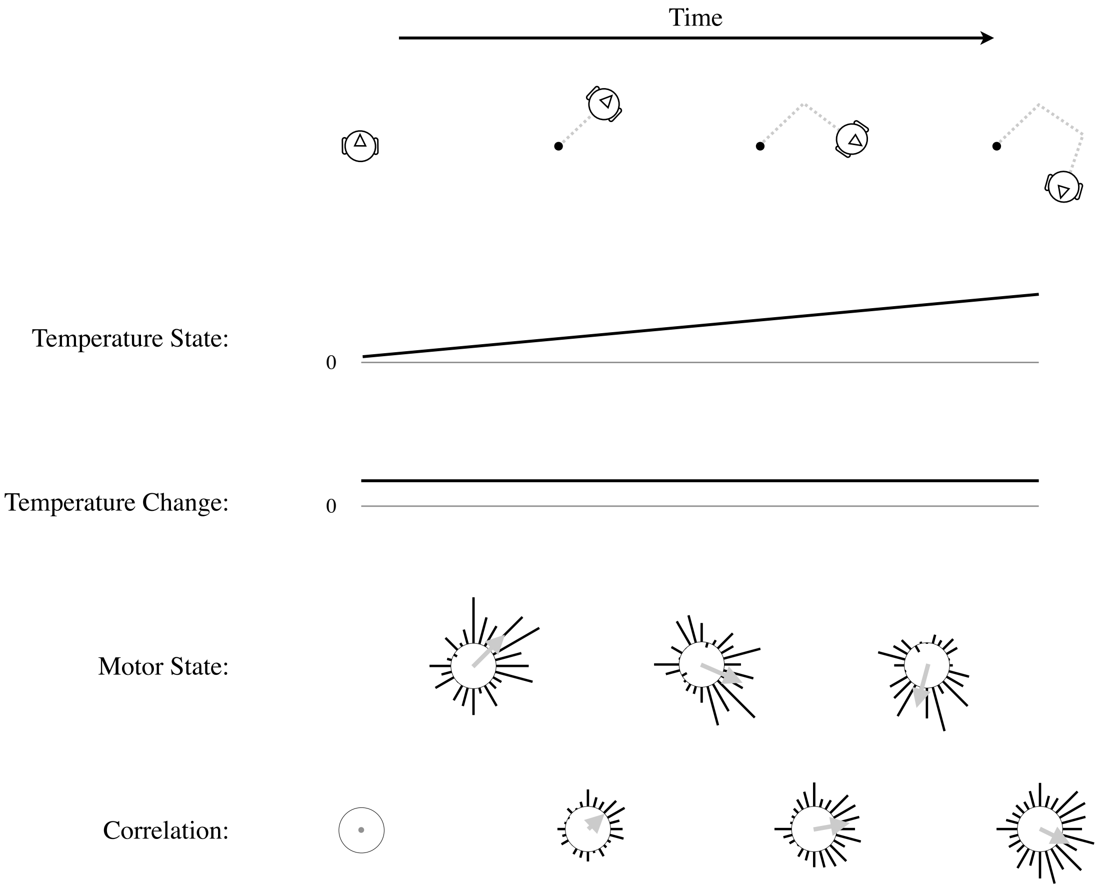


Figure 4-6: Integrated path as correlation between temperature change and motor activity. From top to bottom: As the robot randomly moves around, the temperature constantly increases. Therefore, the correlation learned progressively between the constant temperature change and motor activity represents the integrated path, i.e. the current position relative to the home position, as shown in the bottom row.

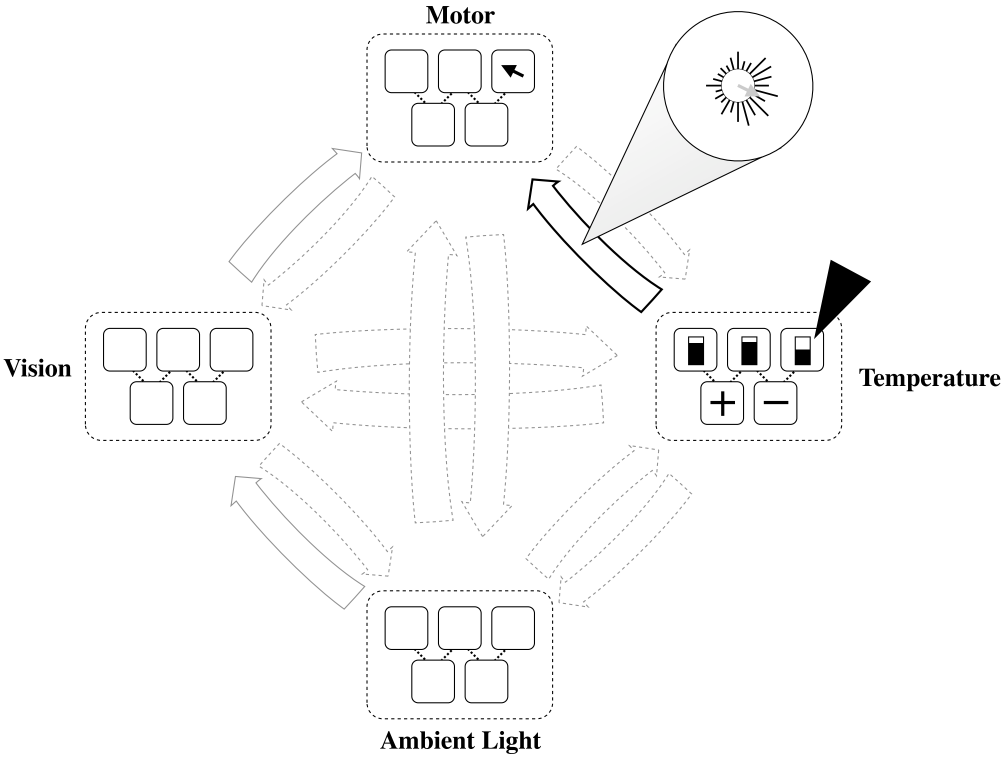


Figure 4-7: Generation of path integration behavior. When the activity of the desired temperature neuron is externally decreased (indicated by the black arrow), neural activity propagates into the motor modality. The activity of the neuron corresponding to desired temperature change is negative, and the synaptic weights coupling the temperature to the motor modality represent the integrated path (see Figure 4-6). The activity generated in the motor modality therefore corresponds to a motion in the opposite direction of the integrated path, i.e. in the direction of the home location.

the actual integrated path (see Figure 4-6). Activity generated in the temperature modality, which corresponds to a low desired temperature, propagates into the motor modality according to the integrated path – with reversed sign –, thus generating the homing behavior (see Figure 4-7).

Note that homing behavior is observed because the motor modality uses a geocentric population coding representation, rather than a direct encoding such as the forward and turning speed of the robot, or the velocity of the individual wheels. Indeed, it can easily be seen that the time-integrated motor activities only correspond to the integrated path in the case of a geocentric population coding[4] (see Figure 4-6).

4.4.2 Visual Homing

This section studies the behavior displayed when the activity of the desired ambient light neuron is decreased. Figure 4-8 (a) and (b) show that, irrespective of the release position, the agent returns to the home position and stops there. This behavior corresponds closely to what is observed with traditional visual landmark homing strategies (see Figure 4-8(c)).

It is worth noticing that agent does not necessarily return along a straight trajectory and most of the time, the agent avoids obstacles on its route – an additional emergent behavior already observed with traditional visual homing strategies (Lambrinos et al., 2000; Hafner et al., 2002).

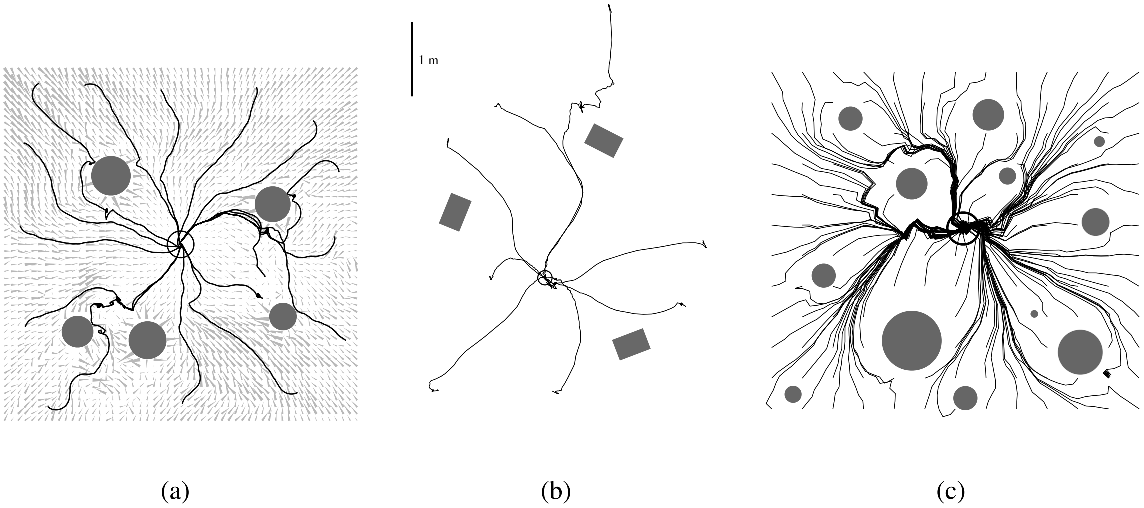
An explanation for this homing behavior requires here to consider neural activity across three modalities, namely the ambient light, the vision and the motor modalities. The neuron in the ambient light modality corresponding to the change of detected ambient light is only active at the beginning of the foraging excursion, namely when the agent exits its dark hole and surfaces into the bright external environment. The effect is that the Hebbian learning rule will thus capture a "snapshot" of the states of all the other modalities in the weights of the synapses leaving the ambient light modality. In particular[5], the synaptic weights connecting the ambient light modality to the visual modality capture a "snapshot" of the visual state from the home location (see Figure 4-9(a)).

Therefore, when the activity of the desired state neuron in the ambient light modality is arbitrarily set to a low value (corresponding to darkness), the synaptic coupling will propagate activity into the desired state population of the visual modality, projecting the view of

---

[4] Consider for instance two agents starting from the same initial position. The first makes a straight motion followed by a U-turn, whereas the second starts with the U-turn followed by the straight motion. Obviously, both agents have the same integrated forward and turning velocities, but are located at completely different locations.

[5] We assume that when the detected ambient light increases, before the foraging excursion, the agent is not moving yet and the detected temperature is arbitrarily set to zero.



(a) (b) (c)

Figure 4-8: Trajectories of the agent corresponding to landmark homing behavior. Objects are shown in dark gray, and the solid circle marks the home location. Each trajectory corresponds to a different release position. (a) Simulation results. (b) Trajectories obtained with the real robot. (c) For comparison, trajectories of an agent using the Average Landmark Vector model (adapted from Hafner et al. (2002) with permission).

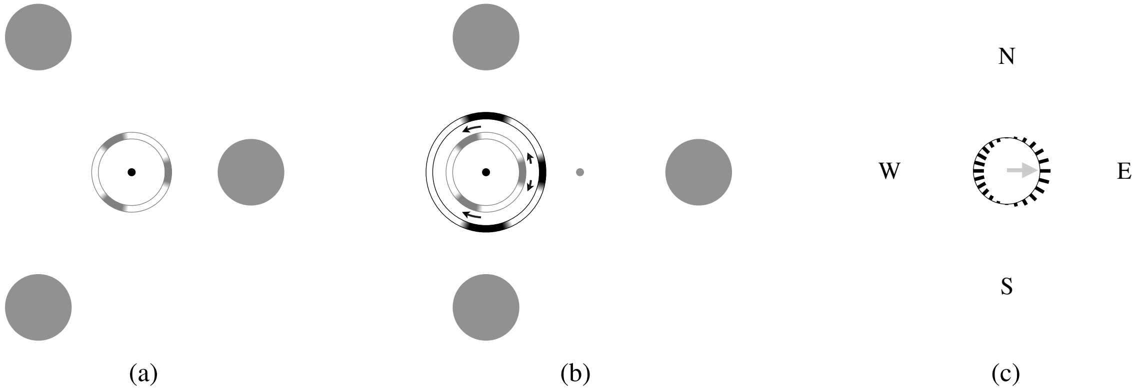
the environment as seen from the home location.[6] The discrepancy between neural activity corresponding to the perceived environment (current visual state) and the “snapshot” view (desired visual state) will lead to activity in the desired change of the visual modality, as illustrated in Figure 4-9(b). This activity will in turn propagate into the motor modality according to the correlation learned by the agent between motor activity and visual flow, thus leading to the observed homing behavior (see Figure 4-9(c)). Figure 4-10 summarizes the propagation of neural activity across the network that produces motion of the agent in the direction of the home location.

4.5 Discussion

The presented series of experiments demonstrates how two different natural homing strategies observed in insects can be observed with an artificial agent situated in a desert-like environment, using the same generic neural architecture as in the previous chapters – i.e. a neural architecture obviously not designed for any navigation task. Neural activity generated in the temperature modality – corresponding to a low desired temperature – will

---

[6] The image is projected with the same sign as when perceived from the home location since the state change neurons in the ambient light always display positive activity (see Footnote 3).



(a) (b) (c)

Figure 4-9: (a) Agent at the home location. The ring illustrates the visual state from this location, and the gray discs indicate the objects, i.e. the visual landmarks. (b) When the agent is located elsewhere, the discrepancy between the current visual state (outer ring) and the desired visual state (inner ring) produces activity in the desired change population corresponding to the visual flow indicated by the arrows (middle ring). (c) Motor activity generated by the propagation of neural activity from the visual modality. The resulting behavior is a motion of the agent in the east direction, i.e. towards the home location.

propagate in the network along a first path into the motor modality, generating an observed behavior corresponding to path integration homing. On the other hand, neural activity generated in the ambient light modality – corresponding to a low desired level of ambient light – will spread along a second path in the neural network, generating in this case a behavior matching visual landmark homing.

4.5.1 Unrelated Sensors and Internal Drive

This study stresses the importance of taking into account various sensors. Indeed, the robot is equipped with temperature and ambient light sensors, two sensory modalities a priori not related at all to navigation. Nonetheless, we showed that these sensors play an essential role in generating the homing behaviors, substituting an explicit "go home" signal that would seem more questionable from a biological and evolutionary perspective.

Note that the modification of neural activity triggering the homing behaviors, presented so far as external interventions, could be generated by the agent itself. Suppose that the agent were provided with an additional "homeostatic" sensory modality measuring say its hunger (as a matter of fact an internal, situated value). The nest being a source of food, the model would learn – as the agent eats inside the nest – a correlation between decrease of hunger and darkness as well as low temperature. As soon as the hunger level of the agent would exceed a given value during foraging, activity would propagate from the "desired" change of the hunger modality into both the ambient light and temperature

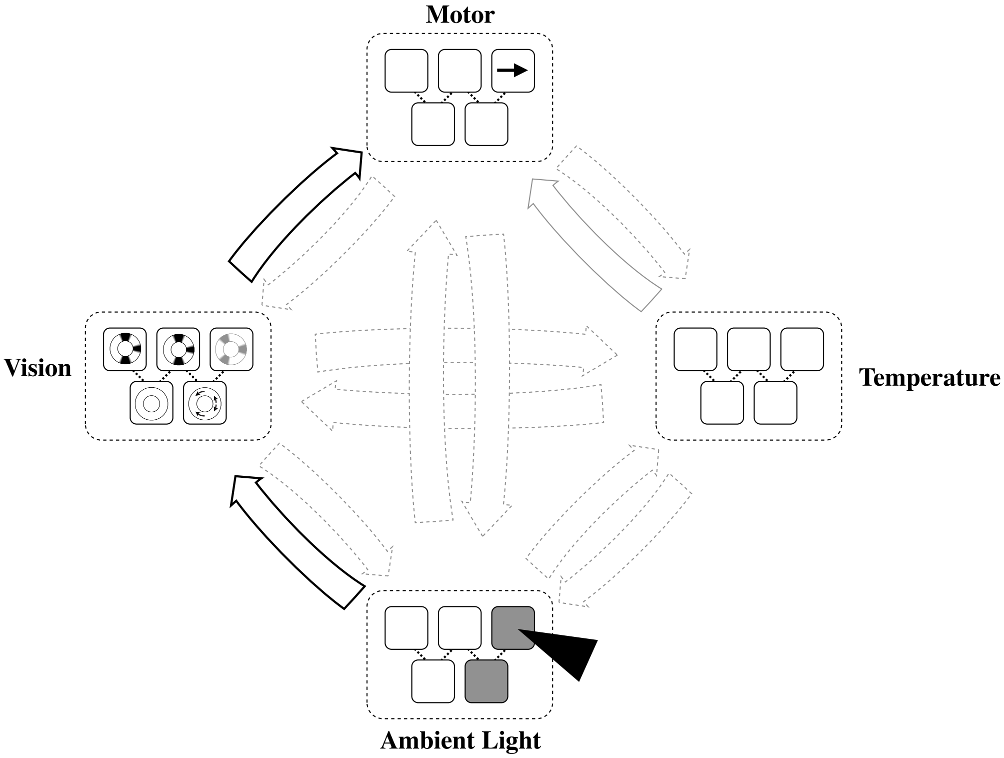


Figure 4-10: Generation of landmark homing behavior. When the activity of the desired ambient light neuron is externally decreased (indicated by the black arrow), neural activity propagates through the visual modality up to the motor modality. See Figure 4-9 and text for details.

modalities, generating the same patterns of activation as the ones studied previously, and thus triggering spontaneously both homing behaviors.

4.5.2 Reconsidering Common Assumptions

The obtained results allow to reconsider some commonly accepted assumptions in models of insect navigation. Indeed, we showed that path integration and visual homing behaviors can be produced with one single underlying mechanism (i.e. a homogeneous neural network with cross-modal Hebbian plasticity). This contrasts with current other models, which – to the best of our knowledge – all suggest that these strategies are based on distinct sets of control "modules"[7]: either, for path integration, an accumulator continuously updating a home vector by integrating all angles steered and all distances covered, as well as a mean of computing the home direction from the components of the integrated vec-

---

[7] Even if some specific parts, such as the weights of a neural network with a dedicated structure, have been shown to be learnable by the agent itself (Hafner and Möller, 2001).

tor; or, for landmark navigation, a unit storing the visual snapshot from the home position, matching it with the current image and computing the current home direction.

In summary, the results obtained in this chapter demonstrate that two different homing behaviors are not necessarily produced by distinct control structures, but can be produced by a single neural substrate. The qualitative difference in the observed behaviors (egocentric vs. geocentric navigation) is solely due to the different dynamics of the physical properties measured by various sensors (temperature vs. ambient light). Once again, this point clearly corroborates the frame-of-reference issues, showing that observed behaviors – resulting from the agent-environment interaction – do not necessarily reveal the underlying mechanisms.

4.6 Conclusions

This chapter describes a second series of experiments performed with a mobile robot endowed with the same homogeneous neural architecture as the one used in the previous chapters. Compared to the previous series of experiment, this chapter illustrates again how new, seemingly goal-directed behaviors are observed when the agent is equipped with a different set of sensory and motor modalities. Two major conclusions can be drawn from the observed results. First, the proposed neural architecture is able to generate, with a particular agent, different behaviors strikingly similar to natural behaviors observed with insects – namely path integration and visual landmark homing. Second, this series of experiments sheds new light on some biologically inspired models of navigation. In particular, it reconsiders the widely accepted assumption that distinct mechanisms (or neural structures) are necessary to account for different homing strategies.

Chapter 5

Homogeneous Coupling and Forward Models

In this chapter, a third series of experiments is conducted with the neural architecture described previously. The goal is here to gain a better understanding of the consequences resulting from the homogeneous connectivity in the proposed neural architecture, i.e. the fact that sensory and motor modalities are coupled together in both directions.

5.1 Introduction

In the experiments discussed in the previous chapters, the observed behaviors were produced by neural activity propagating almost exclusively from sensory modalities into the motor modality. However, the proposed neural architecture is homogeneous in its connectivity, and therefore also allows activity to flow in the opposite direction, namely from motor to sensory modalities. The aim of this chapter is to analyze possible effects resulting from these synaptic connections.

It is widely accepted that the central nervous system is involved in transforming sensory signals to motor commands. Recently, its role in exploiting neural signals flowing in the opposite direction has received an increasing interest in neuroscience, and several concepts such as "forward models," "efference copy" and "corollary discharge" have been developed (Wolpert and Ghahramani, 2000; Webb, 2004; Davidson and Wolpert, 2005). These concepts are closely related in the sense that they all suggest that biological systems use motor signals to predict the sensory consequences of their actions in order to be capable of rapid, robust and adaptive behavior.

A rigorous review of these concepts would exceed the scope of this thesis. We therefore choose to concentrate on one particular research direction, namely neural modeling of cricket phonotaxis (Reeve and Webb, 2003b; Russo et al., 2005), which we believe is detailed enough to extract relevant principles, and focused enough to allow comparison with our simple model. Inspired by this research, we choose a particular ecological niche for a series of experiments with a robotic agent that allows us to outline the role of the synaptic coupling allowing, in the proposed homogeneous neural network, activity to flow from the motor modalities into the sensory modalities. We show that the role of these coupling is at the behavioral level similar to the prediction mechanism of forward models. We finally discuss general principles gained from this series of experiments.

The chapter is organized as follows: Section 5.2 introduces the theoretical concepts of forward model and reafference cancellation, including the special case study of cricket phonotaxis. Then, Section 5.3 describes the series of robot experiments performed to analyze the effect resulting from the homogeneous connectivity of the proposed neural network. Finally, Section 5.4 discusses the relation between the obtained results and the theoretical framework of forward models.

5.2 Forward Models

5.2.1 Neural Mechanisms for Prediction

Descartes was perhaps the first to note explicitly that passive motion of the eye produces an impression of world motion whereas deliberate movements do not (Grüsser, 1995). Helmholtz is usually credited as the first to suggest that the brain, rather than sensing the gaze position of the eye, predicts the gaze position based on a copy of the motor command acting on the eye muscles (Wolpert and Flanagan, 2001). To demonstrate this, he used the following simple experiment. When the eye is moved without using the eye muscles (for instance by closing one eye and gently pressing with the finger on the open eye through the eyelid) the retinal locations of visual objects change, but the predicted eye position is not updated, leading to the false perception that the world is moving.

In the 1950s, this concept was more clearly formulated in the simultaneous and independent work of von Holst and Mittlestaedt (1950) and Sperry (1950). Since then, the idea that sensory consequences of motor commands are predicted internally has emerged as an important theoretical concept in all aspects of sensorimotor control and movement neuroscience (for reviews, see e.g. Wolpert and Ghahramani, 2000; Webb, 2004; Davidson and Wolpert, 2005).

This principle, illustrated in Figure 5-1, is referred to in the current literature as "forward model," termed so because it captures the forward or causal relationship between

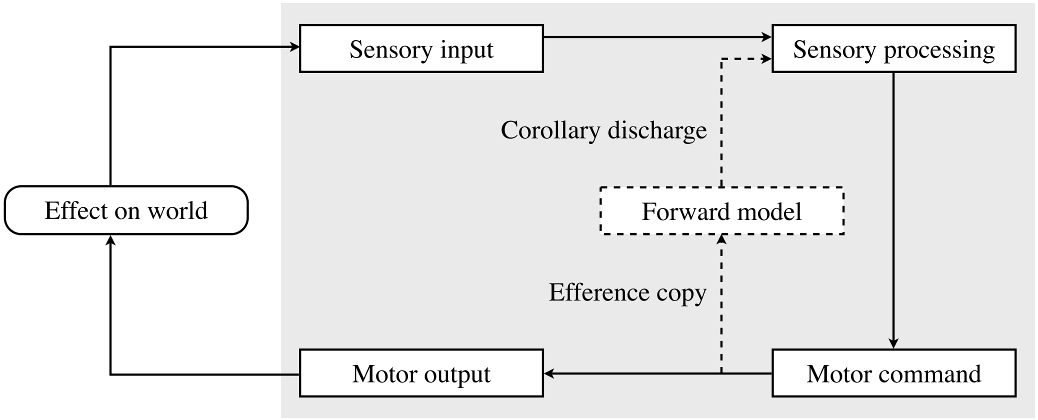


Figure 5-1: Schematic representation of a sensory motor system with a forward model. The unbroken lines indicate the loop by which sensory input is processed by the system to generate motor command, which is translated into motor output, has some effect on the world and causes new sensory input. The forward model is an internal loop (broken lines) that takes the motor command and predicts the expected sensory input, which can be used to modulate the processing of the actual sensory input. A classic example is that moving our eyes causes the image on the retina to move, but we perceive a stable world because the image movement is predictable from the eye movement command (adapted from Webb, 2004).

actions and their consequences. A copy of the motor command (“efference copy”) is used to generate predictions of the sensory consequences (“corollary discharge”) of the motor action. This prediction is then combined with the actual sensory feedback (“reafference”) from the movement.[1]

Several advantages to this prediction ability have been suggested (Miall and Wolpert, 1996). One is to compensate for significant delays sensory signals are subject to, which make sensory feedback too slow to be used for fast and accurate motor control. A prominent example comes from studies on dexterous grasping behavior of human (Johansson and Cole, 1992). When holding an object in a precision grip with the tips of the index finger and thumb on either side, sufficient grip force must be exerted to prevent slip due to the load force exerted by the object. When the object’s behavior is unpredictable (e.g. the load is externally generated by another person moving the object), the grip force is modified reactively in response to sensory feedback from the fingertips, with the consequence that grip force lags behind the load force. However, when the load is increased in a self-generated manner, a predictive mechanism can effectively use an efference copy of the motor command to anticipate the upcoming load force, thereby precisely generating an appropriate

---

[1] In the forward model terminology, the terms “efference copy” and “corollary discharge” are often used as input and output respectively of the forward model. However (as discussed in Webb, 2004), it is not uncommon to find these terms used interchangeably.

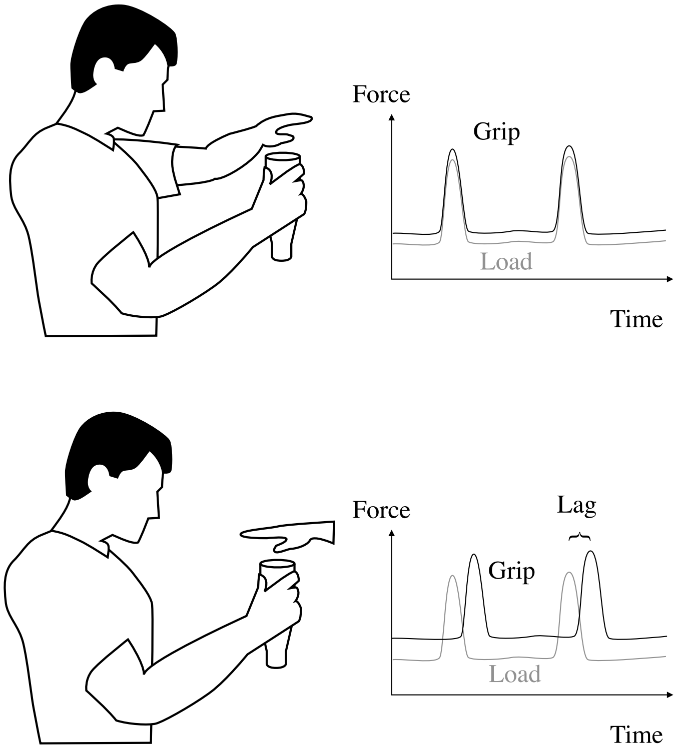


Figure 5-2: To prevent a ketchup bottle from slipping, sufficient grip force must be exerted to counteract the load. When the load is increased in a self-generated manner (left hand strikes the ketchup bottle, top), a predictor can use an efference copy of the motor command to anticipate the upcoming load force and thereby generate grip which parallels load force with no delay. However, when the load is externally generated (another person strikes the bottle, bottom), then it cannot be accurately predicted. As a consequence, the grip force lags behind the load force and the baseline grip force is increased to compensate and prevent slippage (adapted from Wolpert and Flanagan, 2001).

grip force with no delay (see Figure 5-2).

The conceptual idea of a predictive system such as a forward model can be, from a computational perspective, quite attractive, as indicated by the number of theoretical models suggesting further uses of forward models for state estimation, learning or even action planning (Johansson and Cole, 1992; Wolpert et al., 1995; Merfeld et al., 1999; Bhushan and Shadmehr, 1999; Kawato, 1999; Flanagan et al., 2001). However, one must be particularly careful when transferring concepts having their origin in control theory and industrial robotics into neuroscience. Once again, the frame-of-reference problem reminds us that the mechanisms underlying a system cannot be inferred by only observing its behavior. It is therefore interesting to observe in the scientific literature a growing acceptance of the idea that forward models are concrete neural mechanisms located inside the brain, most probably inside the cerebellum (Wolpert et al., 1998; Kawato et al., 2003; McKinstry et al., 2006).

Let us illustrate this issue with a concrete example. A putative "module" that could internally simulate different aspects of an agent and its environment – a theoretical view of internal forward models – could be used to predict the outcome of the system under various conditions. Some authors (Wolpert and Kawato, 1998; Haruno et al., 2001) thus suggests that when we perform an action such as lifting a carton of milk that is either empty or full, multiple forward models are "run" in parallel, simulating the sensory consequences under various possible contexts, e.g. an empty carton or a full carton; as each forward model captures a distinct dynamical behavior of the motor system, their prediction errors can be used during movement to determine in which context the motor system is acting. It is however far from clear how the discrete "contexts" could be defined – an instantiation of the well-known frame problem (Pylyshyn, 1987): Is the carton only half-full? Is it glued to the table? Which parts of the agent body must be "simulated"? etc.

In summary, "forward models" are increasingly recognized as a crucial explanatory concept in animal motor control. There is growing psychophysical evidence that motor output is used by the central nervous system to modulate its sensory processing. However, the neurophysiological evidence is less widespread and the mechanisms responsible for this predictive capability are still debated.

5.2.2 Reafference Cancellation

Let us return to the concept of forward models highlighted by our initial example – Helmholtz's experiment illustrating the perceived distinction between self-generated and external sensory stimulations –, which will constitute the main focus of this chapter. Comparison between predicted and actual sensory input can be used to distinguish results of environmental influences on the body (afference) from sensory changes induced by self-motion (reafference). Such a mechanism, termed reafference cancellation, has been extensively studied in the electric fish (for reviews, see e.g. Bell et al., 1997; Bastian, 1999; Bell, 2001). Electric fishes possess both electroreceptors for sensing current, and electric organs driven by motor commands for discharging current. To prevent the discharge of a fish's own electric organs interfering with its ability to sense its surroundings, the self-generated component is removed from the output of electrosensory cells. In primates, neurophysiological studies (Duhamel et al., 1992) show predictive updating in parietal cortex, anticipating the retinal consequences of an eye movement. In human, predictive mechanisms are believed to underlie the observation that the same tactile stimulus, such as a tickle, is felt less intensely when it is self-applied (Blakemore et al., 2000). Similarly, a dysfunction in the predictive mechanism is proposed to cause delusions of control in schizophrenic patients, i.e. the misattribution of self-generated actions as externally generated (Shergill et al., 2005).

Another reason why the notion of reafference cancellation is particularly interesting is that it can also be observed in "simpler" animals, whose less complex behavior and interaction with the environment allow for comparison with robotic experiments. Poulet and Hedwig have shown reduced responsiveness in auditory interneurons during singing in the male cricket (2002; 2003). Gebhardt and Honegger report, again in crickets, that interneurons sensitive to movement of the antennae are less sensitive during active movement by the cricket itself (2001). Also, it is well established that proprioceptive sensory neurons in several invertebrate systems, including crayfish (El Manira et al., 1991), stick insects (Bässler and Büschges, 1998) and locusts (Büschges and Wolf, 1999), have responses that are modulated in phase with central pattern generator rhythms.

This chapter will focus on a particular behavior of the cricket resulting from the interaction of different sensorimotor systems, which is described in more details in the following subsection.

5.2.3 Cricket Phonotaxis and Optomotor Behavior

Female crickets can locate conspecific males by moving towards the species-specific calling song the males produce by opening and closing their wing. The neuroethology of this system has been extensively studied (for a review, see Pollack, 1998), and an interesting series of robot models have been built (Webb, 1995; Webb and Scutt, 2000; Reeve and Webb, 2003a), showing that surprisingly simple controller can produce the same kind of selective approach behavior in a robot as is observed in the cricket.

Beside this phonotactic reflex, another well-studied behavior is observed as response to visual stimuli, known as optomotor reflex (Götz, 1975). Rotation of the entire visual field is usually produced by rotation of the cricket's body. Consequently, the optomotor reflex corrects unintentional self-rotation by turning in the opposite direction, thus compensating external disturbances and maintaining a straight trajectory.

Note, however, that integrating these two reflexes on an agent is not as straightforward as it might first seem. Indeed, the acoustic reflex tries to align the trajectory towards the sound source, whereas the visual response tries to correct for any change in the trajectory, thus counteracting the alignment attempted by the auditory system. Each turn towards the sound source (phonotactic reflex) produces a rotation of the visual field, which causes the agent to "correct" itself (optomotor reflex) and turn away from the sound again.

Different solutions have been proposed to solve this sensory conflict (Böhm et al., 1991;

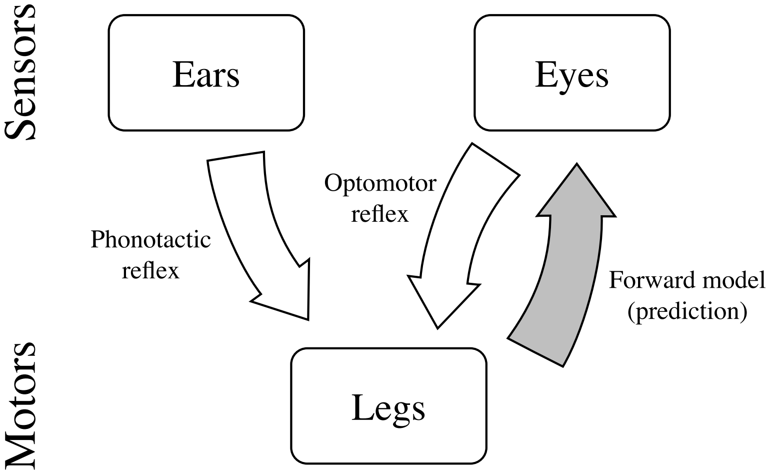


Figure 5-3: Schematic representation of a control network inspired by cricket phonotaxis. The phonotactic reflex system processes auditory sensory input and generates motor commands to orient the agent towards a particular sound source. The optomotor reflex system detects lateral optic flow in the visual sensory input and generates motor commands to correct rotational disturbances and thus maintain a straight trajectory. However, the two reflex systems produce conflicting behaviors: a turn generated by the phonotactic reflex produces a rotation of the visual field, which in turn is corrected back by the optomotor reflex. To smoothly combine the two competing reflexes, a forward model is introduced, which predicts the visual input from the motor commands and inhibits the optomotor reflex.

Reeve and Webb, 2003b). Even though a simple suppression of one sensory system by the other is a reasonably effective mechanism, a more efficient control method consists in using a forward model that predicts the perceived visual input from the motor commands and uses the prediction to inhibit the optomotor reflex, so that it is smoothly combined with phonotaxis (Figure 5-3). Russo et al. (2005) showed that this control strategy could successfully be implemented on a mobile robot with a spiking neural network, suggesting a plausible mechanism to explain real cricket behavior.

5.3 Robot Experiments

The set of experiments presented is this chapter aims at comparing properties of our neural architecture with existing theoretical models that explicitly integrate forward models for predictive sensory cancellation. We therefore choose an agent and an environment that are similar to those used in research on models of cricket phonotaxis.

The robotic platform used in this chapter is the AMouse robot, already introduced in Chapter 3. Note that instead of mounting extra sensors detecting a sound source, we choose to use the built-in light sensors of the Khepera platform. Phonotactic behavior (i.e. turning toward a *sound* source) is therefore substituted in our series of experiments by an equivalent phototactic behavior (i.e. turning toward a *light* source).

The main goal of the following experiments is to outline various behavioral properties resulting from the homogeneous connectivity of the neural network. More specifically, we investigate the behaviors observed with our system under different conditions – e.g. some parts of the neural network being temporarily removed, similar to neurophysiological lesion studies – in order to look into the following key questions:

- Does the agent display any behavior that can be identified as phototactic reflex, i.e. turning towards the light source?
- Does the agent display any behavior that can be identified as optomotor reflex, i.e. correcting its orientation when subject to unintended slip?
- If so, how do the two reflex-like behaviors interact with each other? Can some part of the network dynamics be interpreted as a forward model, i.e. a predictive mechanism used to cancel sensory consequences of self-generated action?

5.3.1 Agent and Environment

The AMouse robot, illustrated in Figure 5-4, is equipped for this series of experiments with the following sensory and motor modalities:

Light Modality

The current state population consists of 6 neurons, whose activities correspond to the readings of the six light sensors in the front of the Khepera. The state change is defined as the horizontal components of the corresponding visual flow (as described in Appendix 3.A).

Optic Flow Modality

The current state consists of an array of 110×40 neurons whose activity corresponds to the horizontal optic flow field detected by the omnidirectional camera. Note that in contrast to the previous experiments, the state in this modality does not represent the pixel values of the image from the omnidirectional camera. Rather, the state represents the horizontal optic flow computed from the successive images captured by the camera. In other words, the camera is used in this experiment to emulate a set of elementary motion detectors.

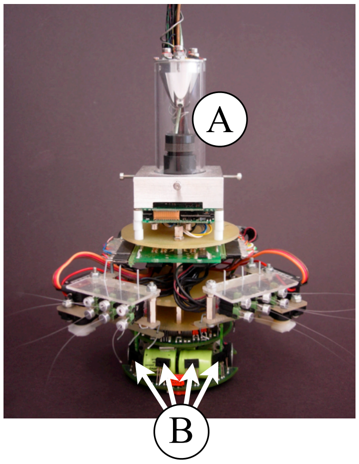


Figure 5-4: The "AMouse" robot used for the series of experiments presented in this chapter. The robot is equipped with an omnidirectional camera (A) used as an optic flow sensor, as well as with six light sensors (B) integrated on the Khepera basis. Note that the whisker sensors are not used in this series of experiments. The specifications of the robot are provided in Chapter 3.

Battery Level ("Reward") Modality

This experiment introduces a new simulated sensory modality, whose value is increased when the robot externally receives some kind of "reward".[2] Here, the reward is delivered whenever the robot reaches the light source. One could metaphorically consider this sensory modality as a measure of battery level, with a battery charging station located at the light source.

Motor Modality

As in the experiment presented in Chapter 3, the current state in the motor modality consists of two neurons, indicating respectively the forward and turning speed of the robot.

Environment

The environment is shown in Figure 5-5. The light source is a lamp positioned on one side of the arena.

---

[2] At first sight, the term "reward" might be slightly confusing given the context of our research. However, it is important not to confuse (a) an adjective used to refer, from an *observer* perspective, to a particular sensory signal (which is increased at particular points in time, *as if* the robot receives a reward) with (b) the purpose of internal mechanisms (there is indeed no reward in the neural architecture: the "reward" signal plays the very same role as any other sensory or motor signals). The term "reward" – similar to the adjective "desired" – is used here only to help the reader better understand the model.

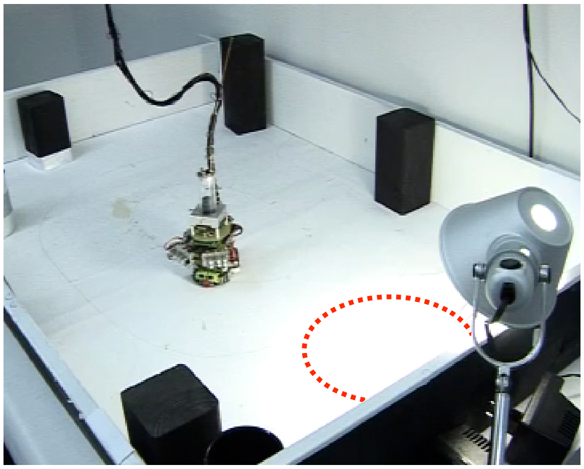


Figure 5-5: The environment consists of a flat arena, the light source being a lamp placed on one of its sides. Whenever the robot reaches the light source (dotted area), its battery level is increased.

5.3.2 Observed Behaviors

During the initial phase, the robot is randomly driven across the arena, so that the neural network, whose synaptic coupling are initially all set to zero, can capture the various cross-modal correlations. Whenever the robot reaches the light source, it receives an electronic "reward" as described above, i.e. its battery level is increased. The neural network connectivity shaped after this initial exploration phase is given in Figure 5-6.

After that, the activity of the neuron corresponding to the desired state in the battery level modality is kept at a high level, and the robot, placed at a given position in the arena, is let to move on its own. We then observe the behaviors resulting from the propagation of activity across the neural network under four different conditions.

First, we investigate what happens when the optic flow sensor is switched off, i.e. discarded from the neural network. We observe that the robot moves toward the light source, maintaining its orientation towards the light source. This first result, shown in Figure 5-7(a), reveals the phototactic behavior displayed by the robot.

Second, random perturbations are introduced into the system by adding noise to the wheel motors. The unintended slips provoked by the perturbations make the trajectory of the robot more winding, as shown in Figure 5-7(b). This result is in fact not surprising, since the robot, deprived of its optic flow sensor, cannot detect the random changes of orientation.

In a third phase, we reintroduce the optic flow sensors into the neural network to see if the robot, now also equipped with a sensor detecting unintended rotations, can compensate for the perturbations. In this third phase, we also remove the synaptic couplings allowing

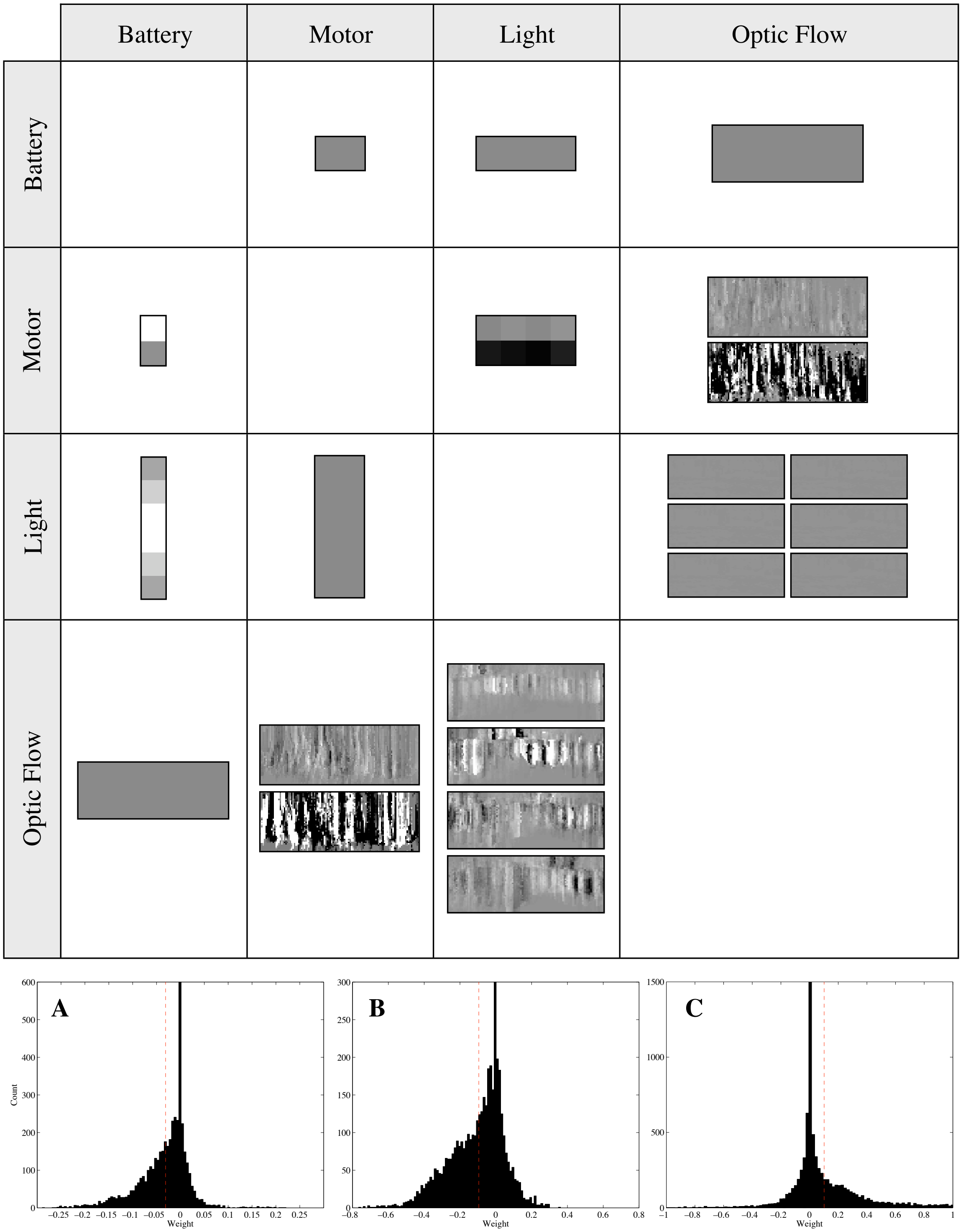
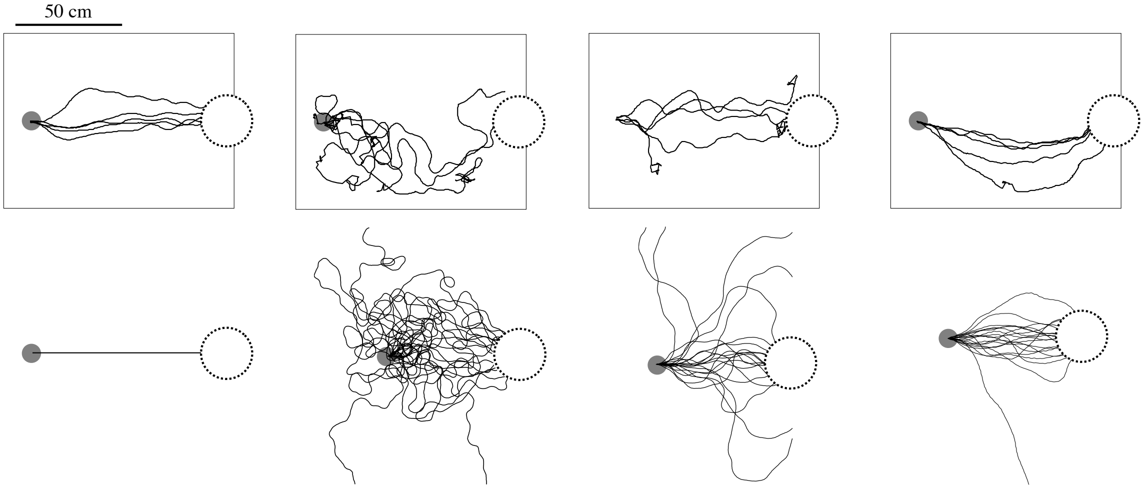


Figure 5-6: Top: Connectivity of the neural network after the initial phase ($\eta = 0.01$, $\varepsilon = 0$). The synaptic weights are displayed for the cross-modal couplings connecting each modality (column) to all other modalities (rows). Positive values are represented in lighter colors, negative values in darker color, and zero values in gray. Bottom: Histograms of weight distributions with average (dotted line) for some of the cross-modal couplings. A: optic flow → motor; B: motor → optic flow; C: light → optic flow. See text for details.



(a) (b) (c) (d)

Figure 5-7: Trajectories of the agent observed under various conditions, either with the AMouse robot (top) or in simulation (bottom). The trajectories of the robot were obtained by tracking a marker on the robot with an overhead camera. The gray disk indicates the starting location and the dotted circle indicates the light source. (a) Phototaxis: the robot moves towards the light source. At this first stage, the optic flow sensor is turned off. (b) Perturbations: noise is introduced in the motors at this second stage. The trajectories are more winding and the robot often misses the light source. (c) Optomotor reflex: when the optic flow sensor is reintegrated into the system, the robot corrects the perturbations detected through rotation of the visual field. At this third stage, the synaptic connections leaving the motor modalities are removed from the network. The zigzag in the trajectories reveals the competition between phototactic and optomotor reflexes: the robot alternatingly turns toward the light source and corrects the perceived rotation by turning away from the light source. (d) Reafference cancellation: the synaptic connections leaving the motor modalities are reintegrated into the network at this fourth stage. As a result, the robot's trajectories get straighter and the robot almost always reaches the light source despite the perturbations.

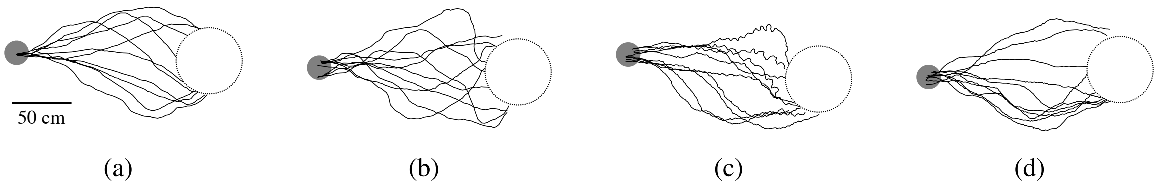


Figure 5-8: Trajectories of a robot modeling aspects of cricket phonotaxis, controlled by a network that explicitly contains distinct modules responsible for the phototactic and optomotor reflexes, as well as for the reafference cancellation. As in Figure 5-7, the gray disk indicates the starting location and the dotted circle indicates the sound source. (a) Trajectories obtained when the robot is driven by phonotaxis only, without perturbations. (b) The trajectories are more winding with perturbations, introduced as random noise in the motors. (c) When the optomotor reflex module is enabled, the zigzags in the trajectories reveal the conflict between phonotactic and optomotor reflexes. (d) Trajectories obtained with an additional forward model performing reafference cancellation, thus allowing phonotaxis and optomotor reflexes to be smoothly integrated despite the perturbations (adapted from Russo et al., 2005).

activity to flow out of the motor modality into the different sensory modalities. These couplings will be integrated back in the next phase, so that their role can better be identified. The behavior displayed by the robot shows that it is indeed able to correct unintended slips: the trajectories toward the light source get more straight, as shown in Figure 5-7(c). The observed behavior can be described as follows. On the one hand, the robot makes use of the light sensors to turn towards the light source, i.e. to create a behavior that can be identified, from the observer perspective, as phototactic reflex. On the other hand, the robot employs the optic flow sensor to correct the random perturbations, i.e. to create a behavior that can be identified as optomotor reflex. Note however that the robot sometime still misses the light source, and that the trajectories reveal zigzags, which are typical of two antagonist reflexes acting against each other: when the robot turns toward the light source (phototactic reflex), the rotation of the visual field generates optic flow that in turn is corrected – as any random perturbation – by turning in the opposite direction (optomotor reflex), i.e. away from the light source.

In the fourth and last phase, the synaptic couplings leaving the motor modalities, which were removed from the neural network in the previous phase, are reintegrated back, thus restoring the original neural architecture where all modalities are homogeneously coupled to each other. The surprising consequence of reintegrating all synaptic couplings can be seen in Figure 5-7(d): despite the perturbations, the robot trajectories towards the light source do not oscillate anymore, and get almost as straight as they were before the perturbations were introduced. Also, the agent hardly ever misses the light source. The role of the synaptic couplings from the motor modalities to the sensory modalities can thus be deduced from the comparison of the trajectories shown in Figures 5-7(c) and 5-7(d). Without these couplings, the zigzag trajectories reveal the conflict between phototactic and optomotor reflexes. However, this conflict disappears when these couplings are reintegrated in the network. From an observer perspective, their role can thus be identified as a forward model, i.e. a predictive cancellation mechanism of self-generated sensory input (reafference cancellation), preventing the robot from correcting rotation of the visual field generated when actively turning toward the light source.

In summary, different behaviors are observed with the robot endowed with the proposed homogeneous network architecture: phototactic and optomotor reflex-like behaviors, as well as predictive cancellation of self-generated sensory input. Of particular interest is the comparison of the behaviors observed so far to behaviors obtained under similar conditions with a robotic system where the two reflexes and the forward model for reafference cancellation are, in contrast to our approach, explicitly built into the system as distinct "modules," i.e. specific neural networks with individual structures and connectivity. Figure 5-8 shows trajectories obtained with such a system, namely a robot using a biologically inspired neural network to model multisensory integration in cricket phonotaxis (Russo et al., 2005). A comparison between Figures 5-7 and 5-8 shows that despite completely different underlying mechanisms, the behaviors observed in both systems bear striking similarities.

The following subsections investigate the results obtained with our system more in details, and describe how the different behaviors described so far are generated by neural activity flowing across the sensory and motor modalities in the homogeneous neural network.

5.3.3 Phototactic Reflex

We first discuss how the dynamics of the network leads to a behavior that corresponds to a phototactic reflex, where the agent moves toward the light source (Figure 5-7(a)).

This behavior can easily be explained by looking at the propagation of neural activity between the battery level, light, and motor modalities (see Figure 5-9). During the initial phase, the agent is – on average – facing the light source when it receives the reward. The agent therefore learns a correlation between increase of reward and strong stimulation of the central light sensors. By increasing the activity of the desired reward neuron, neural activity will thus propagate to the desired light population, projecting a sensory input corresponding to a light source perceived from the front of the agent. Discrepancy between current and desired states in the ambient light modality makes neural activity propagate further to the

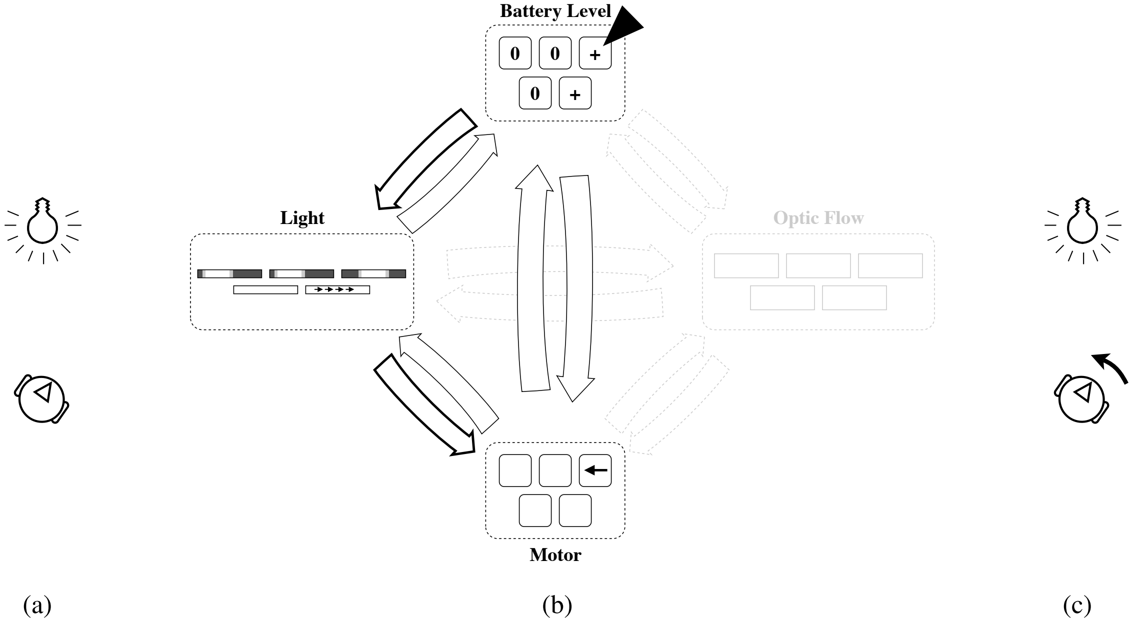


Figure 5-9: Propagation of neural activity in the network leading to phototaxis. (a) In this example, the light source is perceived on the left of the visual field of the agent. (b) The optic flow sensory modality is grayed out, indicating that this modality is discarded from the network at this stage of the experiment. The neuron corresponding to the desired reward level is externally activated (black triangle). Activity thus propagates into the visual modality. Discrepancy between current and desired states in the light sensory modality produces activity that further propagates into the motor modality. (c) The resulting motion is a turn to the left, i.e. towards the light source.

motor modality, leading to the observed behavior.

As a matter of fact, this behavior closely matches the object-following behavior described in Chapter 3, where the agent moved toward an object (i.e. a source of tactile stimulation) when the neuron corresponding to the desired tactile input was activated. In the present chapter, the battery level sensor plays the role of the tactile sensor, and the light sensors substitute the camera.

5.3.4 Optomotor Reflex

A second behavior corresponding to the optomotor reflex was observed as noise was introduced in the motors. We showed that the noise in the motors, which perturbs the trajectory, often even preventing the agent from reaching the light source (Figure 5-7(b)), is at least partly compensated when the agent is equipped with visual flow sensors detecting rotation of the visual field (Figure 5-7(c)).

To account for this behavior, which corrects self-rotation perceived visually, let us con-

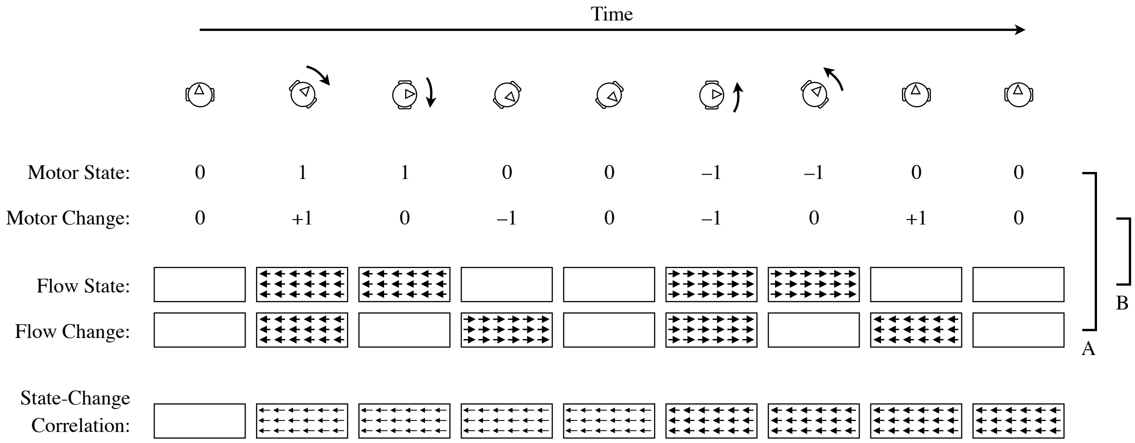


Figure 5-10: Qualitative correlations between motor and optic flow modalities. In this figure, only the activity of the motor neuron corresponding to rotation is shown. The motor and optic flow changes reflect the difference between the current and previous states. (A) Activity in the optic flow modality corresponding to an accelerating visual flow to the left is correlated to motor activity corresponding to rotation of the agent to the right. (B) Activity in the motor modality corresponding to angular acceleration to the right is correlated to optic flow activity corresponding to a visual flow to the right. The bottom row shows the qualitative correlation capture between either (A) motor state and change of visual flow or (B) change of motor state and visual flow. Quantitative measurements of these correlations are given in histograms A and B of Figure 5-6.

sider the neural dynamics between visual flow and motor modalities. The weights of the synapses coupling change of visual flow (i.e. acceleration of visual flow) and motor state reveal a correlation between a visual flow increasing in one direction and a motor state corresponding to rotation in the opposite direction (see Figure 5-10). An accelerating visual flow to the left is correlated to a rotation of the agent to the right, and vice versa. Let us examine what happens when the agent, subject to perturbations, turns in one direction, say to the right (see Figure 5-11). Rotation to the right will obviously elicit a perceived visual flow in the opposite direction, i.e. to the left. The difference between detected visual flow (left) and desired visual flow (essentially zero) will lead to activity in the desired change population corresponding to an accelerating visual flow to the right, which is correlated to motor activity producing rotation to the left. The robot will thus correct the angular perturbation to the right and turn to the left – exactly the behavior an optomotor reflex is supposed to elicit.

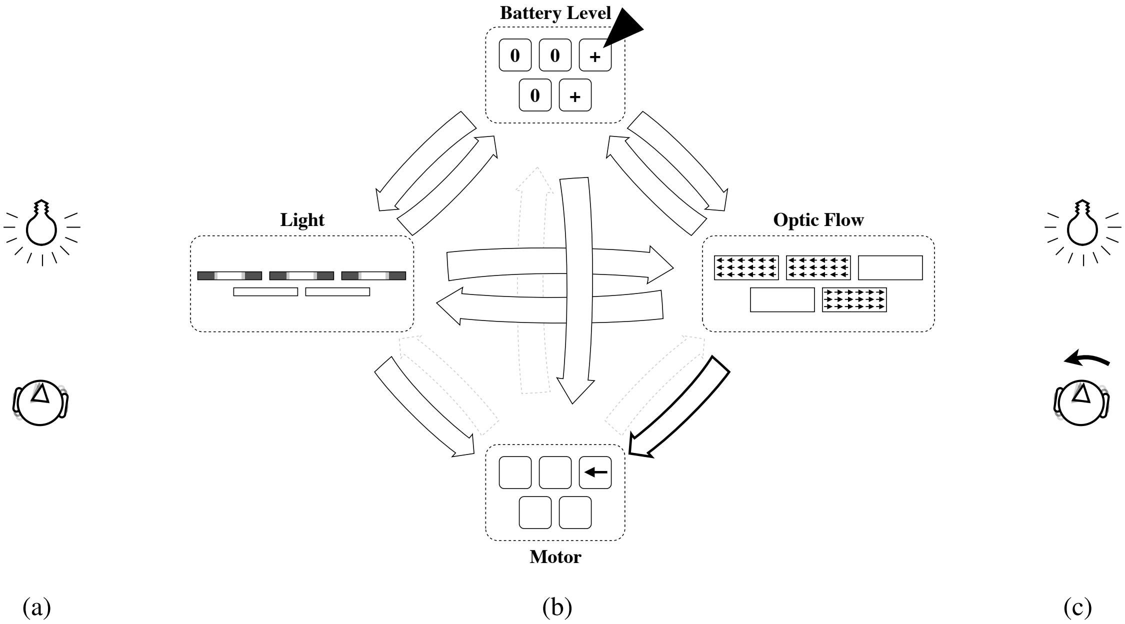


Figure 5-11: Propagation of neural activity in the network leading to optomotor reflex. (a) In this example, the agent faces the light source, but is subjected to unintended self-rotation to the right. (b) The difference between the current perceived optic flow (to the left) and the desired optic flow (essentially zero) leads to activity in the desired change population corresponding to an accelerating visual flow to the right. This neural activity propagates into the motor modality and generates the correlated activity, i.e. (c) a turning motion to the left that corrects the unintended self-rotation to the right.

5.3.5 Forward Model

We have shown so far that neural activity flowing from the sensory modalities into the motor modalities generates two distinct reflex-like behaviors. We also have observed that when the synaptic couplings allowing neural activity to flow in the opposite direction (namely from the motor to the sensory modalities) are removed, these two "reflexes" produce competing and alternating behaviors resulting in zigzag trajectories (Figure 5-7(c)). However, this competition of antagonist reflexes is not observed with the original, homogeneous neural architecture that allows activity to flow in both directions between sensory and motor modalities (Figure 5-7(d)).

The synaptic couplings leaving the motor modality can thus be identified as playing the role of a forward model, canceling self-generated sensory input. Indeed, when the robot starts to turn towards the light source, the difference between current and desired motor states generates neural activity – in the desired change population – that propagates out of the motor modality into the different sensory modalities. In fact, activity propagates then essentially from the motor modality into the visual flow modality (the other synaptic coupling, namely from the motor modality to the battery level modality or to the light modality, are essentially zero since there is no correlation between change of motor state and battery level or light state). The consequence of neural activity flowing into the visual flow modality when the robot starts to turn, is that it reduces the optomotor reflex, thus preventing the robot from correcting back the self-generated turning motion towards the light source.

Let us consider the case where the robot starts turning towards the light source perceived on its left. As the robot accelerates to the left, the desired visual flow population gets activated with a pattern of activity corresponding to visual flow in the opposite direction (since motor activity corresponding to acceleration to the left is correlated to visual flow to the right, see Figure 5-10). Metaphorically speaking, the synaptic coupling from the motor modality to the visual flow modality "predicts" a perceived visual flow to the right when the robot accelerates to the left. The difference between current and desired visual flow is reduced, and so is the optomotor reflex too.

Figures 5-12 and 5-13 illustrate graphically the neural dynamics of the network for the case where the two reflexes produce a conflicting behavior (when the synaptic couplings leaving the motor modality are removed), and for the case of the predictive cancellation of the optomotor reflex (with the original homogeneous connectivity in the network), respectively.

5.4 Discussion

The main conclusion from the series of experiments presented in this chapter is that both reflexive and predictive cancellation mechanisms can be produced by one single, unbiased neural network architecture. Indeed, our homogeneous neural architecture is able to reproduce the very same behaviors as the ones observed with a control system explicitly implementing different reflexes in combination with a predictive forward model. Important principles can be drawn from these results.

5.4.1 Segmentation of Behavior

The first one is directly related to the notorious frame-of-reference problem. From an observer perspective, the overall behavior of the agent can be segmented in different tasks. First, the agent moves toward the light source (phototatic reflex). Second, it corrects self-rotation by turning in the direction opposite to the perceived rotation of the environment

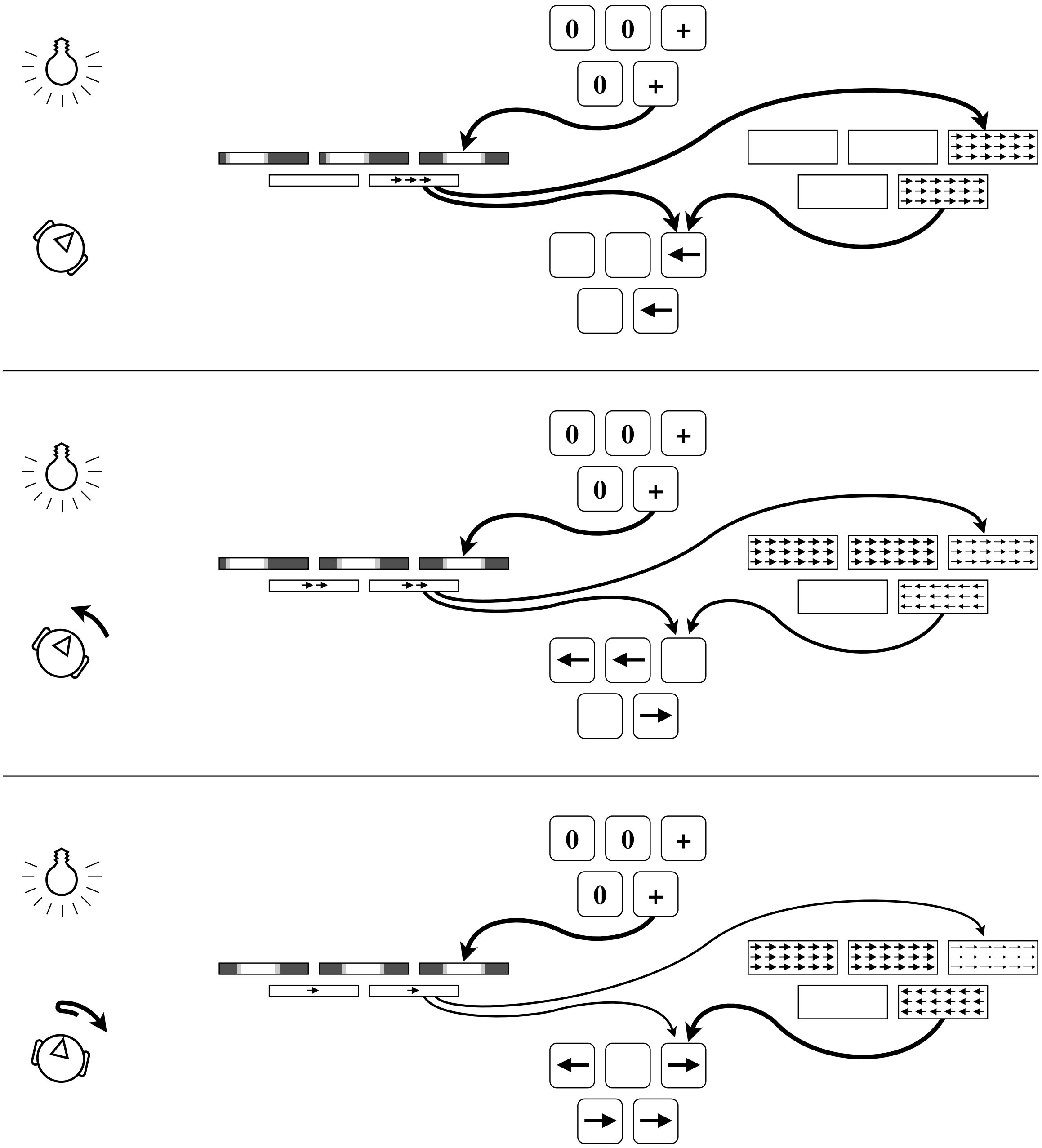


Figure 5-12: Competition between phototactic and optomotor reflexes. Activity generated in the battery level modality flows into the light modality and propagates further, both directly into the motor modality and through the optic flow modality, to produce the phototactic turning behavior (top). Note that motor activity does not propagate further in the network, since in this experiment the synaptic couplings leaving the motor modality are removed. As the robot turns towards the light source (middle), the difference between current and desired optic flow changes in sign, producing motor activity corresponding to the optomotor reflex behavior, which progressively cancels the phototactic reflex and eventually produces a counter rotation of the robot away from the light source (bottom).

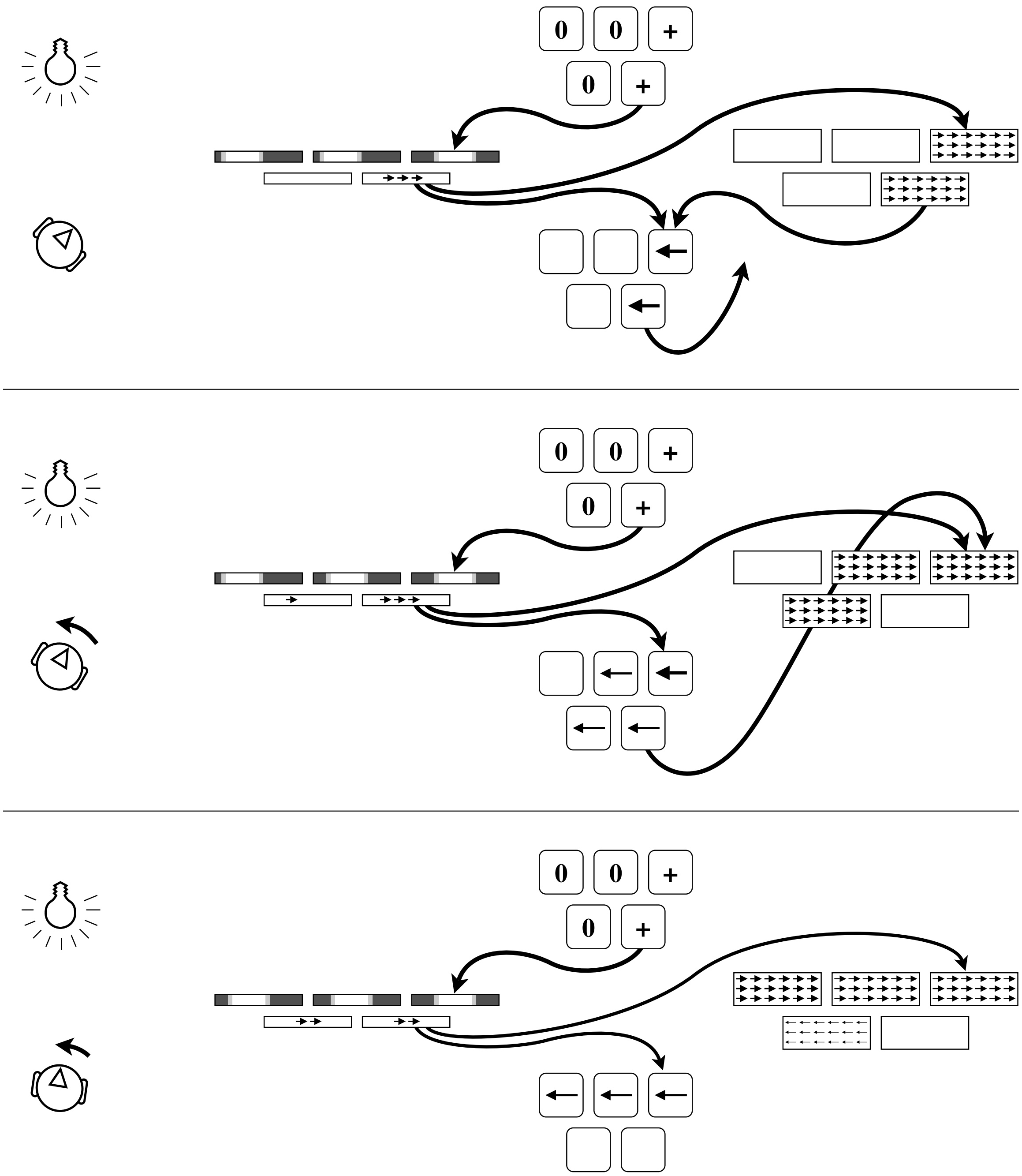


Figure 5-13: Cancellation of optomotor reflex during phototaxis. As in Figure 5-12, activity generated in the reward modality flows through the network and produces the phototatic turning behavior (top). However, motor activity can now propagate back into the network. As the robot starts turning towards the light source (middle), activity from the motor modality increases the desired optic flow, therefore reducing the difference between current and desired optic flow, and thus canceling the optomotor reflex (bottom).

(optomotor reflex). Also, it is able to cancel self-generated sensory stimulation (forward model): the agent does only correct the rotation perceived visually when it is generated externally, not when the agent is actively turning towards the light source. However, the corresponding processes can only be identified when considering the whole agent-environment interaction, together with the internal dynamics of the neural network. In other words, specific "modules" do not need to exist – such as distinct neural networks with dedicated structure and functionality – that explicitly perform the observed reflexes and cancellation of self-generated sensory stimulation. Our results demonstrate that the same neural structure, i.e. the same mechanism, can underlie phototactic and optomotor reflexes as well as forward model for reafference cancellation.

5.4.2 Reflexes and Forward Models

Secondly, our approach provides an innovative perspective into multimodal integration, and allows us to identify potentially artificial problems resulting from unwarranted designer assumptions. A typical and broadly accepted bias is that internal signals generally flow from the sensors into the motors. Even in models containing various feedback loops or reentrant connections, control networks receive mostly input from the sensors, and output is used to actuate the motors. Despite the growing interest in considering behavior as a result of sensory-motor coordination, the vast majority of control models in the current literature are still built on such an assumption – a more or less direct reminiscence of the computational paradigm.

When observing crickets, it seems for instance straightforward to attribute goal seeking and slip correction behaviors to simple built-in reflexes, i.e. connections from the sensors to the motors. However, this approach turns out to require additional mechanisms of completely different nature to compensate undesired behavior resulting from the combination of both reflexes. In contrast, our approach avoids this bias. In our neural architecture, modalities are coupled in a homogeneous way, irrespective of whether they represent sensors or motors. As a result, it turns out that apparently competing processes (the phototatic and optomotor reflexes) can be smoothly integrated without resorting to extra means (such as a predictive cancellation mechanism).

This raises the following compelling question: to what extent are concepts such as forward models mere artificial constructs postulated as a way to compensate for side-effects resulting from a biased approach – i.e. an approach that initially views behavior as an input-ouput process where sensory signals are processed to generate motor actions?

5.4.3 Role of Past Interaction with the Environment

Thirdly, the results presented in this chapter demonstrate that in order to understand the behavior of an agent, its past interaction with the environment cannot be neglected. The internal dynamics of the described neural network is not only shaped by the different physical processes characterizing the various sensors and motors, but also by the cross-modal correlations experienced by the agent throughout its lifetime.

Overlooking this yet obvious principle can induce misleading conclusions. An interesting example can be found in a series of studies on the cancellation of self-generated stimulation conducted on electric fish and human, which suggests different underlying mechanisms. However, a reinterpretation of the results from the new perspective gained in this chapter can lead to the opposite conclusion. Indeed, the various observations reported in these studies can be accounted for by applying the same neural network architecture on two systems corresponding to electric fish and human. Dissimilarities in observed sensory cancellation might then rather be attributed to the different interactions that electric fish and human have with their environment. A detailed argumentation is provided in Appendix 5.A on page 101.

5.5 Conclusion

Inspired by research done on cricket phonotaxis, and by the theoretical framework of predictive forward models, we performed a series of experiments to analyze the role of the homogeneous connectivity of the proposed neural architecture. The results show that synaptic couplings allowing neural activity to flow from sensor modalities to motor modalities generate reflex-like behaviors, whereas the role of synaptic couplings in the opposite direction can be identified as predictive reafference cancellation.

Even though the proposed neural architecture does not aim at modeling any realistic aspect of biological systems, our study suggests that the same mechanism could possibly be underlying both reflexes and forward models observed with natural agents.

Furthermore, our approach sheds new light on some neurophysiological findings, suggesting that disparities in sensory attenuation observed with electric fish and human could be attributed to differences in their interaction with the environment, rather than to distinct cancellation mechanisms.

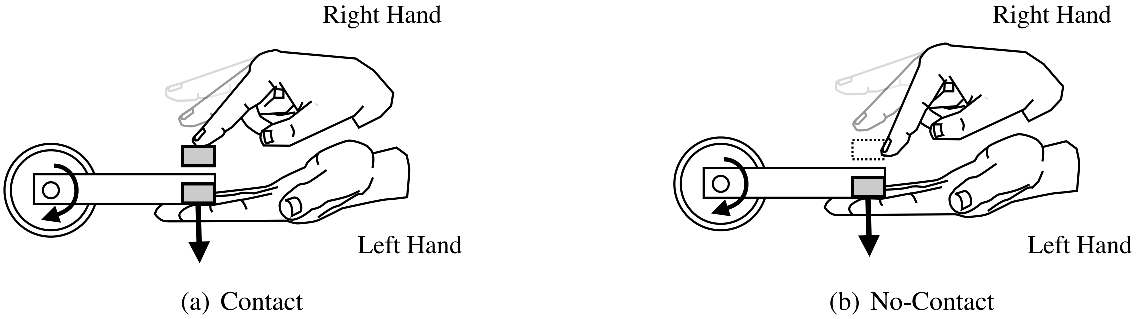
Appendix 5.A

Do Electric Fish and Humans Need Distinct Reafference Cancellation Mechanisms?

Electrosensory cells of electric fish are less sensitive to an electric current generated by the animal itself than to a current produced externally (Bell, 1982, 2001). In this cancellation mechanism, a prediction of the expected electric sensory input is generated based on the fish's own motor command. This prediction is then subtracted from the output of the electrosensory cells. Further studies revealed that this predictive sensory cancellation is also observed for sensory inputs that are predictable only on the basis of incoming sensory information such as proprioception. For example, a consistent pairing of electrical stimulus with a particular phase of a passive tail bend is sufficient to produce sensory cancellation (Bastian, 1995, 1999).

Interestingly, Bays, Flanagan and Wolpert performed a similar kind of study with human. They showed that when one finger touches the other, the resulting tactile sensation is perceived as weaker than the same stimulus externally imposed (Bays et al., 2005). However, they showed that when one finger makes a tapping movement above a finger of the other hand, sensation in the passive finger is only attenuated when contact is expected between the fingers (Bays et al., 2006), i.e. when the subjects are habituated to the situation where a tapping movement leads to tactile perception in the same – active – finger (see Figure 5-14).

This means that when a sensory stimulation (electric or tactile) is consistently paired with a particular proprioceptive event (tail bend or finger flexion), sensory attenuation is observed in electric fish, but not necessarily in human. Consequently, Bays et al. suggest that



(a) Contact (b) No-Contact

Figure 5-14: Schematic of the apparatus used by Bays et al. (2006) to measure attenuation of self-generated tactile sensations. On contact trials (a), participants produced a brief force pulse with their right index finger on a force sensor fixed above their left index finger. A similar force pulse was delivered to the left index finger by a torque motor. On no-contact trials (b), the force sensor was replaced by an optic sensor detecting the tapping movement, so that the right index finger did not make contact. Participants trained to produce a tapping movement on the force sensor (a) perceived attenuation of self-generated tactile sensations in their left finger, even in occasional no-contact trials (b). Participants trained to produce a similar tapping movement, but with no contact (b), did not perceive any significant attenuation of self-generated tactile sensations.

> “the cancellation mechanism in the human somatosensory system is not identical to that in electric fish.” (Bays et al., 2006, p. 28)

However, a reinterpretation of these results from the new perspective gained in this chapter can lead to the exact opposite conclusion.

The adaptive sensory cancellation observed in the electric fish strongly suggests that the animal is able to associate different sensory or motor events occurring simultaneously. This closely corresponds to the effect of the Hebbian learning rule used in our network architecture, which captures various cross-modal correlations. It is thus legitimate to investigate how our neural architecture might suggest principles explaining the various findings described so far.

Figure 5-15 describes how sensory cancellation can be accounted for using our neural architecture with an agent corresponding to a model of an (obviously very simplified) electric fish. Interestingly, the same model is able to reproduce sensory attenuation similar to the one observed in human, as explained in Figure 5-16. The key feature that explains the difference in sensory cancellation as observed between electric fish and human is, in our approach, the different cross-modal correlations experienced by the agent.

In the case of the electric fish, our model predicts sensory cancellation after a substantial correlation between tail bend and electric stimulation (Figure 5-15(b)). Correspondingly, the experiments performed by Bastian (1999), where sensory cancellation was observed in

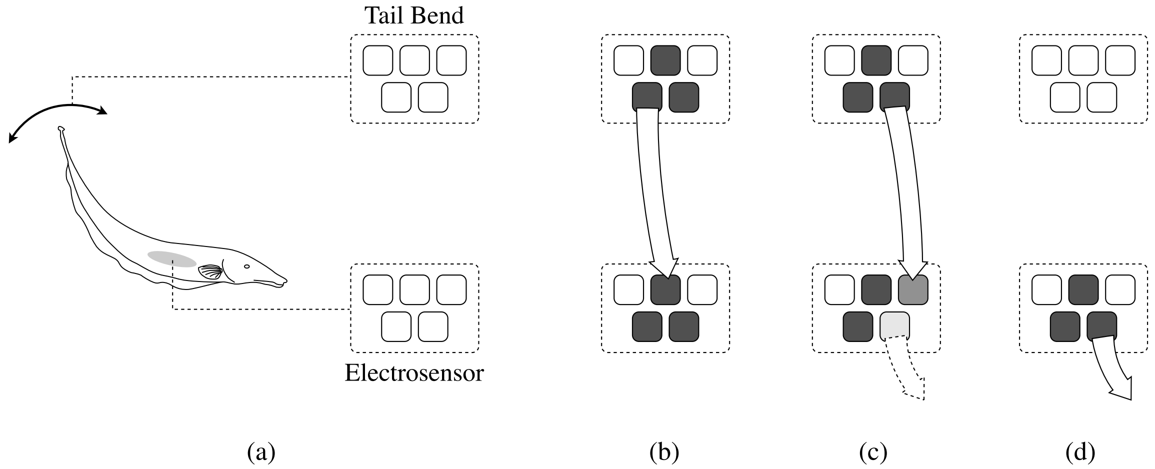


Figure 5-15: Model of sensory cancellation in electric fish. (a) The motor modality corresponds to the bending motion of the tail, and the sensory modality indicates the detected electric current. (b) Consistent pairing of tail bend with electric stimulation produces a cross-modal correlation between activity in the motor modality and in the sensory modality. As a consequence, neural activity generated by a tail bend propagates into the sensory modality (c), thus reducing the difference between current and desired sensory states. The neural output of the electrosensory modality (indicated by the arrow pointing out from the sensory modality) is thus attenuated when electric current is detected while the tail is bent (c), as compared with the case where electric current is detected with no tail bend (d).

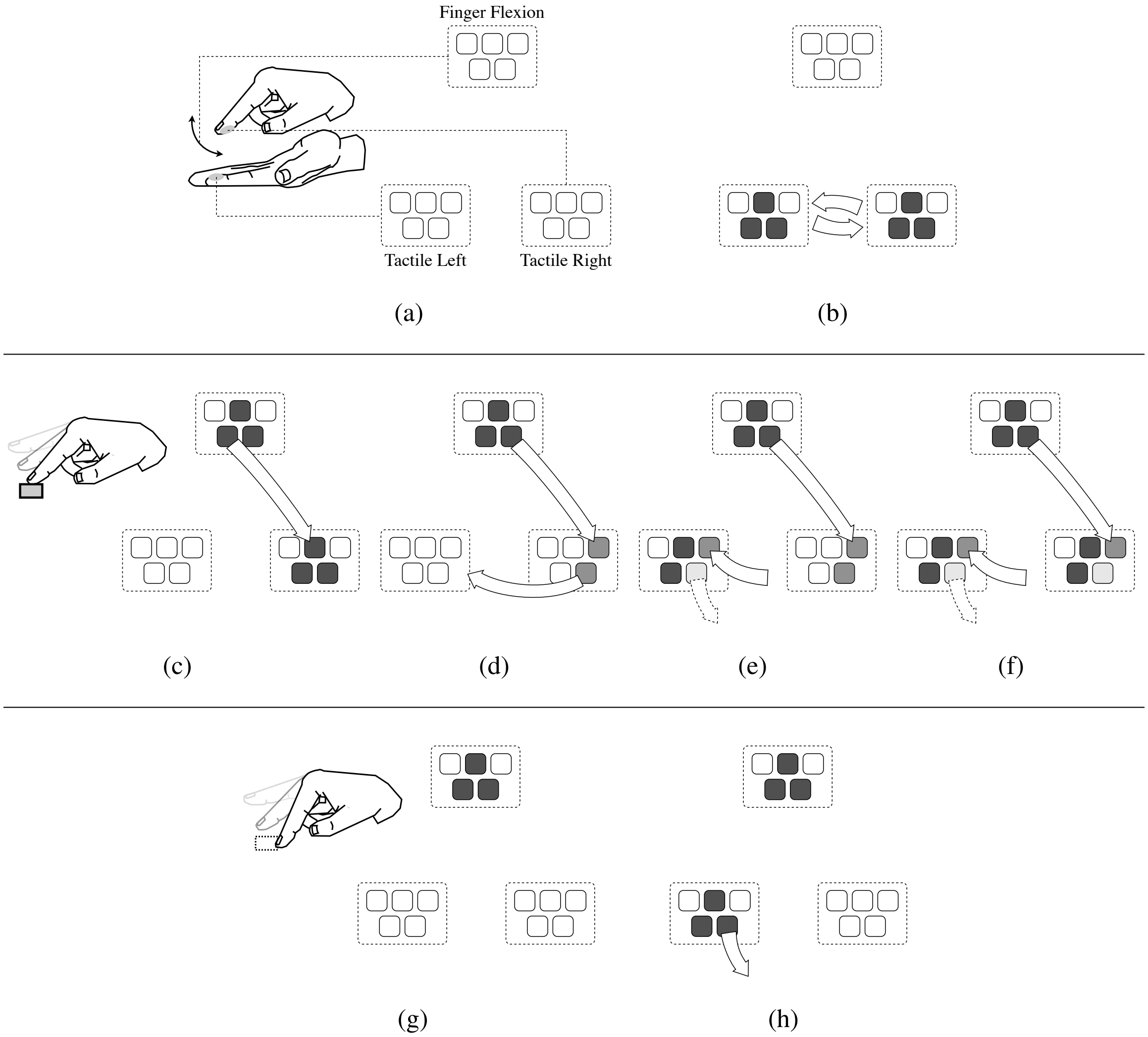


Figure 5-16: Model of sensory cancellation in human. (a) The motor modality corresponds to the flexion of the right finger, and the two sensory modality indicates the detected tactile stimulations in the left and right fingers, respectively. (b) It is assumed that humans, when their hands are closely facing each other, experience simultaneous tactile stimulation in both fingers. Accordingly, the model captures correlations of activity between the two sensory modalities. (c) A first group of subjects is trained to produce a tap on a force sensor. Consequently, the model captures a (possibly short-term) correlation between finger flexion and tactile stimulation of the same finger. Because of the correlations learned by the subjects of the first group (b, c), neural activity produced by a finger flexion flows into the tactile modality of the same finger, and propagates further into the other tactile modality (d). The neural output from the tactile modality of the passive finger is thus attenuated, irrespective of whether the active finger detects contact (e) or not (f). In contrast, a second group of subjects is trained to only produce a motion of the finger, but with no contact (g). In this case, the neural output from the tactile modality of the passive finger is not attenuated during flexion of the active finger (h).

the electric fish, precisely generated this correlation by consistently and repeatedly pairing tail bend with electric stimulation.

In the case of tactile attenuation in human, it is necessary to consider multiple cross-modal correlations. First, it is fairly reasonable to assume that human do not experience any significant correlation between movement of a finger of one hand and tactile stimulation in a finger of the other hand. In contrast, it is not too unrealistic to assume, when both hands are closely facing each other, a certain correlation of tactile stimulation between two fingers of different hands. With these assumptions, our model predicts tactile attenuation when a correlation is learned between flexion and tactile stimulation of the same finger (Figure 5-16). Let us now have a closer look at how the study was conducted by Bays et al. (2006) with the human subjects. One group of participants was first trained during a practice session to produce a tap on a force sensor with their active finger, whereas a second group was first trained to only produce a similar movement of their finger – but with no contact. Since correlation between finger flexion and tactile stimulation on the same finger is only experienced during the practice session by the first group, the predictions of our model qualitatively match the results of the psychophysical experiment: tactile stimulation on the passive finger is only attenuated for the first group, i.e. when contact is expected between the fingers.

Even though our neural architecture does not aim at providing any realistic model of the actual neural mechanisms of electric fish and human, it outlines the importance of considering the whole – both present and past – agent-environment interaction when comparing the two systems. At a short-term time scale, the difference between the sensory attenuation in electric fish and human can only be explained by postulating different mechanisms. However, a different account can be given when taking past interaction into consideration, i.e. from a longer-term time scale perspective. Indeed, we showed that dissimilarities in sensory attenuation observed with electric fish and human can be attributed to obvious differences in the way they interact with the environment, rather than to different underlying mechanisms.

Chapter 6

Dynamic Locomotion

The previous chapters described experiments conducted with the proposed neural architecture on robotic platforms equipped with various sensory systems and situated in various environments, but having essentially the same motor system. Consequently, this chapter presents a fourth series of experiments showing how the neural architecture can also be successfully used on a robot with complex body dynamics – namely on a running four-legged robot – to generate coherent, seemingly goal-directed behavior. In particular, we discuss how the presented approach allows the robot to discover an unexpected, but particularly stable strategy to achieve turning motion.

6.1 Introduction

In the previous chapters, we explored the behaviors that we could observe with the proposed minimal neural architecture on robotic platforms equipped with different sensory systems and situated in different environments. However, the motor systems of the robots described so far – i.e. the physical processes producing locomotion – remained essentially the same: forward and turning motion was simply produced by two rotating wheels in contact with the ground.

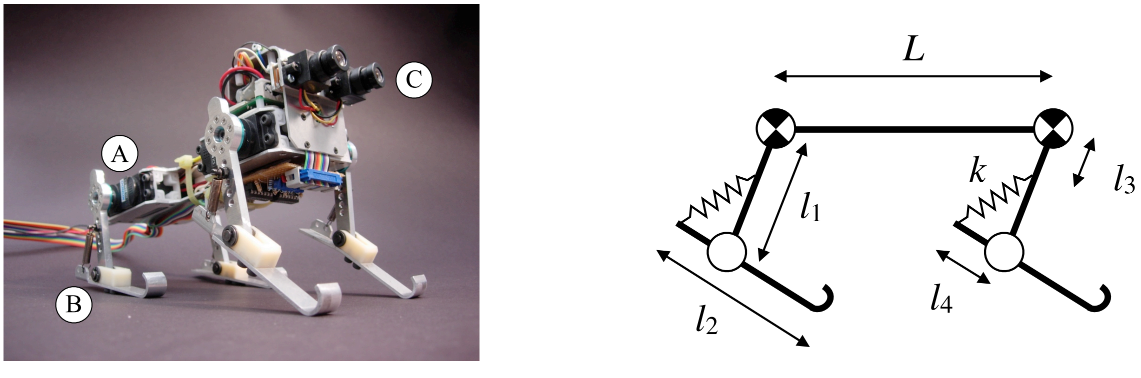
The reasons why we first – and up to now exclusively – investigated robots with wheeled motor systems are relatively obvious: wheeled robots are readily available, intuitive and easy to use, and particularly well suited for flat surfaces – the typical kind of environments obtained with little cleaning effort in an office room. The same reasons also explain why wheeled robots are so often the tools of choice for researcher working with mobile robots.

There are however several fundamental issues inherent to wheeled locomotion that one has to be aware of. First, wheeled robots can only operate in neatly prepared environments (essentially flats and obstacle-free environments). Even though this is not a problem per se, this might severely restrict our view when trying to understand natural agents living in much more complex and dynamic environments.

Second, wheeled robots can only display a limited behavioral diversity: no matter how complex the sensory and control systems are, the robots still just drive around on a flat ground. They cannot explore other, possibly unexpected ways to interact with their environment and achieve more adaptive behaviors. As a consequence, wheeled agents also have limited autonomy (limitations many of us have had the unfortunate experience as one's car got stuck in snowy or icy conditions).

Third, wheeled systems are carefully designed to produce specific behaviors. For instance, the positions and the sizes of the wheels in a differential steering are explicitly designed so that forward and turning motions are proportional to the average and difference of wheel speeds, respectively. This somehow artificial "grounding" of behavior in motor activity – which evokes the problem of grounding symbols in sensory activity (Harnad, 1990) – has a definite cognitivistic flavor. This also strongly contrasts with the dynamical systems approach to locomotion and behavior (Smith and Thelen, 1993; Kelso, 1995; Kuniyoshi et al., 2004): for instance, the speed of a wheeled robot is proportional to the turning speed of the wheels and can be thus varied arbitrarily, whereas in dynamical systems (such as animals and humans) there are only few, typically discrete preferred speeds that correspond to the different attractor states of their specific physical systems; even though the speed at which animals can move their legs can continuously vary over a certain range, animals usually settle to a particular, clearly identifiable gait with preferred speed such as walking, trotting or galloping (McMahon, 1984).

Finally, since wheeled robots have admittedly very simple body dynamics, the intrinsic dynamics of the sensorimotor signals remains at a limited level of complexity. For instance, a camera mounted on a wheeled robot will provide a much more stable, and much less noisy visual sensory input than a camera mounted on a running legged robot and being thus constantly shaken. Also, some information structure in the sensorimotor signals – such as the correlation, picked up by the neural network, between visual flow and motor activity – was observed due to the straightforward physical processes underlying the motion of wheeled systems. For instance, activity of the turning motor neuron almost always generated – because of the simple body dynamics of the system – a proportional rotation of the robot, which in turn could produce a relatively homogeneous and noise-free optical flow detected by the camera. The complexity in the dynamics of the system was increased in Chapter 5 as a significant amount of noise was introduced in the motor signals. Nevertheless, one can still justifiably wonder whether such artificially limited complexity in the body dynamics of the agent is a necessary condition to generate the cross-modal correlations on which the



(a)

(b)

Figure 6-1: (a) Photograph of the quadruped robot "Puppy". The robot has four identical legs, each of which consists of one servomotor (A) actuating a series of two limbs connected through a passive elastic joint (B). The robot is also equipped with a vision system (C). Only one of the two cameras pointing forward is used for the experiment. (b) Schematic of the robot. Empty circles denote passive joints and crossed circles denote joints controlled by servomotors. The specifications of the robot are given in Table 6.1.

proposed neural architecture relies to eventually produce the coherent observed behaviors.

The focus of this chapter is therefore to study whether coherent behaviors can still be observed when our homogeneous neural architecture is implemented on a robotic agent with more complex body dynamics. In what follows, we first introduce the running quadruped robot used for the experiment. Then, we describe and analyze how coherent and seemingly intentional behavior is generated as the robot interacts with its environment. In particular, we examine the learned connectivity pattern in the neural network that produces the turning behavior of the robot, as it turns out to reveal an original and particularly stable strategy.

6.2 Robot Experiment

6.2.1 Agent and Environment

The robotic platform used in this chapter is the robot "Puppy," a running robotic dog designed and built by Fumiya Iida (2005a). The robot, shown in Figure 6-1, has four identical legs, each of which consists of one servomotor actuating a series of two limbs connected through a passive elastic joint.

The quadruped robot displays interesting characteristics resulting from its particular morphology (Iida and Pfeifer, 2004; Iida et al., 2005; Iida, 2005b), which we briefly discuss here. First, locomotion can only be achieved dynamically. Indeed, since each leg has one

Parameter	Description	Approx. Value
L	Distance between axes of front and hing legs	130 mm
W	Distance between left and right legs	100 mm
l_1	Length of upper leg limb	50 mm
l_2	Length of lower leg limb	75 mm
l_3	Spring attachment on upper leg limb	25 mm
l_4	Spring attachment on lower leg limb	20 mm
k	Spring constant	700 N/m
m	Mass of the robot	550 g

Table 6.1: Mechanical specifications of the robot.

single actuated degree of freedom, the only way the robot can lift its legs off the ground is by delivering enough energy through the motors to make the whole body jump. This first property is characteristic of a dynamical system. It shows not only the non-trivial relation between actuation and behavior, but also that the observed behaviors correlate with typically discrete attractors of the physical system: for instance, the attractor corresponding to the robot staying put is only escaped when the legs are actuated with enough energy.

Second, stability is achieved through the material properties of the legs – especially the compliance of the passive joints – rather than by actively controlling the positions of all limbs. For instance, an inadequate position of the lower limb of a leg (which is only passively attached to the upper limb) during the flight phase will automatically be corrected by the spring on contact with the ground. In particular, this characteristic allows the robot to be controlled in an open-loop manner (i.e. without any sensory feedback) over a continuous range of control parameters. By simply actuating periodically the motors back and forth, the quadruped robot put on the ground will automatically settle after a few steps into a natural and stable running rhythm.

Third, the elasticity of the legs, partially absorbing and releasing energy during contact with the ground, allows to achieve not only stable, but also rapid and energy efficient locomotion. The importance of such elastic properties in muscle-tendon systems has been long recognized in biomechanics, where it has a particular significance in theoretical models for the locomotion of legged animals (Alexander, 1990; McMahon and Cheng, 1990).

In summary, the locomotion characteristics of the robot “Puppy” nicely illustrate an important and challenging principle in adaptive robotics: by properly exploiting its body dynamics and its ecological niche, a robotic system can achieve efficient behavior with very little control (Yamamoto and Kuniyoshi, 2001; Collins et al., 2005), or even no control at all, as in the case of passive dynamic walkers (McGeer, 1990; Collins et al., 2001).

Motor Modality

As discussed previously, dynamic locomotion is achieved by periodically moving back and forth the servomotors actuating the legs of the robot. Thus, rather than representing the instantaneous position or speed of each servomotor in the motor modality, we choose instead an oscillatory position control scheme where the rhythmic oscillations of the leg motors are modulated by steady activity of the neurons in the motor modality. This scheme – generally referred to as central pattern generators – is inspired by neurophysiological studies on how animals produce motions to run, swim or fly. A central pattern generator (CPG) is a neural unit that spontaneously produces a rhythmic output signal whose amplitude, frequency, phase and baseline offset is typically modulated by steady input signals. CPGs are observed in various animals (for a review, see e.g. Grillner, 1996), and also widely used to control multi-segmented robots (e.g. Ijspeert, 2001).

The target angular position $P_i(t)$ of motor i at time t is given by

$$P_i(t) = A_i \sin(2\pi f \cdot t + \phi_i) + B_i$$

where A_i is the amplitude, ϕ_i the phase offset, and B_i the set point of the oscillation.

Instead of specifying directly the parameters of oscillation for each individual motor, we choose a set of motor variables that represent the *differences* of parameter values between left-side and right-side motors, as well as between fore and hind motors. For instance, the oscillation amplitudes A_i of the four motors are defined as follow:

$$\begin{aligned}
A_{\text{fore,left}} &= A_0 - \frac{1}{2}\Delta A_{\text{lat}} - \frac{1}{2}\Delta A_{\text{long}} \\
A_{\text{fore,right}} &= A_0 + \frac{1}{2}\Delta A_{\text{lat}} - \frac{1}{2}\Delta A_{\text{long}} \\
A_{\text{hind,left}} &= A_0 - \frac{1}{2}\Delta A_{\text{lat}} + \frac{1}{2}\Delta A_{\text{long}} \\
A_{\text{hind,right}} &= A_0 + \frac{1}{2}\Delta A_{\text{lat}} + \frac{1}{2}\Delta A_{\text{long}}
\end{aligned}$$

ΔA_{lat} and ΔA_{long} are the lateral and longitudinal differences of amplitude, and A_0 is the average amplitude. The other motor parameters (i.e. the set points B_i and the phase offsets ϕ_i) are defined accordingly. A description of the 8 state components in the motor modality – whose values are represented by the activity of 8 neurons in the motor modality – is provided in Table 6.2.

Note that the frequency of oscillation f is constant for all motors. There are two reasons for that. First, the range of exploitable frequencies is limited by the physical properties of

Component	Description
A_0	Average amplitude
ΔA_{lat}	Lateral amplitude difference
ΔA_{long}	Longitudinal amplitude difference
B_0	Average set point
ΔB_{lat}	Lateral set point difference
ΔB_{long}	Longitudinal set point difference
$\Delta \phi_{\text{lat}}$	Lateral phase offset difference
$\Delta \phi_{\text{long}}$	Longitudinal phase offset difference

Table 6.2: State components of the motor modality.

the motor system. Indeed, because of the limited power of the servomotors, the maximum reachable oscillation amplitude drops significantly with frequencies higher than approximately 3 Hz. Also, at frequencies lower than about 2 Hz, the oscillation of the upper parts of the legs gets absorbed by the springs. As a result, the lower parts of the legs do not lift off the ground, and the robot, not running anymore, does not move any longer. Second, the phase difference, which will reveal interesting effects described later, becomes mathematically ill-defined as soon as the motors oscillate at individual frequencies. Therefore, we choose to set all oscillation frequencies to a constant value of $f = 2.5$ Hz.

Visual Modality

The quadruped robot is equipped with a vision system consisting of a camera attached to the body and pointing in the forward direction (see Figure 6-1(a)). The current state in the visual modality is represented by an array of neurons, whose activities correspond to the 32×24 pixel values of the grayscale image extracted from the camera. Similar to the previous experiments, the state change in the visual modality is defined as the estimated horizontal and vertical components of the visual flow approximated at each pixel.

"Reward" Sensor

The robot is also equipped, as in the previous chapter, with a simulated sensor whose value is increased when the agent receives an external "reward" (see below, as well as Foonote 2 on page 87). Since the activity of the state change neuron in this modality is relatively sparse compared to the other modalities, the learning rate for the corresponding synaptic couplings is increased to $\eta = 0.1$. (The learning rate for all other synaptic couplings takes the usual value of $\eta = 0.01$.)

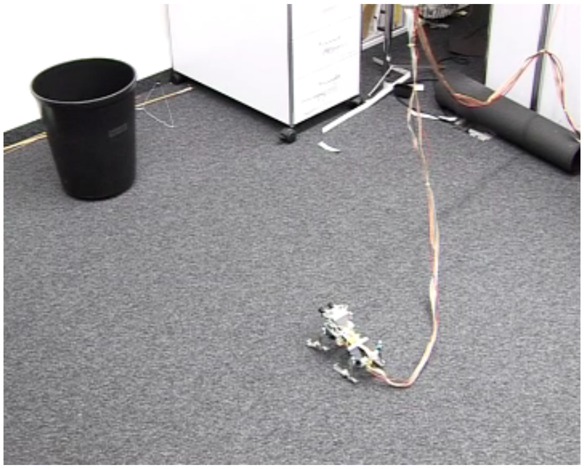


Figure 6-2: The environment for the experiment discussed in this chapter is an unmodified office floor. The "reward" signal is delivered to the robot when it faces a large black bin, as shown in the figure.

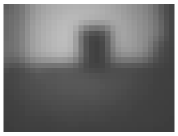


Figure 6-3: Graphical representation of the synaptic weights coupling the reward modality to the visual modality, showing the visual state correlated to increase of reward signal.

Environment

The environment consists of an unmodified office floor, as shown in Figure 6-2.

6.2.2 Observed Behavior

During the initial phase of the experiment, the neurons in the motor modality are randomly activated, thus producing arbitrary motions of the robot. This initial phase – during which, metaphorically speaking, the robot randomly explores its interaction with the environment – allows the neural network to learn the basic cross-modal correlations. In the presented experiment, the reward is delivered when the robot is facing a large black bin placed in the environment (as illustrated in Figure 6-2). The synapses coupling the reward modality to the visual modality therefore learn a correlation pattern between increase of reward signal and a visual input corresponding to a black object in the center of the visual field (see Figure 6-3).

After this initial phase, the neuron corresponding to the desired state in the reward modality is externally activated, and the robot is let to move on its own. The observed behavior, generated from the propagation of neural activity across the network, is illustrated in Figure 6-4. The robot systematically turns towards any black object that is placed in the center of its field of view, and follows the object as it is moved around.

As a matter of fact, this behavior closely resembles the object following behavior observed with the AMouse robot and analyzed in Chapter 3. Indeed, the agent learns in both cases an association between the onset of one sensory input (tactile signal in Chapter 3, or reward signal in this chapter) and a visual input corresponding to black pixels in the center of the visual field. The propagation of activity across the neural network follows therefore a qualitatively similar structure (see Figure 6-5): neural activity externally generated in the reward modality propagates to the visual modality, projecting the image correlated to increase of reward, i.e. black pixels in the center of the desired visual field; the neural activity then further propagates to the motor modality, according to the difference between current and desired visual states. These are the only paths for propagation of neural activity, since there is no other significant cross-modal correlation.

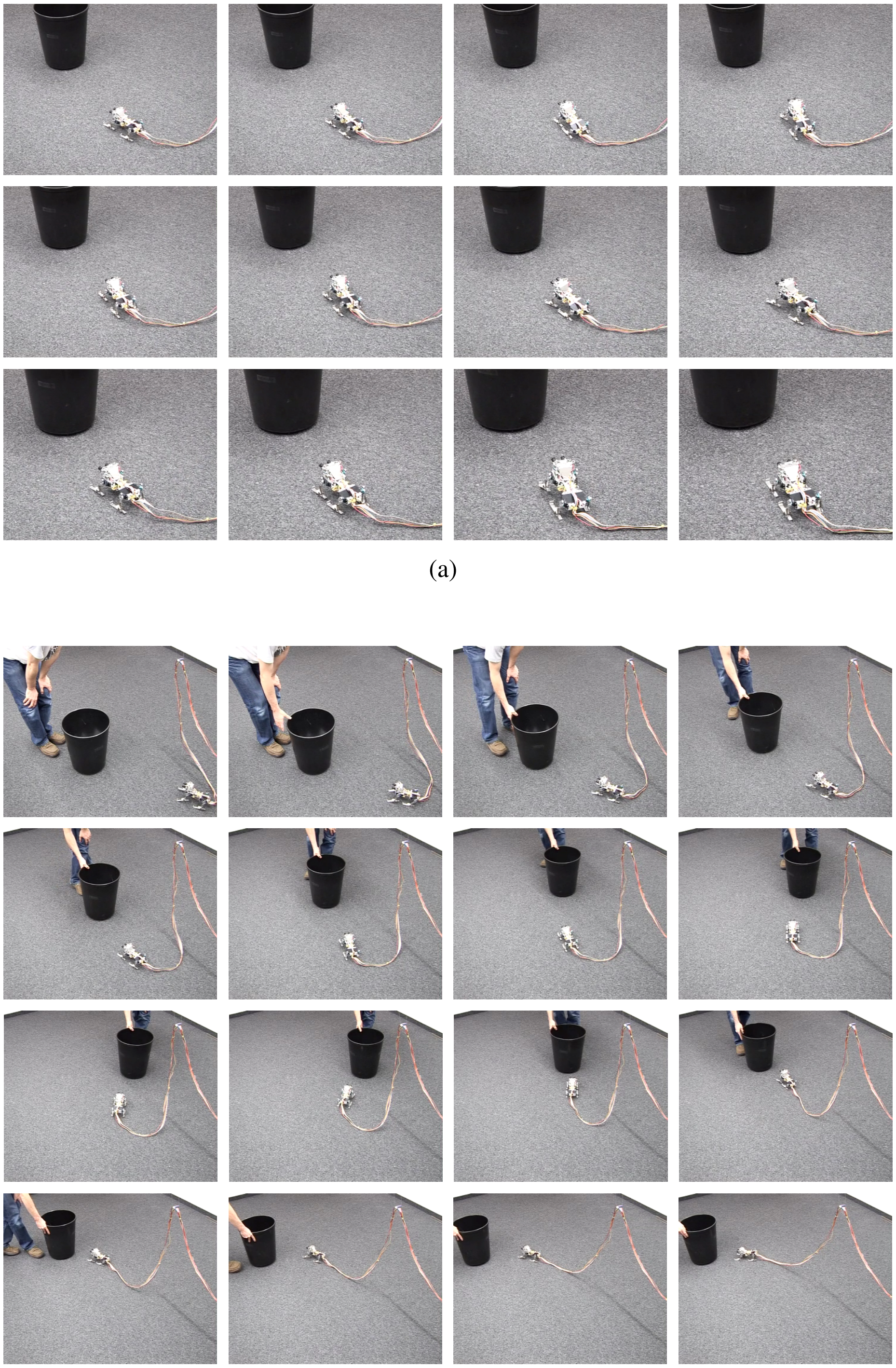
The following section focuses on the second cross-modal correlation learned by the neural network, namely the correlation between visual flow and motor state, which allows the robot to adjust its heading direction towards the target object.

6.2.3 Turning Behavior

From an observer perspective, the quadruped robot learned how to turn: indeed, when the target object is moved laterally, motor activity – and thus the oscillations of the legs – is modulated by the neural network in such a way that the robot turns toward the object. This section investigates which pattern of motor activity the neural network generates when the robot is observed to turn.

The key aspect of the network connectivity is the correlation between perceived visual flow and motor activity, which is captured by the synaptic weights coupling the visual modality to the motor modality. Figure 6-6 shows a graphical representation of these weights, illustrating the visual flow correlated to each motor control parameter. Clearly, the neural architecture captures a significant correlation only between visual flow and the motor parameter corresponding to lateral phase offset difference ($\Delta\phi_{\text{lat}}$). This means that the quadruped robot learns a control strategy for turning that consists in modifying essentially the phase difference between the oscillations of the left and the right legs.

This result is particularly surprising since this strategy contrasts with the intuition that turning is most easily achieved by modulating the amplitudes, i.e. by increasing the amplitude of leg oscillations on one side while decreasing it on the other side.



(a)

(b)

Figure 6-4: Observed "object following" behavior. (a) The robot adjusts its orientation towards a black object in its field of view. (b) When the object is moved around, the robot follows it.

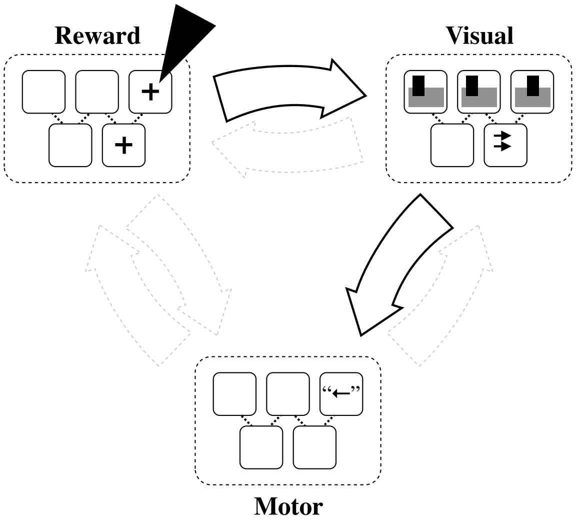


Figure 6-5: Propagation of neural activity in the network leading to the observed “object following” behavior. The black arrow indicates activity generated externally in the desired population of the reward modality. This neural activity propagates into the visual modality, projecting the image of a black object in the center of the visual field. If an object is visible, activity propagates further into the motor modality, generating the object following behavior. In this example, the object is perceived slightly on the left of the visual field. The pattern of activity generated in the motor modality producing the appropriate turning motion (illustrated by a quoted arrow) is discussed in Section 6.2.3.

To better understand this result, we systematically quantify the turning rate of the robot as a function of various motor control parameters. Figure 6-7 clearly shows that the turning speed is indeed most easily and robustly controlled with the lateral phase difference, the relation between the two quantities being almost linear in the considered range (Figure 6-7(c)). In contrast, when the other motor control parameters are varied, the turning speed of the robot either does not change significantly, or displays no linear relation: for instance, as the lateral amplitude difference is steadily increased (Figure 6-7(a)), the turning speed of the robot even changes its sign – the robot is turning in one direction for a small amplitude difference, and in the opposite direction for a larger amplitude difference. Moreover, Figure 6-8 shows that when both phase and amplitude differences are varied simultaneously, the turning speed of the robot is predominantly influenced by the phase difference.

Why is it then that the turning speed of the robot is – contrary to intuition – more influenced by the phase, rather than by the amplitude of leg oscillations? The main reason comes certainly from the compliance of the legs. On the one hand, most difference in oscillation amplitude gets absorbed by the springs. Rotation of the body is then only observed when the difference of amplitude is so large that the legs on one side almost do not move anymore, thus providing a constant contact point for pivoting around it. On the other hand,

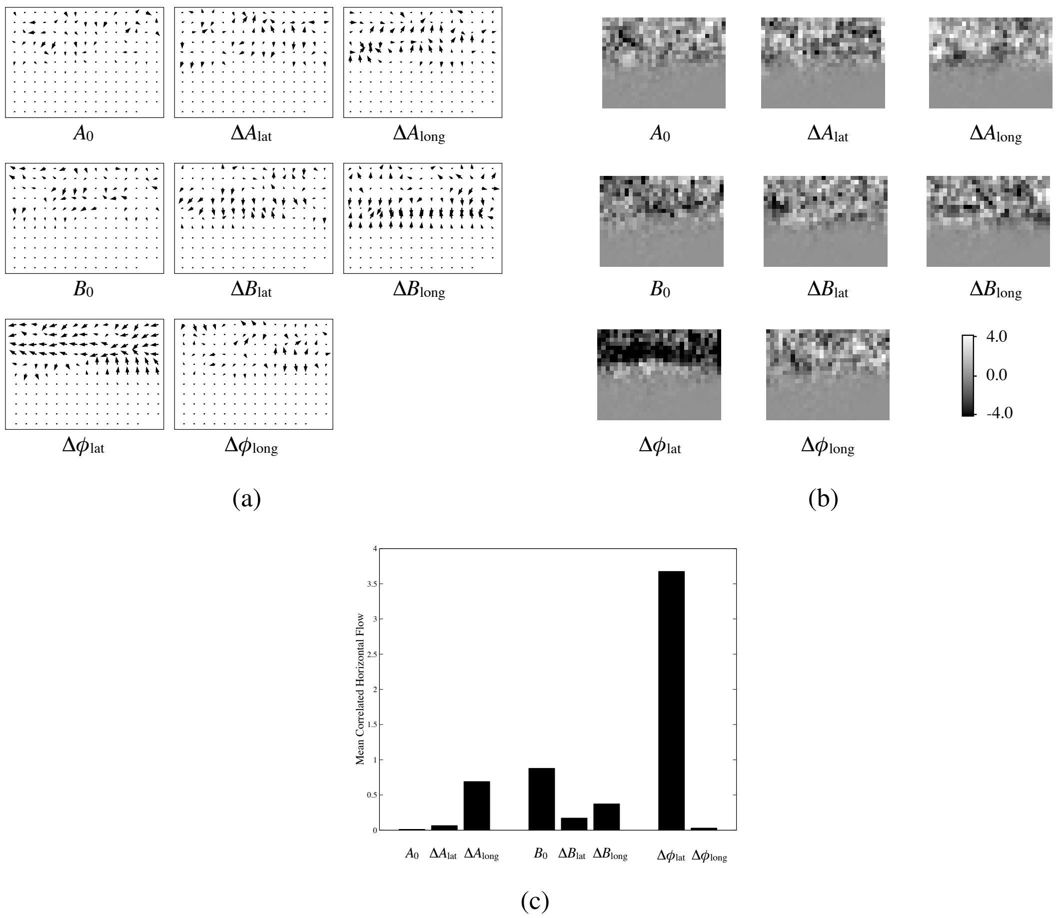
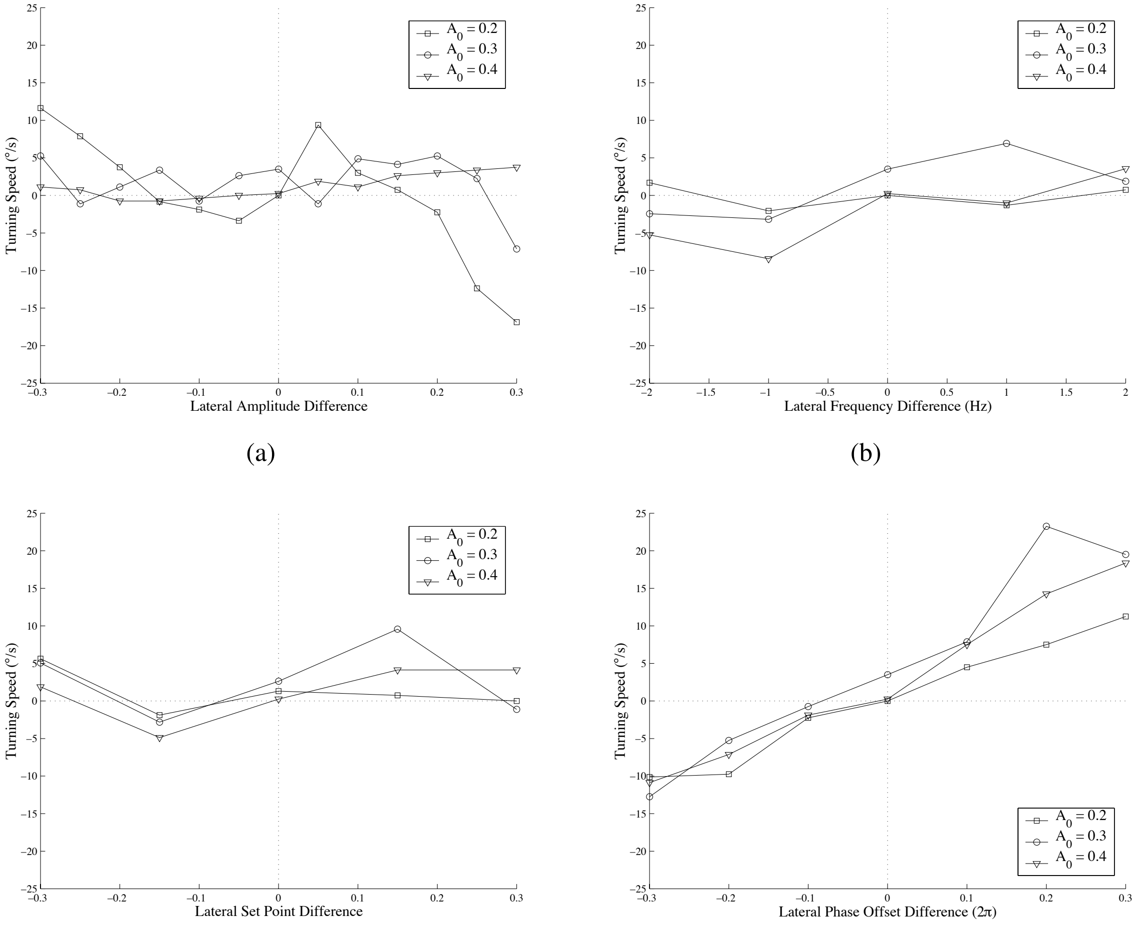


Figure 6-6: Graphical representation of the synaptic weights coupling the visual modality to the motor modality, showing (a) the visual flow field, and (b) only the horizontal component thereof, correlated to each of the eight components of the motor state. (c) Average horizontal component of the visual flow correlated to each motor component (absolute value). The learning and forgetting rates are $\eta = 0.01$ and $\varepsilon = 0$, respectively.

the same compliance is exploited for generating rotation of the body when the legs do not move in phase. Then, because of the shifted times of contact with the ground between left and right legs, the body starts swinging laterally, thus creating a dissymmetry in the reaction forces with the ground – a break of symmetry that eventually produces the observed turning motion.

6.3 Discussion

In contrast to the series of experiments conducted in the previous chapters, where the locomotion of the agents was the result of quasi-static processes (i.e. wheels turning in contact with the ground), locomotion of the quadruped robot used in this chapter is the prod-



(a) (b) (c) (d)

Figure 6-7: Turning speed of the robot as a function of (a) amplitude difference ΔA_{lat}, (b) set point difference ΔB_{lat}, (c) phase offset difference $\Delta\phi_{\text{lat}}$, and (d) frequency difference between left and right leg oscillations. The turning speed is measured for three different values of the average amplitude A_0. (The angular position of each motor is scaled such that 1 unit corresponds to approximately 90°.)

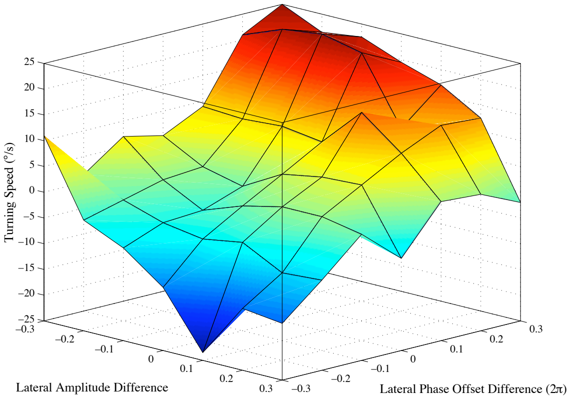


Figure 6-8: Turning speed of the robot as a function of both lateral amplitude difference ΔA_{lat} and later phase offset difference $\Delta \phi_{\text{lat}}$.

uct of highly dynamical, non-linear physical processes: oscillations of the legs generate through the elasticity of the passive joints a running movement of the robot body, which is in turn transformed into translational or rotational movement by the frictional ground reaction forces.

The results presented in this chapter demonstrate that, despite the complex body dynamics and the high level of noise in sensory signals (the camera is constantly shaken due to the rhythmic running of the robot), the proposed homogeneous neural architecture can nevertheless capture some cross-modal correlations that lead to the generation of coherent, seemingly intentional behavior: the robot is observed to move towards and follow an object, placed in its visual field, which has previously been associated with sensory stimulation corresponding to a reward signal.

A point of particular interest is the way the robot actually generates turning motion when it follows the target object. Intuitively, we would expect turning motion to be achieved by swinging the legs on the left side and on the right side with different amplitudes (or frequencies). This is indeed often how legged robots are controlled for turning. However – and somehow surprisingly – it turns out that the quadruped robot discovers and uses a completely different strategy. The learned connectivity of the neural network reveals that rotation of the body is actually achieved through modulation of the relative phase of oscillation between left and right legs – a control strategy shown furthermore to be especially robust, in fact much more robust than the amplitude or frequency modulation strategies.

Finally, the presented experiment illustrates how the body dynamics of an embodied system can be exploited – despite its apparent complexity – to achieve robust behavior with a minimal amount of control. The elastic properties of the legs not only confer an inherent stability to the quadruped robot, but also generate correlated sensorimotor data:

activity of the motor neuron representing the lateral phase difference of the leg oscillations is – because of the particular physical interaction of the legs with the ground – correlated to visual sensory activity representing horizontal visual flow. By simply picking up and exploiting such correlations, the proposed neural architecture is able to generate a stable and robust behavior.

Before concluding this chapter, it is worth briefly discussing how observed behavior – in particular turning behavior – is "grounded" in the system. From an observer perspective, it perfectly makes sense to make statements such as "when the robot sees the target object on one side, it activates the corresponding behavior, namely a turn in the direction of the object." However, using a similar statement to describe the underlying mechanisms of the system would constitute a category error. Indeed, what is interpreted as the "image of the target object" (i.e. an input in a single modality) is for the situated robot a pattern of visual sensory stimulation correlated to an increase of reward signal (i.e. an association between two different sensory modalities). Similarly, the observed turning behavior cannot be simply reduced to motor activity. The pattern of neural activity produced in the motor modality as the robot is observed to turn can only be understood when considering the physical interaction of the agent with its environment: modulating the phase of leg oscillations only gives rise to a rotation of the robot body because the particular morphology and elasticity of the legs produce certain reaction forces of the ground. In contrast to the experiments described in the previous chapters, where the motor system of the wheeled robots was carefully designed, this chapter illustrates the fact that with embodied systems, there is not necessarily a straightforward relationship between motor activity and observed behavior. In other words, turning behavior is not "grounded" in motor activity as it often is the case with wheeled robots, but emerges from the sensorimotor interaction of the agent with its environment: the pattern of motor activity produced by the neural network arises from the correlation between sensory and motor activity captured during the physical interaction of the robot with its environment. Metaphorically speaking, "turning" is for the situated agent not a particular motor activity but rather a complete multi-modal sensorimotor experience.

6.4 Conclusion

This chapter extends the previous series of experiments by showing that the same neural architecture can also be used on a robot with a more complex body dynamics, namely a running quadruped robot with elastic legs, to generate a stable object following behavior. In particular, the presented experiment shows how a simple and unbiased learning mechanism can, by exploiting multi-modal correlations generated by the embodied agent, discover an original and remarkably stable strategy for generating turning motion.

Chapter 7

Delayed Reward Learning without Working Memory

The main goal of this chapter is to investigate the behaviors that are observed when a robotic agent, using the same neural architecture as in the preceding experiments, is engaged in a more complex environment. Inspired by maze experiments performed to study navigation and learning in rodents, we choose an environment consisting of a T-shaped maze. At the junction, a tactile cue indicates which arm contains the reward. This chapter not only describes how the robot is observed to spontaneously develop a strategy to solve the delayed reward learning task, but also demonstrates that the robot does so without possessing any explicit working memory of previous events and actions. By investigating the underlying mechanisms, namely the dynamic and reciprocal interaction between the neural and the physical dynamics of the system, this chapter provides an original case study shedding new light on the neural basis of memory, and supporting recent concepts in memory research developed as alternatives to the common but problematic view of memory as stored information.

7.1 Introduction

The preceding chapters of this thesis described various behaviors that were observed, under different conditions, with different agents endowed with the proposed neural architecture. In each experiment, the observed behavior resulted from neural activity propagating in the network through a cascade of cross-modal associations. These cross-modal associations were correlations of sensorimotor activity captured by the network essentially during an initial exploratory phase, during which the agent randomly interacted with its environment. However, once the basic connectivity of the neural architecture was shaped during the initial phase (i.e. once the cross-modal correlations were learned), the qualitative structure of the neural network did not change anymore: even though the Hebbian learning mechanism was continually modifying the synaptic weights throughout the whole experiment, the patterns of cross-modal correlations remained qualitatively the same – and so did the observed behaviors.

In contrast, this chapter investigates how the observed behavior can progressively change as the agent autonomously interacts with its environment. The set of experiments presented in this chapter illustrates how – in a particular and more dynamic environment – the initial behavior produced by the agent will generate, through interaction with the environment, qualitatively new cross-modal correlations, which in turn will modify the behavior displayed by the agent – and so on further generate new cross-modal correlations. This reciprocal interplay between the internal neural dynamics and the physical interaction with the surrounding world, taking place at different time scales, illustrates how the proposed neural architecture, even though it "only" captures instantaneous sensorimotor correlations, is able to generate adaptive and more complex behaviors.

In this chapter, behavioral adaptivity is observed as the agent is engaged in a more complex task environment. The chosen environmental set-up is inspired by typical maze experiments performed with rodents to study their navigation, memory and learning capabilities: it consists of a T-shaped maze, with a tactile cue at the junction indicating which arm of the maze contains the reward. The reason for this particular inspiration stems from the context in which the AMouse robot was devised. The AMouse robot – that we already met in Chapters 3 and 5 – was indeed developed as a model to study particular aspects of rodents[1] such as their remarkable whisker tactile system (Fend et al., 2006; Hipp et al., 2006; Arabzadeh et al., 2006).

On the one hand, using robots to study natural behaviors is obviously a challenging scientific venture (Webb, 2001). For instance, the current technological constrains clearly prevent robots from reproducing more than just a tiny fraction of the sophisticated sensorimotor morphology of any animal. On the other hand, finding the appropriate level of abstraction can turn artificial embodied systems into unique conceptual platforms from which innovative insights into natural behaviors can emerge. The present chapter illustrates this latter point by showing how the investigation of an artificial agent, engaged in a behavioral paradigm used to study rodent behavior, can indeed reveal new and intriguing perspectives on learning and memory.

The results presented in this chapter – i.e. the behaviors of the AMouse robot observed in the T-maze – are remarkable in many respects. First, the robot learns with time to solve

---

[1] Hence the name Artificial Mouse.

the task, i.e. learns to rely on the tactile cue to turn into the arm of the maze where the reward is located. This observed behavior is in itself already quite surprising, since the robot does not know anything about the environment or the task, nor that a particular sensory stimulus indicates the position of the future reward. Second, a closer inspection into the mechanisms underlying this behavior allows reconsidering some generally accepted assumptions about delayed reward learning. Third, the results demonstrate that, contrary to the generally accepted view, memory is not necessarily information stored in some way or other inside the agent. Rather, this chapter shows that memory can be, at least partly, off-loaded into the environment. Finally, this series of experiments shows that even a priori irrelevant aspects of the environment can play a crucial role in shaping the robot learning behavior.

The outline of the chapter is as follows. Section 7.2 briefly introduces the maze and the task paradigm used during the subsequent series of experiments, highlighting the close relation between mazes and the study of learning and memory in rodents. Then, Section 7.3 describes in details the actual experiments performed with the robotic and simulated agents engaged in various T-maze tasks. The implications of the results are discussed in Section 7.4. In addition, Section 7.5 focuses on their connections to traditional and more recent concepts of memory. Finally, Section 7.6 concludes the chapter.

7.2 Mazes, Learning and Memory

Rodents are extremely gifted navigators. In particular, they are amazingly talented at solving experimental mazes[2], as shown since at least the beginning of the 20th century by the thousands of studies that examined how rodents run different types of mazes, from T-mazes to radial arm mazes (Olton, 1987) to water mazes (Morris, 1984). These maze experiments have been used to study spatial learning and memory in rats and mice (for reviews, see e.g. Hodges, 1996; D'Hooge and De Deyn, 2001; Crusio and Schwegler, 2005) and have helped uncover general principles, such as the existence of so-called place cells – neurons whose activity only depends on the spatial position of the rat in its environment (O'Keefe and Dostrovsky, 1971; O'Keefe, 1979).

This chapter focuses on one of the simplest kinds of maze, namely a T-shape maze as the one illustrated in Figure 7-1, and considers the following scenario. The rat is initially placed in the central arm (the base of the 'T'). The problem for the rat, when it reaches the

---

[2] This holds especially for rats. Their maze-running ability comes most probably from their evolutionary history: rats are small burrowing rodents that have spent millennia digging and finding their way around underground tunnels. It's no wonder they have a knack with mazes.

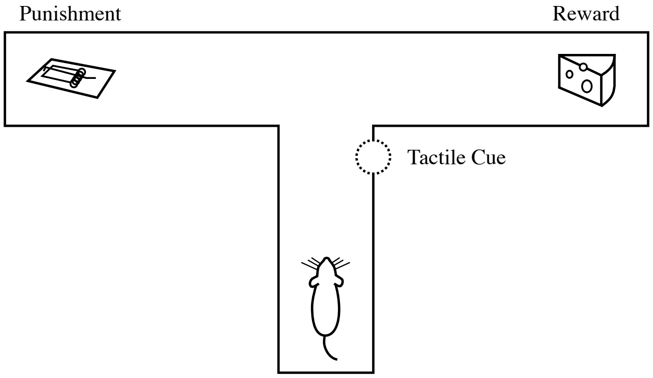


Figure 7-1: Schematic illustration of a typical T-maze. The rat is initially placed in the central arm of the maze. At the junction, a tactile cue indicates the side of the reward, which is randomly reassigned to the left or to the right. The task for the rat is to learn that the position of the tactile cue indicates the arm of the T-maze where the reward is, and by using that cue to always turn into the correct arm. Note that the rat cannot see or otherwise perceive the reward from the junction.

junction, is whether to turn left or right in order to get the reward (e.g. a pellet of food), which is either at the end of the left arm, or at the end of the right arm. At one of the corners of the junction, there is a tactile cue (e.g. a vertical stick) that the rat can detect with its whiskers. If the tactile cue is on the left, the reward will always be at the end of the left arm of the T-maze; if it is on the right, the reward will be at the end of the right arm. During the experiment, the position of the reward, together with the tactile cue, is randomly switched from left to right, and each time the rat is given a chance to find the reward. The task for the rat is to learn that the position of the tactile cue indicates the arm of the T-maze where the reward is, and to use that cue at the junction to turn into the correct arm.

This delayed reward learning paradigm, where the difficulty is to find out at what point in time the correct or incorrect decision has been taken – also referred in the literature to as "credit assignment problem" (Minsky, 1961) or "road sign problem" (Rylatt and Czarnecki, 2000) – illustrates the tight relation between learning and memory. For instance, a typical assumption is that if the rat can learn the task, it must somehow remember the position of the cue and its decision (whether it turned left or right) until it receives the reward: otherwise, how can the rat learn anything if it does not remember, when it receives the reward, in which direction it just turned, and on which side the tactile cue was perceived?

This assumption, which constitutes the base of reinforcement learning (Connell and Mahadevan, 1993; Kaelbling et al., 1996; Sutton and Barto, 1998), seems so obvious that it is, to the best of our knowledge, accepted unquestioningly[3]. In short, delayed reward

---

[3] For instance, Kim (2004) even analysed the memory requirements for T-maze tasks, i.e. how many bits

learning requires a working memory – where “working memory” refers here to the ability to remember previous decisions and events. Yet, the series of experiments presented in the next sections will shed a new and unexpected light on this very assumption.

7.3 T-Maze Experiments

In this section, we describe in details the behavior of a particular robotic platform engaged in a T-maze task similar to the one described previously. The robot, having the same neural architecture as in the previous chapters, is observed to progressively improve its performance and to eventually learn to solve the delayed reward task. In other words, the robot seems able to learn that the tactile cue indicates the position of a later reward, and to use this information to always reach the reward.

We then analyze the mechanisms underlying the observed performance. In particular, we show that the adaptive and seemingly intentional behavior of the robot is actually produced by a reciprocal interaction between the internal dynamics of the neural architecture and the physical dynamics of the agent behaving in its environment.

7.3.1 Agent and Environment

Experiments are performed with the AMouse robot, shown in Figure 7-2, equipped with the following modalities:

Tactile modality

The left and right whisker arrays are used to detect tactile stimulation on either side of the robot. The tactile state consists of two components, whose binary values indicate whether any whisker is stimulated on the left side or on the right side, respectively. Note the slightly modified morphology of the robot (Figure 7-2) compared to Chapter 3 (Figure 3-1 on page 41): the two whisker arrays are more oriented to the sides in the present chapter, so that the robot can better detect the lateral tactile cue.

Visual modality

The current visual state consists of an array of 100×50 neurons, whose activities correspond to the pixel grayscale values of the panoramic image extracted from the omnidirectional camera. As in all previous chapters, the state change is defined in the visual modality

---

of information the agent needs to store to solve the task.

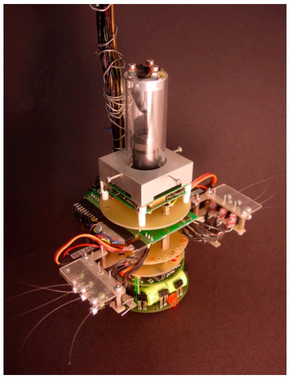


Figure 7-2: The AMouse robot.

as the horizontal and vertical components of the local visual flow computed at each pixel (see Appendix 3.A).

Reward modality

To detect the reward (or conversely, the punishment), the robot is supplied with an extra "reward sensor," similar to the one used in Chapters 5 and 6. Its value is increased when a reward is given, and decreased when a punishment is given.

Motor modality

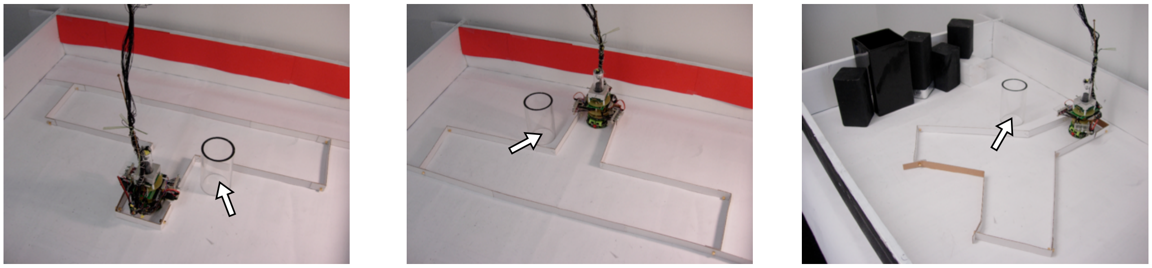
The state in this modality consists of one motor neuron whose activity corresponds to the steering angle of the robot. The forward velocity of the robot is kept at a constant positive value. That way, the robot never stay put but has a drive to always move forward.

Infrared modality

The Khepera platform, which constitutes the base of the AMouse robot, is equipped with six infrared proximity sensors, regularly arranged on the front half of the body. For the sake of clarity, this sensory modality will not be mentioned in the discussion before Section 7.3.5, where its role in the neural architecture will then be analyzed.

Network parameters

As usual, the default learning rate of the neural network is set to $\eta = 0.01$, and the learning rate of the synapses coupled to modalities with sparse activity (here, the tactile and the reward modalities) is set to $\eta = 0.1$. We choose in this chapter a non-zero forgetting rate of $\varepsilon = 0.1$ (see Equation 2.9).



(a) (b) (c)

Figure 7-3: Different T-maze environments used for the series of experiments discussed in the present chapter. The robot is shown at its initial position, and the arrow indicates the transparent tactile cue.

Environment

The experiments are conducted in the three different T-maze environments shown in Figure 7-3. The tactile cue is a transparent cylinder, and the walls of the maze are sufficiently low so that they do not get in contact with the whiskers of the robot. The whisker sensors can thus only be stimulated by the tactile cue.

7.3.2 Observed Behaviors

During an initial phase, the robot is randomly driven across an empty maze with no tactile cue nor reward signal, thus allowing the neural architecture to capture some basic sensorimotor correlations. After that, the tactile cue and the reward delivery system are added to the environment, and the robot, placed in the central arm of the T-maze, is let to move on its own. The activity of the neuron corresponding to the desired state in the reward modality is kept to a fixed positive level, mimicking the idea that the artificial mouse "wants" to get the reward. After the robot reaches the end of an arm where it receives either a reward or a punishment, it is placed back to the starting location in the central arm for the next trial. During the experiments with the real robot, the position of the cue and the reward is changed after every third trial. The trajectories of the robot are recorded for each trial from an overhead camera and are shown in Figure 7-4. The results are summarized in Figure 7-5, were the outcome ($+1$ for reward, -1 for punishment) is plotted for each consecutive trial.

The observed behavior can be described as follows. In the very first trial, the robot moves straightforward until it reaches the wall of the maze. The robot turns away from the wall and eventually reaches an arbitrary arm of the maze. In the next two trials, where the cue and the reward are kept on the same side, the robot systematically turns into the correct arm, irrespective of whether it succeeded or failed in the first trial. But when the side of

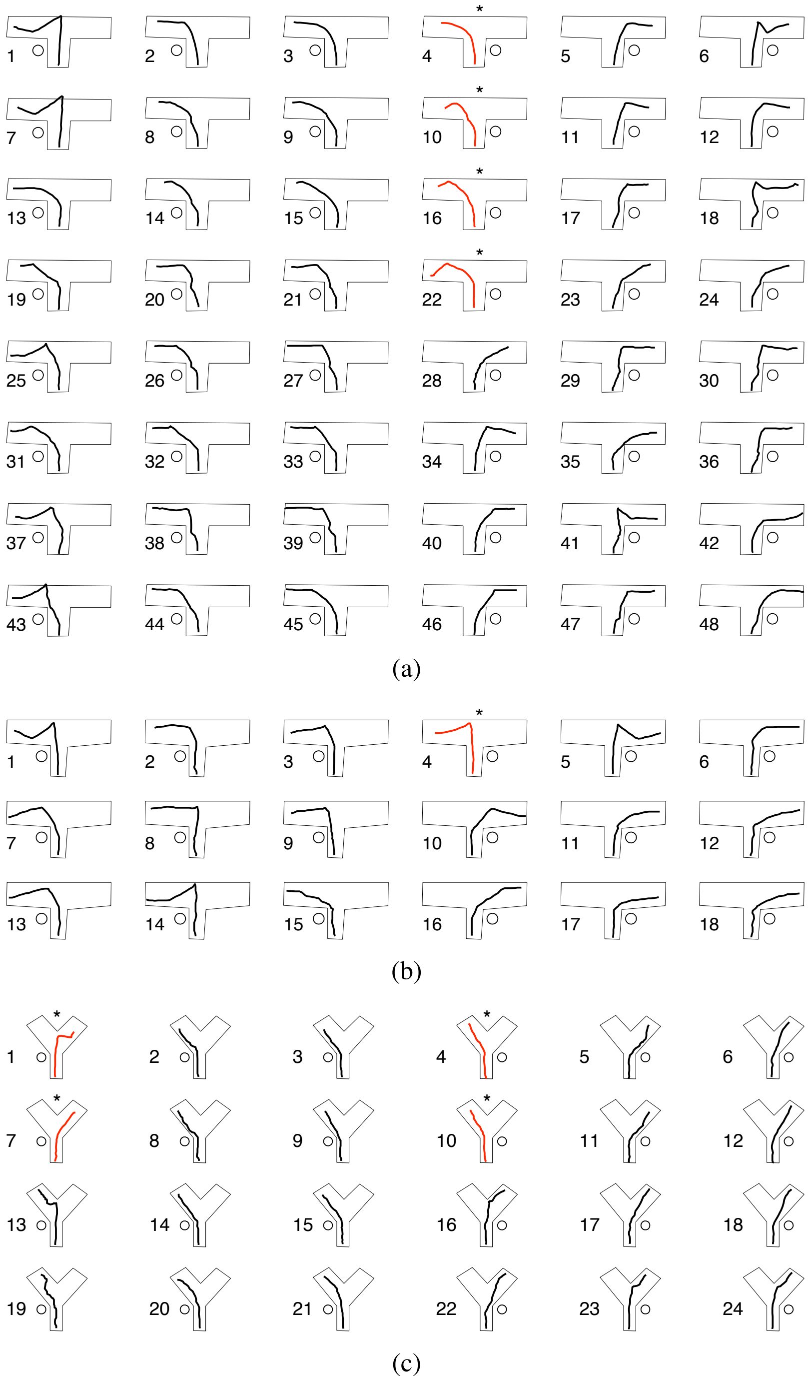


Figure 7-4: Trajectories of the robot recorded from an overhead camera in the first successive trials of three runs performed in each maze configuration shown in Figure 7-3. The position of the tactile cue (circle) is changed after each third trial. The trials where the robot fails to turn in the correct arm are marked with asterisks.

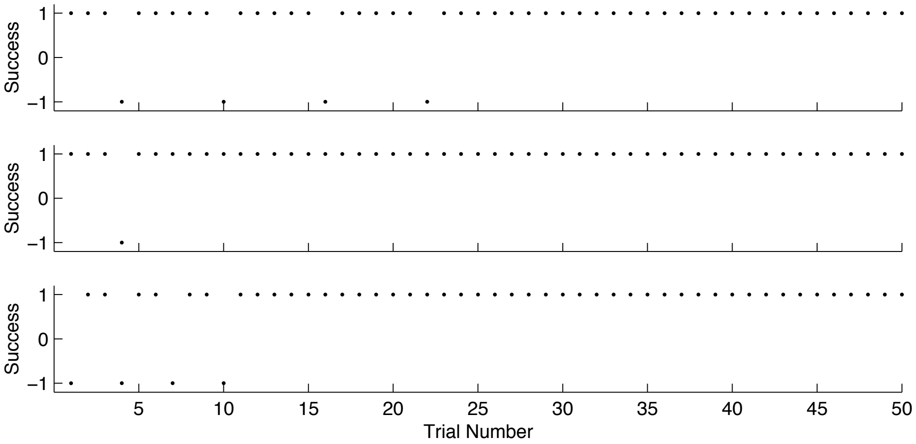


Figure 7-5: Evolution of the robot performance in the T-maze tasks. For each successive trial, a dot indicates whether the agent succeeds ($+1$) or fails (-1) to turn in the arm where the reward is located. The three graphs correspond to the three runs shown in Figure 7-4.

the cue and the reward is changed, the robot typically fails to turn into the correct arm – at least during the first dozens of trials (e.g. trials 4, 7 and 10 in Figure 7-4(c)). Nevertheless, the performance of the robot progressively improves, as it turns less and less often into the wrong arm. Eventually, the robot is observed to "solve" the task, i.e. it consistently turns into the arm of the maze indicated by the tactile cue and thus always reaches the reward.

In summary, the robot seems able to learn that the tactile cue indicates the position of a later reward, and to exploit this information to always reach the reward.

In particular, the robot displays a behavior that changes as the robot repeatedly interacts with its surrounding maze environment. Despite the quantitative differences observed between the three environments, the overall qualitative behavior of the robot can be segmented, from an observer perspective, into the following three strategies.

First, the robot only moves forward and avoids collisions with the walls, thus turning by chance into an arbitrary arm of the maze. This initial strategy is observed during the first trial.

Then, starting with the second trial, the robot is observed to turn into a particular arm of the maze depending on the outcome of the immediate previous trial: if the robot reaches the reward on one trial, it will turn again in the same direction on the next trial; conversely, if the robot receives a punishment on one trial, it will turn into the opposite arm of the maze on the next trial.

Finally, this second strategy is progressively replaced by a third and effective strategy where the robot systematically turns into the arm of the maze indicated by the tactile cue and thus always reaches the reward.

7.3.3 Underlying Mechanisms

In this subsection, we analyze the mechanisms underlying the observed behavior described so far. We show that the robot actually does not reinforce its behavior by remembering its decisions or other events until it receives a reward or a punishment. Rather, the final behavior of the robot is produced by a series of reactions, catalyzed by multi-modal correlations with task-neutral sensory stimuli, between internal neural structures and the external interaction of the agent with its environment.

Absence of working memory

One of the first and most obvious questions to ask, since the robot is observed to learn the delayed reward task in the T-maze, is the following (cf. Section 7.2): how does the robot remember the side of the tactile cue and its decision – i.e. whether it turned left or right – until it receives a reward or a punishment? In other words, where is the information about the events occurring during a trial stored inside the neural architecture of the robot?

The answer sounds perhaps as simple as it is surprising: nowhere. At the end of a trial, the robot does not remember its decision, nor the position of the tacile cue. The neural network does not store such information. To show this, let us consider the evolution of the internal state of the robot during the very first trial.

The robot starts by moving straight along the central arm of the maze. When the robot reaches the junction, the whisker sensor starts detecting the tactile cue on one side: the Hebbian learning mechanism thus captures a correlation between the increase of tactile stimulation on one side and the current state of all other modalities. However, as soon as the robot moves past the cue, the tactile signal decreases back to its original value, thus canceling out all previously learned correlations. This shows that the internal state at the end of the first trial does not depend in any way on the position of the tactile cue.

Finally, when the agent receives the reward or the punishment at the end of the trial, the neural network captures a correlation between change in reward signal and the current visual input (there is no activity in the current states of the other modalities). One could argue that indirectly, the information about the turn the robot has taken previously is contained in such correlations with the visual input[4]: for instance, the fact that a wall of a particular color is detected on one side of the visual field implies that the robot has turned in a particular direction before. This is however a frame-of-reference issue: there is no way to deduce in which direction the robot turned without knowing where the colored wall

---

[4] The same argument holds for the correlations captured by the network when the robot turns to avoid collision with the boundary of the maze.

is in the environment – something that the robot obviously does not know. This demonstrates that the robot does not keep any trace of information about the cue, nor any explicit information about the action taken before the reward or punishment is received.

On the role of task-neutral sensory stimuli

The environment stays visually the same throughout the whole experiment. Indeed, the only elements of the environment that change are the reward and the punishment, which are clearly invisible to the robot, and the position of the tactile cue, which is a transparent object that the camera of the robot cannot detect (see Figure 7-3). The sensory input from the camera thus does not provide any information about the location of the reward nor about the cue: vision is a task-neutral sensory modality.

One might therefore intuitively think that the camera, not providing any useful information to the robot about the task, could just as well be discarded. The robot, deprived of its camera, would indeed display no different behavior during the first trial of a run: it would still move straightforward and avoid collisions with the walls of the maze.

However, the behavior of this sightless robot would not change anymore during the subsequent trials (the reason for this will soon become clear), and the robot would keep turning by chance in a random arm of the maze, irrespective of the possible outcome. In other words, the robot cannot learn to solve the T-maze task without its camera – even though this sensory modality provides no information whatsoever about the task.

The reason why the visual modality, despite its a priori irrelevance, plays a critical role in enabling the robot to solve the task is the following. As the robot interacts with the T-maze, the neural network will capture some momentary cross-modal correlations with the visual stimuli, which temporarily modify the behavior of the robot. This biased behavior will in turn generate, because of the particular way the robot then interacts with its environment, some new sensorimotor correlations that will further modify the observed behavior. In short, activity in the visual modality acts as a catalyst that initiates a series of internal and behavioral changes – described in the next subsection – that eventually lead to the observed learning of the task.

A behavioral chain reaction

The environment is visually not uniform. As shown in Figure 7-3, one side of the arena is darker than the rest: either it is a color-painted wall, or it contains various random objects. Let us consider the environment with a color-painted wall shown in Figure 7-3(a) in two successive trials where the reward is placed twice on the same side, say without loss of generality on the left.

If, by chance, the robot succeeds in the first trial (i.e. if the robot turns into the left arm and reaches the reward), the neural network will momentarily learn a correlation between an increase of reward signal and the image of the colored wall on the right part of the visual field[5]. Therefore, the neural activity generated in the reward modality will then propagate into the visual modality, projecting in the desired population the image of a colored wall on the right of the visual field. When, in the beginning of the next trial, the robot faces the colored wall, the difference in the visual modality between current state (colored wall in the center of the visual field) and desired state (colored wall on the right of the visual field) will elicit a visual flow to the right. As a result, the activity will propagate further into the motor modality and generate a turn of the robot to the left. The propagation of neural activity across the different modalities described in this paragraph is illustrated graphically in Figure 7-6. In fact, the reverse situation holds too: the robot will also turn to the left in the second trial if the robot *fails* in the first trial, as shown in Figure 7-7.

This seemingly goal-directed behavior is similar to the behaviors observed in the previous chapters: the robot is observed to move towards a position where it perceives again the visual stimulus associated to the reward signal – in this case, the view of a the colored wall on its right.

This shows how the static sensory input (the image of the environment) – which is entirely neutral with respect to the task – gives rise to a behavior that actually increases the probability for the robot to turn into the correct arm when the reward is placed consecutively on the same side. This explains the behavior observed in the trials following the very first trial of a run, where the robot only fails during each third trial, i.e. when the side of the reward is changed.

The reason why the robot eventually learns to always turn in the correct arm is that it eventually captures a sensorimotor correlation (between tactile and motor activity), which is initially non-existent, and which progressively surfaces from the biased behavior described so far. Since the robot turns more often to the left when the tactile cue is perceived on the left, and vice versa, a sensorimotor correlation is produced between neural activity in the tactile and motor modalities. Change of tactile stimulation on one side gets with time correlated to motor activity corresponding to a turning motion on the same side (see Figure 7-8 (a) and (b)). As a consequence, when the robot perceives the tactile cue on one side, the neural activity generated in the tactile modality directly propagates into the

---

[5] This correlation is indeed only momentary since over several trials, reward or punishment signals occur as often with the colored wall on the left part than on the right part of the visual field. We also assume Hebbian learning with a positive forgetting rate $\varepsilon > 0$, see Equation 2.9.

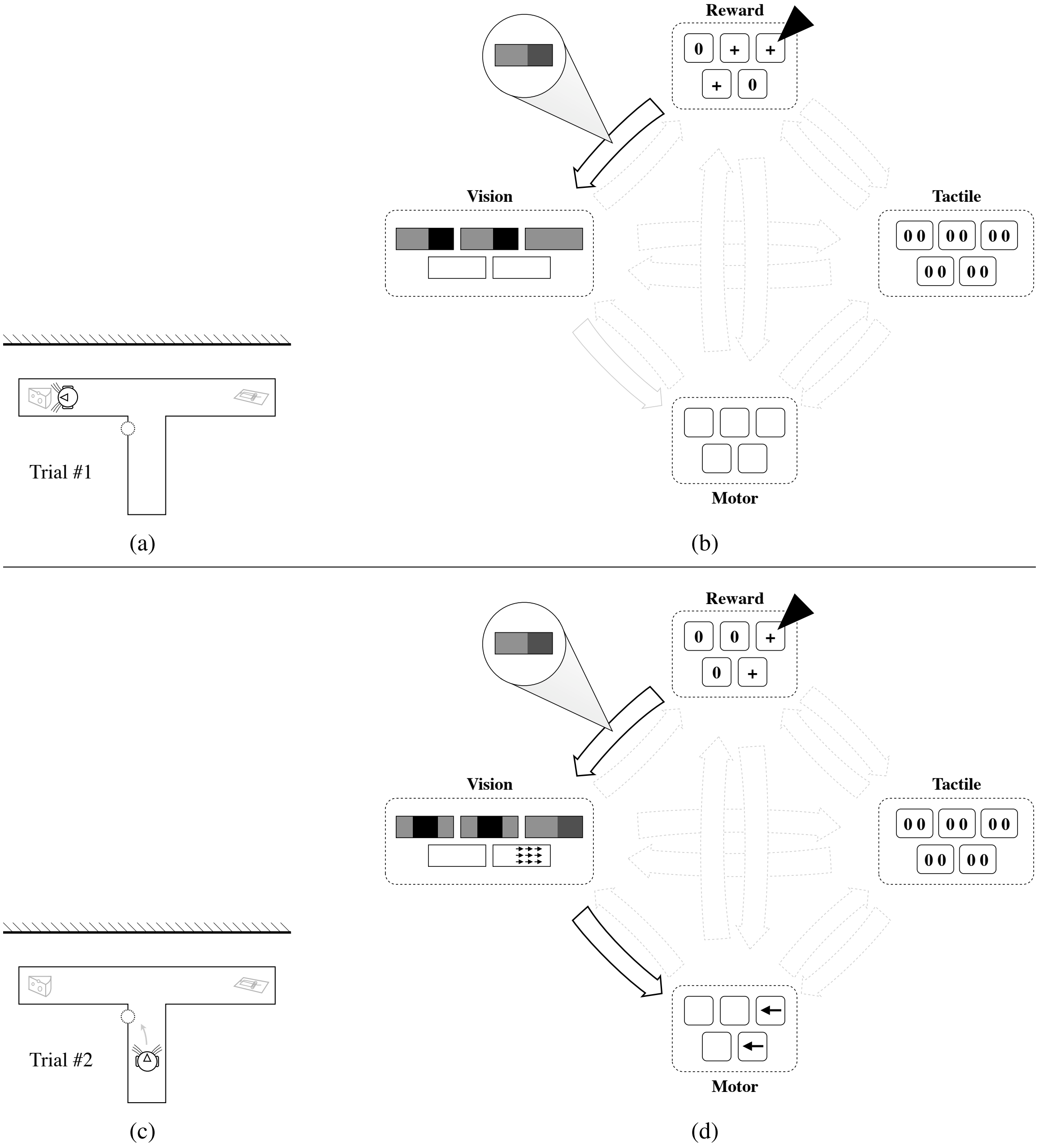


Figure 7-6: Graphical illustration of the internal network dynamics during the first two trials of a run. The robot succeeds in the first trial (a). The neural network thus learns a temporary correlation between increase of reward signal and the image of the wall on the right part of its visual field (b). During the next trial, activity generated in the reward modality will therefore propagate into the visual modality. When the agent faces the wall (c), the difference between the current and desired visual states elicit a visual flow to the right (d), and the neural activity propagates into the motor modality, making the robot turn to the left, i.e. into the arm where the reward was previously received.

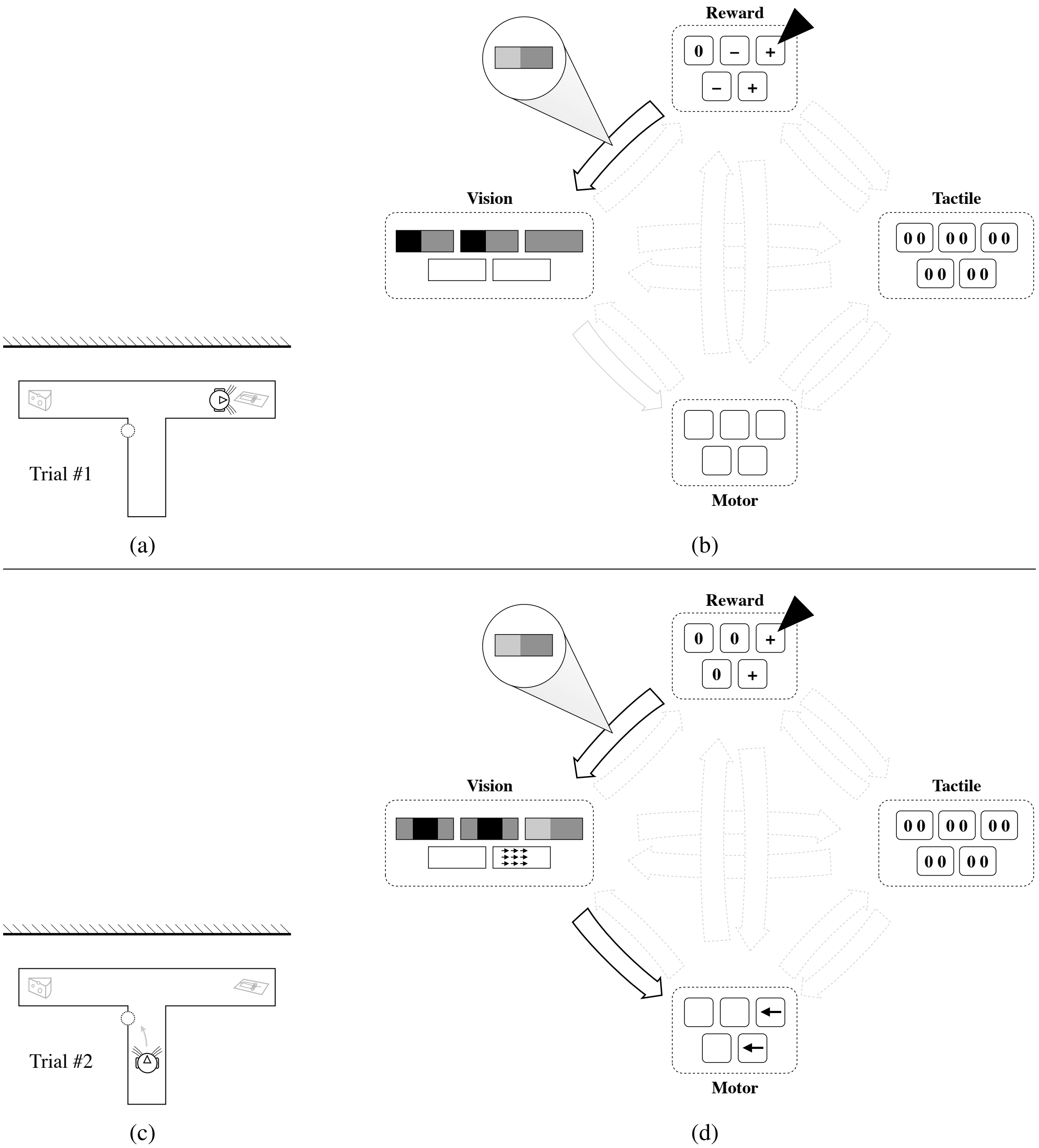


Figure 7-7: If the robot fails in the first trial (a), the neural network will learn a temporary correlation between decrease of reward signal and the image of the wall on the left part of its visual field (b). When the agent faces the wall on the next trial (c), neural activity generated in the reward modality propagates, as in the case illustrated in Figure 7-6, first into the visual modality (with opposite sign and opposite side of the visual field) and then further into the motor modality (d), generating again a turning motion of the robot to the left, i.e. into the arm opposite to the one where the punishment was previously received.

motor modality, inducing a turning motion of the robot in the appropriate direction (see Figure 7-8(c)).

Note that one behavior – namely turning again into the arm where the reward was received, irrespective of the tactile cue – results from the *temporary* correlation captured between reward signal and visual input. In contrast, the other behavior – namely turning on the side where the tactile cue is perceived (and therefore into the correct arm) – results from an initially non-existent correlation of tactile and motor activity that gradually increases with each trial. During all trials, the turning motion of the robot is induced by both cross-modal associations. However, the latter behavior, initially imperceptible, starts to become dominant after a certain number of trials, and the robot is then observed to always turn into the correct arm of the maze. This gradual shift between two competing behaviors in a real and thus noisy environment also explains why the robot is sometimes observed to learn faster with the cue on one side, and slower with the cue on the other side of the junction (see e.g. Figure 7-4(a)).

Parallel pathways for the propagation of neural activity

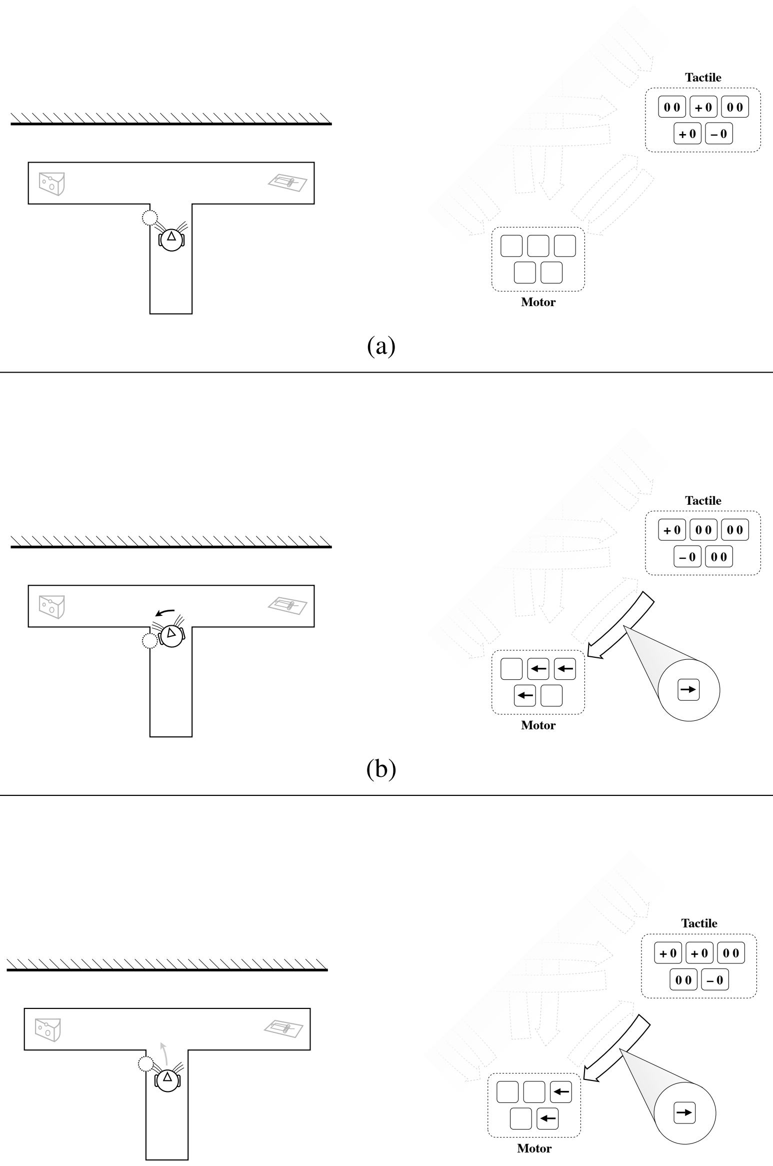
What about the remaining cross-modal correlations? Do they interfere with the phenomena described so far? Interestingly, it turns out that the only significant cross-modal correlation that has not been mentioned so far actually enhances the second observed behavior where the robot turns into the correct direction.

Figure 7-9 (a) and (b) shows the evolution of the synaptic weights connecting the tactile modality to the visual modality during a trial where the tactile cue is perceived on the left. Change of tactile input on the left gets correlated to a pattern of visual activity corresponding the contour of the colored wall shifted to the right.

As a consequence, when the robot detects a tactile cue, activity propagates from the tactile modality into the motor modality simultaneously along two different pathways – first directly into the motor modality, and second through the visual modality – both inducing a turning motion of the robot in the appropriate direction (see Figure 7-9(c)).

7.3.4 Simulation Experiments

To assess the robustness of the model, further experiments are performed in simulation under various conditions. First, the probably to reassign the side of the reward after each trial – i.e. the average frequency at which the side of the reward is changed – is varied between 0% (the reward always remains on the same side) and 100% (the reward is randomly reassigned after each trial). Second, the experiment is repeated with the reversed task condition,

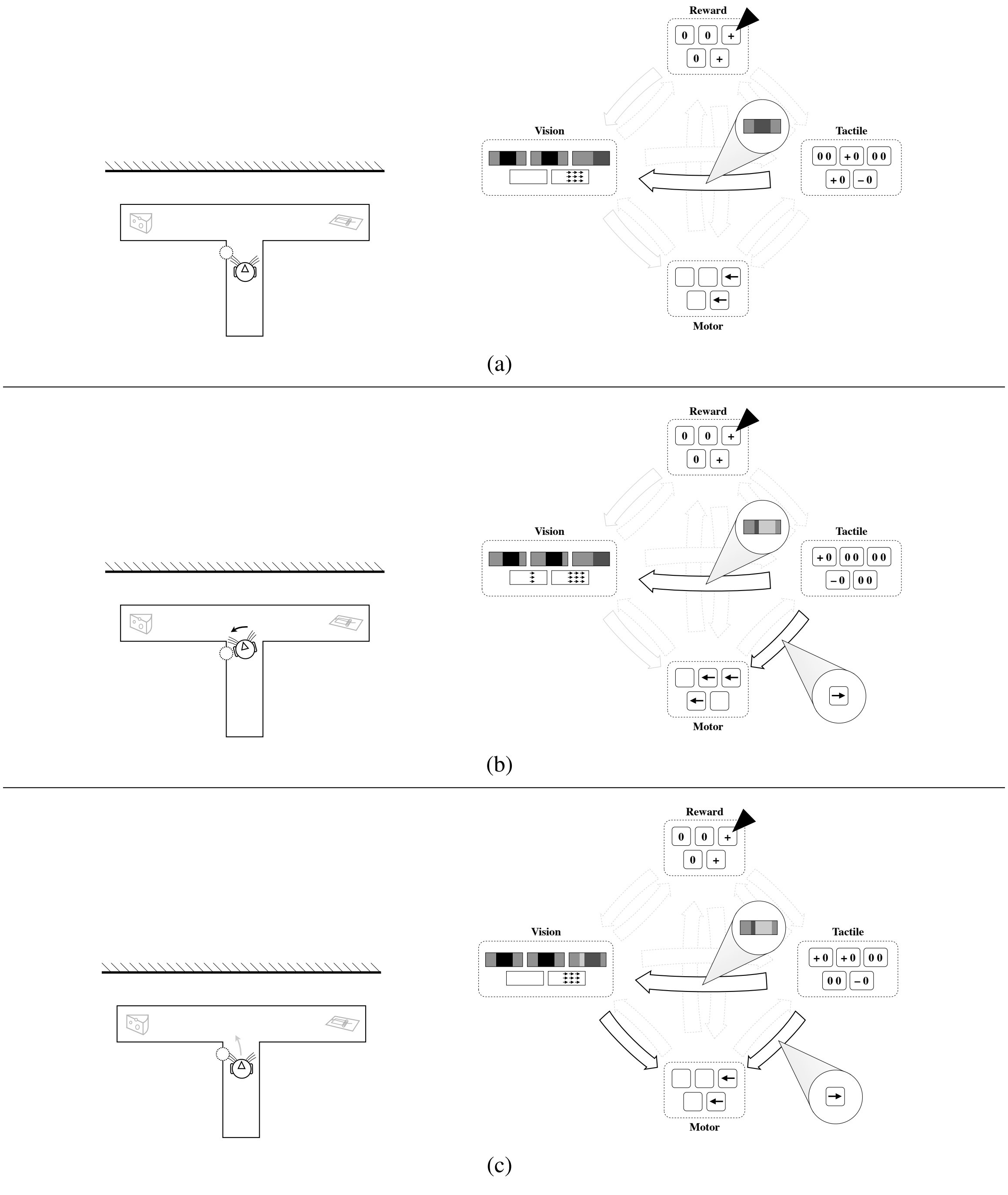


(a)

(b)

(c)

Figure 7-8: Sensorimotor correlation created when the robot turns at the junction into the correct arm of the maze. This figure only illustrates the tactile and motor modalities of the neural network. (a) The whiskers on one side get in contact with the tactile cue, creating an increase of tactile sensory activity. (b) As the robot further advances, the contact between whiskers and tactile cue is lost. Since the robot is then turning, the network captures a correlation between decrease of tactile sensory input and motor activity. (c) During subsequent trials, when the whiskers detect again the tactile cue on the same side, activity generated in the tactile modality propagates into the motor modality, creating a turning behavior of the robot to the side of the cue.

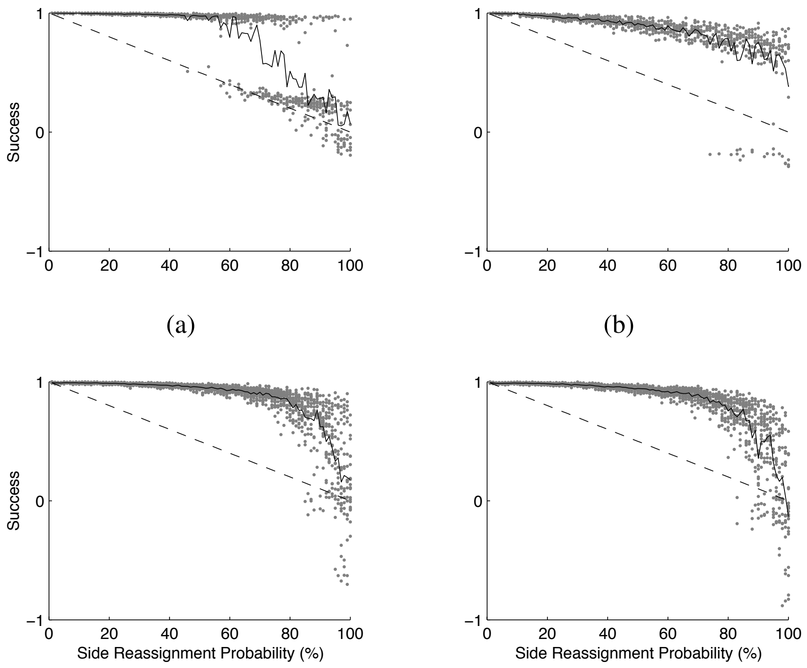


(a)

(b)

(c)

Figure 7-9: This figure illustrates the same conditions as Figure 7-8, but shows all modalities of the neural network. (a) When the whiskers on one side get in contact with the tactile cue, the network learns a correlation between increase of tactile sensory activity and visual input. (b) As the robot passes by the tactile cue, the network learns not only a correlation between decrease of tactile input and motor activity (already shown in Figure 7-8), but also further modify the synaptic weights between the tactile and visual modalities. (c) When the whiskers detect again the tactile cue, activity generated in the tactile modality propagates across two pathways – directly into the motor modality, and through the visual modality – both inducing motor activity corresponding to a turning motion on the side of the cue.



(a) (b)

(c) (d)

Figure 7-10: Simulation results. Each run is indicated by a dot, the abscissa indicating the probability at which the reward is reassigned after each trial, the ordinate giving the success rate averaged over 1000 successive trials. The camera vision used in (a) and (b) is replaced with temperature sensors in (c) and (d). The cue is placed on the same side as the reward in (a) and (c), on the opposite side in (b) and (d). The solid curves indicate the mean success rate. The dotted lines indicate the maximum possible success rate for strategies ignoring the tactile cue.

namely where the tactile cue is placed on the side of the junction *opposite* the arm of the maze with the reward. Finally, the vision system of the robot is replaced with temperature sensors placed on the front of the agent. We then assume a uniform temperature gradient across the environment, so that the temperature is in one arm of the maze higher than in the opposite arm[6].

The results are shown in Figure 7-10, and can be summarized as follows. First, the robot is always able to solve the task in the long run as long as the position of the reward is not changed too often (probability of side reassignment less than approx. 70 to 80%). As the probability increases towards 100%, the average performance drops steadily towards the maximal performance level achieved by a strategy ignoring the tactile cue. This means that the robot needs ever more initial trials until its network captures enough sensorimotor correlation between tactile and motor activity to produce the correct turning strategy.

---

[6] The temperature sensors are calibrated to read positive values in one arm of the maze, and negative values in the opposite arm.

Second, the robot is observed to solve both variants of the task, where the tactile cue is consistently placed either on the same side as the reward, or on the side opposite to the reward. The quantitative difference in the performance observed when the side of the reward is frequently reassigned (as seen for instance by comparing graphs (a) and (b) of Figure 7-10) is due to the particular morphology of the tactile system of the agent and its physical interaction with the tactile cue. Indeed, the whiskers of the agent stays longer in contact with the tactile cue when the robot turns on the side of the cue, as compared to when the robot turns away from the cue. The likelihood that the robot is not turning anymore (or even turning is the other direction to avoid collision with the maze) during the *offset* of tactile stimulation is thus increased. In this case, the neural network captures less rapidly the appropriate sensorimotor correlation between tactile and motor activity. This explains the reduced performance in the case where the cue is placed on the side where the robot has to turn to reach the reward.

Finally, the results show that the robot is also able to solve the maze when the visual modality is replaced with another task neutral sensory modality – in this case with sensors detecting a temperature gradient. The key aspect is that both task neutral sensory modalities allow the network to capture some temporary correlation with the outcome of one trial – reward with a colored wall on one side of the visual field, or reward with high or low temperature – to trigger the chain of behavioral changes described previously that eventually leads to the correct turning strategy.

7.3.5 Collision Avoidance

This section describes the role of the infrared proximity modality in the neural architecture, and how this sensory modality produces the collision avoidance behavior observed when the robot moves around in the T-maze environment.

Collisions with the walls of the maze constitue an important issue as they prevent the AMouse robot from moving any further – unless of course the robot actively turns away. This problem could obviously be solved with ad hoc solutions, such as providing the robot with a built-in avoidance reflex or, in simulation, neglecting the friction and thus allowing the robot to slide against the walls of the maze.

Surprisingly, it turns out that this problem vanishes by itself if the infrared proximity sensors, that the robot is anyway equipped with, are incorporated into the neural architecture. In fact, including this infrared sensory modality into the neural network does not disrupt or otherwise interfere with the neural dynamics or with the observed behaviors discussed so far – this is the reason why, for the sake of simplicity, the infrared modality has deliberately not been mentioned in the discussion so far. Indeed, the only significant cross-modal correlation that the neural network captures with the infrared modality is the one, described in the following paragraph, which gives rise to the collision avoidance behavior.

During the initial phase of the experiment, where the robot is randomly driven across the maze, the synaptic weights connecting the proximity modality to the motor modality learns the following correlation. A motor state corresponding to a turn on the left will be correlated both to an increase of proximity input on the left side (whenever the robot approaches and bumps into a wall), to a decrease of proximity input on the right side (whenever the robot turns away from a wall), and vice versa. Consequently, when the robot approaches a wall, activity generated in the infrared proximity modality will propagate into the motor modality, leading to the observed collision avoidance reflex (see Figure 7-11).

In summary, it is interesting to observe that the robot, equipped with proximity sensors, will spontaneously avoid collision with the walls of the maze. Again, this behavior is not explicitly programmed into the system, but is produced by the interaction between the agent, the neural architecture and the environment.

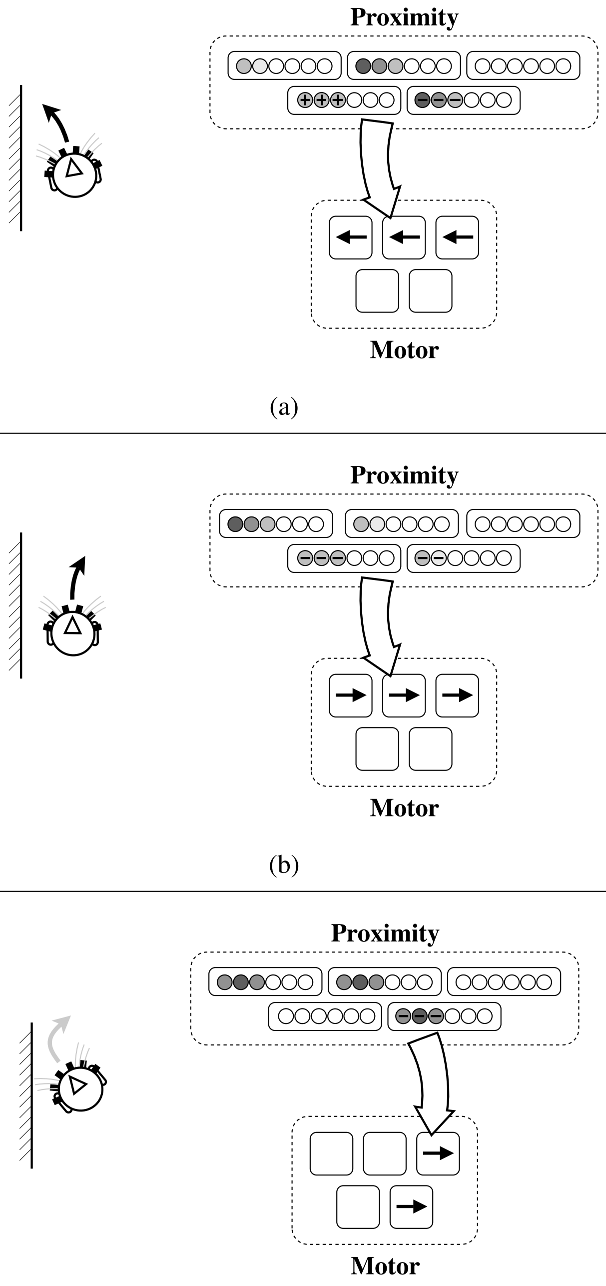
7.4 Discussion

The series of experiments described in this chapter provides an additional case study illustrating how the homogeneous neural architecture proposed in this thesis – i.e. a neural system not specifically designed for any particular task – can lead to surprisingly coherent and seemingly goal-directed behaviors. Indeed, we described how the AMouse robot, engaged in a T-maze task, could learn to rely on a tactile cue indicating the side of the reward to always turn into the correct arm of the maze.

The obtained results are interesting in many respects, which are discussed in the following paragraphs. In particular, they demonstrate how the approach put forward in this thesis can offer a fruitful level of abstraction, enabling the investigation of an artificial agent, engaged in a maze environment used to study animal behavior, to shed new light on the neural basis of learning and memory.

7.4.1 Adaptive Behavior

The neural architecture of the robot is the same as the one used in all other chapters of the thesis. The neural network has thus clearly no particular structure related to the task. In other words, the robot does not “know” anything about the environment and the task, nor that a specific sensory signal (namely tactile stimulation) indicates what action has to be



(a)

(b)

(c)

Figure 7-11: Wall collision avoidance with the robot equipped with infrared proximity sensors. (a) When the robot turns towards a wall – in this example on its left – the neural network learns a correlation between increase of proximity input in the left sensors and motor activity corresponding to turning motion to the left. (b) Conversely, decrease of proximity input in the left sensors is correlated to motor activity corresponding to turning motion to the right. (c) When the robot gets close to a wall, the difference between the current and desired states in the proximity modality generates neural activity that propagates into the motor modality, generating an avoidance behavior. This figure only illustrates the proximity and motor modalities of the neural architecture.

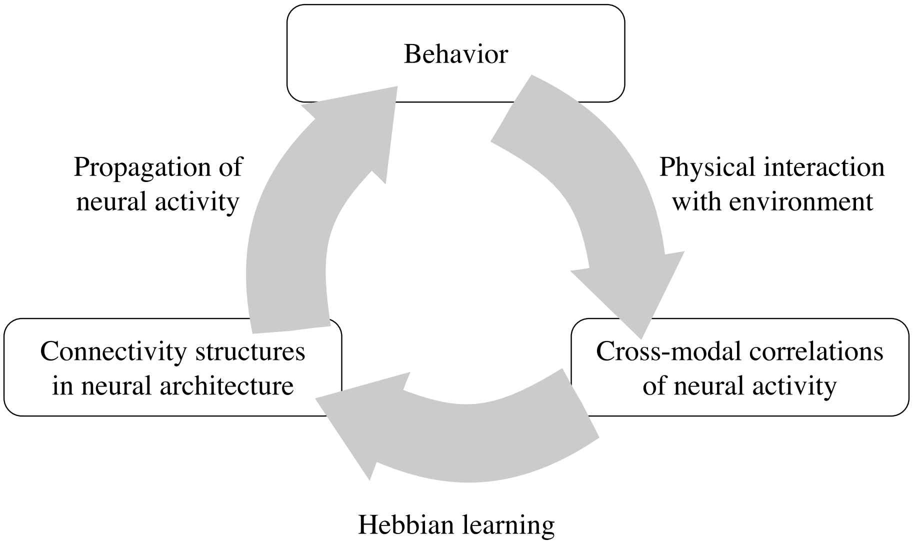


Figure 7-12: Adaptive behavior seen as a continuous and reciprocal interplay between the internal neural dynamics and the physical interaction of the agent with its environment.

taken to get the reward. Actually, the robot isn't even told or instructed in any way what is has to learn: the reward or punishment signals, which look from an observer perspective as value signals that the robot could exploit to reinforce its behavior, are in fact just another sensory signal treated in the neural architecture as any other sensory or motor signals. Nevertheless, the robot is progressively observed to behave as if it would learn different strategies – relying initially on the outcome of the immediate previous trial, and eventually on the tactile cue – to consistently turn into the arm of the maze containing the reward.

The reason why the behavior of the robot is observed to qualitatively change with time – i.e. why the robot learns different strategies to improve its performance – is the continuous and reciprocal interplay between the internal neural dynamics and the physical interaction of the robot with its environment (see Figure 7-12). The interaction of the robot with the environment generates cross-modal correlations of neural activity. These correlations, captured by the Hebbian plasticity of the synapses, modify the connectivity structure of the neural network and thus give rise to qualitatively new behavior – i.e. in turn a different interaction of the robot with the environment.

The chain of reactions producing the observed behavioral adaptivity, which we analyzed in details in the previous section, can be summarized as follows. The initial sensorimotor correlations provide a preliminary structure to the network that allows the robot to freely move in the T-maze without getting stuck against a wall. Then, during the early trials, the network captures some temporary cross-sensory correlations, which momentarily bias the behavior of the robot from one trial to the next. In turn, this new behavior creates, as the agent interacts with the T-maze environment, a further sensorimotor correlation that, once captured by the neural network, produces the final observed behavior.

In summary, the results obtained in this chapter illustrate how adaptive behavior can be conceptualized as a self-organizing process with no explicit value system, and involving the interaction between the agent's body, neural system and environment. In particular, the reason why the observed behavior qualitatively changed with time – in contrast to the behaviors observed in the previous chapters of this thesis – is not a more complex neural mechanism, but rather a more dynamic environment.

7.4.2 Increased Complexity

This chapter also illustrates the potential of the proposed neural architecture to scale up with increasing complexity of the system. Not only is the robot equipped with more sensory modalities, but also the environment is more dynamic than in the previous experiments. Nevertheless, the results show that the neural architecture is still able to produce a coherent and adaptive observable behavior.

On the one hand, the results show that – at least to a certain extent – more modalities can be included in the network without perturbing its intrinsic dynamics. Indeed, the AMouse robot is equipped in this chapter with five different modalities (motor, tactile, visual, reward and proximity). However, out of the twenty cross-modal synaptic couplings, only four are shown to learn a significant correlation of neural activity, and just one to capture some temporary correlation. This shows that even though all modalities are coupled together, the inherent structures in the patterns of sensorimotor activity – generated through interaction with the environment – automatically constrain neural activity to only propagate along a sparse subset of synaptic connections.

The investigation of the internal connectivity of the network even reveals some interesting characteristics arising from the increased number of modalities. Section 7.3.5 described how activity from the infrared proximity sensors produces a collision avoidance behavior – a necessary behavior that would otherwise require some ad hoc mechanisms – without disrupting the remaining dynamics of the network.

Moreover, we observed that some neural activity propagates along multiple parallel pathways in the neural architecture, thus providing a certain redundancy to the neural system. More specifically, we showed that the turning behavior of the robot in the presence of the tactile cue is triggered by two sets of associations (cf. Figure 7-9). Neural activity generated by the cue in the tactile modality propagates both directly into the motor modality as well as through a cascade of synaptic associations between the tactile, visual and motor modalities. An immediate benefit of this redundancy is an increased robustness of the system: the robot will continue to turn into the correct arm of the maze even if the synaptic coupling between tactile and motor modalities is disrupted, or if the vision system of the robot breaks down.

On the other hand, the complexity of the system is increased by making the environment more dynamic: during the experiment, the side of the tactile cue and the reward is regularly changed. The repeated interaction of the robot with this non-static T-maze environment produces thus a richer interaction between internal and external dynamics. In particular, the fact that the neural dynamics is continuously evolving – because of the Hebbian plasticity – allows a chain reaction to take place between neural and physical dynamics that leads, as described previously, to the observed adaptivity in the behavior of the robot.

The results thus suggest that more complex behaviors – such as the capability of the robot to learn and exploit the relation between an early tactile cue and a delayed reward, i.e. something that some might refer to as "higher-level cognition" – do not necessarily require more complex internal structures or processes, but can simply be produced by a richer interaction between the agent, the neural system and the environment.

7.4.3 Delayed Reward Learning with No Working Memory

The study presented in this chapter demonstrate how an embodied agent, whose neural architecture only captures *simultaneous* correlations (i.e. who only learns the relation between pairs sensory or motor events happening at the same time), is nevertheless observed to behave as if would learn the temporal relationship between an early tactile cue and a delayed reward.

The results allow us to reconsider a widely accepted assumption about delayed reward learning. Delayed reward learning refers to situations in which the subject – an animal or robot – has to make a particular decision, e.g. whether to turn left or right in a maze, but the feedback whether the decision was right or wrong is only provided later, when the reward or the punishment is given. The assumption is that in order to solve the problem, the subject must somehow remember the sequence of sensory or motor events until the reward or punishment is given. For instance, if the subject receives a reward after having perceived the cue on the right (previous state) and after having turned to the right (previous action), a reasonable strategy is to increase the probability of turning again to the right next time the cue is perceived on the right.

This assumption is so intuitive that putting it into question seems almost pointless: how can the animal or the robot learn anything if it does not remember what happened before the reward is received? Indeed, algorithms developed in the field of reinforcement learning to solve this kind of problem all assume that the learning agent has the ability to record a sequence of observed states and actions taken until the reward is given (Connell and Mahadevan, 1993; Kaelbling et al., 1996; Sutton and Barto, 1998; Doya, 2000; Doya et al., 2001). Similarly, the ability to *retain* information to guide behavior, referred in neurobiology to as "working memory" (Baddeley, 1986; Pasternak and Greenlee, 2005), is believed to be the neural basis of an animal's performance in a delayed reward task (Goldman-Rakic, 1995). However, this assumption about a memory consisting of stored information – as intuitive as it may sound – probably raises more issues than it solves (see discussion in Section 7.5). How does the situated agent knows what information needs to be stored? How is the information kept internally? How and when is the information retrieved back? And how is the retained information eventually processed?

In contrast, the present chapter demonstrates how an embodied agent can solve a delayed reward task in a T-maze environment *without* retaining any information about where the cue is perceived or which turn is taken until the reward or punishment is received.

7.4.4 On the Importance of Task-Neutral Stimuli

Section 7.3.3 analyzed the mechanisms underlying the observed increasing performance of the robot. In particular, it showed that the learning behavior of the robot actually emerges from a chain reaction between the neural and the physical dynamics of the system. Interestingly, this chain reaction was shown to be catalyzed by some task-neutral sensory stimuli, i.e. sensory activity produced by structures of the environment – such as the surrounding arena being painted on one side in a different color – that provide no information whatsoever about the location of the reward.

In fact, the learning process can even be catalyzed by task-neutral sensory stimuli that are only temporary. For instance, once the robot has learned to turn into the correct arm of the maze using the tactile cue, the colored wall or the temperature gradient can be removed from the environment without disrupting the observed behavior.

This demonstrates that the behavior of an embodied agent dynamically interacting with its environment can possibly exploit any sensory modalities, and might therefore critically depend on a priori completely irrelevant aspects of the environment. This suggests some potentially profound consequences concerning the study of natural behavior. When experiments are performed to study animal behaviors (for instance, how a rat can learn to efficiently find food in a maze by relying on tactile cues), great care is taken to remove significant aspects of the environment, i.e. aspects that can interfere with the task (e.g. ol-

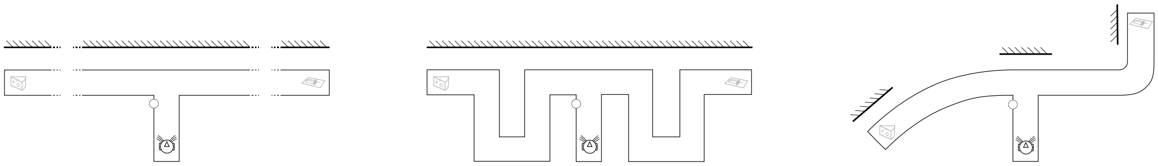


Figure 7-13: Various shapes of mazes in which the robot can learn to solve the delayed reward task.

factory or visual cues that would disclose the food location). However, a priori irrelevant aspects of the environment are typically not taken into account, such as the environment being slightly and temporarily brighter or cooler on one side – something that obviously can never be avoided under real-world conditions. Similarly, the aspects of the internal neural system that neurophysiology typically focuses on are *significant* patterns of activity, i.e. patterns of neural activity that are correlated in some way or other with the task or the behavior being studied. Yet, the present chapter demonstrates that ignoring task-neutral aspects of the environment (such as the colored wall) or a priori insignificant sensory stimuli (such as visual activity) can preclude the understanding of the behavior observed with a situated, embodied agent.

7.4.5 Mazes the Agent Can Solve

What kind of mazes is then the robot able to solve? By understanding the various processes, analyzed in Section 7.3.3, that lead to the observed behavior, it is possible to extract and summarize the two key features of the mazes in which the robot is able to solve the task. First, the task-neutral structure of the environment perceived when the reward or punishment is received must be so that, next time the robot reaches the junction, the robot is biased to turn into the arm where the reward was located. And second, the tactile cue must be placed at the junction of the maze.

This means for instance that there is no constraint whatsoever about the shape of the maze between the junction and the location of the reward. The arms of the T-maze can thus be infinitely elongated or arbitrarily bent, as shown in Figure 7-13.

Also, the condition for the robot to be able to solve the task is not so much a limited complexity in the shape of the maze, but rather an appropriate arrangement of the task-neutral structures in the environment. Figure 7-14 illustrates different environments where, depending on how the task-neutral wall is perceived when the agent receives the reward or the punishment, the agent is able or not to solve the delayed reward task.

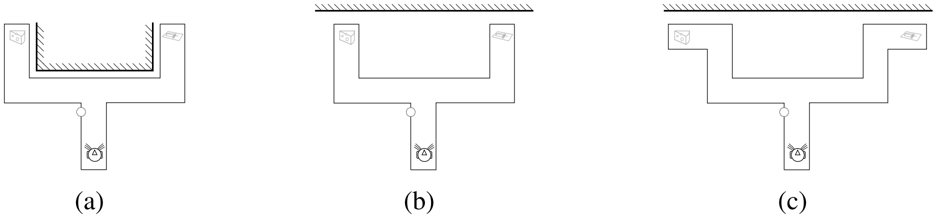

(a) (b) (c)

Figure 7-14: Depending on how the wall is perceived when reaching the reward of the punishment, the agent is either able to solve the delayed reward task (a, c) or not (b).

7.5 Memory

One of the chief contributions of the study presented in this chapter is the new light shed on the neural basis of memory. The obtained results not only demonstrate that memory is not necessarily information stored inside the agent, but also show that memory can be at least partly off-loaded into task-neutral structures of the environment. Consequently, before concluding the chapter, this last section discusses how the results gained from the T-maze experiments are related to major concepts of memory found in both classical and modern scientific literature.

Because of the huge body of literature on this topic, it is of course impossible here to cover all aspects and to do justice to the richness of the field. An excellent overview of the traditional and more recent ideas concerning memory can be found in Pfeifer and Scheier, 1999b, chapter 15 or in Pfeifer and Bongard, 2007, chapter 10.

7.5.1 The Storehouse Metaphor

The classical notion of memory – and certainly still the most prominent conception – views memory as the ability to store, retain, and subsequently retrieve information (e.g. Atkinson and Shiffrin, 1968; Ashcraft, 1994; Baddeley, 1998). This view is captured by the "storehouse metaphor" (Koriat and Goldsmith, 1996), and comes so naturally that it is hard to see what other possibility there could be. There are indeed several reasons why this notion of memory as storage is so popular. It can easily be described, communicated or represented graphically. Also, analogies (either explicit or implicit) with the memory of computers – nowadays so ubiquitous in our everyday environments – are particularly easy to understand.

There are however several problems with this view of memory (Bartlett, 1932; Rosenfield, 1992; Clancey, 1997). What information (or events) needs to be stored? Consider for instance the solutions to delayed reward learning proposed by reinforcement learning techniques discussed previously. It is assumed that the agent stores previous states (e.g. the side on which the tactile cue was perceived) and previous actions (e.g. the turn which was made), so that the agent can retrieve them when the reward is received, and figure out which action was appropriate or not. But how does the agent initially know that *this* information needs to be stored?

A similar problem is to decide how long stored information must be retained. What happens is the arms of the maze are elongated, or bent several times as illustrated in Figure 7-13? To solve the maze, would the agent need to store and retain *all* turns made before the reward? Clearly, this becomes a critical issue since storage capacity cannot be infinite.

These difficulties are directly related to the homunculus problem (Bursen, 1978; Edelman, 1992). Postulating a memory consisting of stored information implies that someone – the designer, or a homunculus – defines what information needs to be stored, and when it needs to be recalled. In other words, there is some kind of circularity in the argument: a delayed reward task can be learned if the agent stores relevant information – but then, how can the agent learn in the first place what information is actually relevant?

The approach presented in the current chapter provides a promising alternative to the storehouse conception of memory. Indeed, the AMouse robot is able to solve the delayed reward learning task without knowing what sensory signals are relevant, and without any explicit internal representation of relevant events or decisions taken in the past.

7.5.2 Short-Term and Long-Term Memory

Over decades of memory research, different concepts of memory have been proposed. A well-known distinction based on their temporal properties postulates two different kinds of memory called short-term memory (STM) and long-term memory (LTM). The reasons to postulate separate STM and LTM systems come notably from studies with brain-damaged patients: typically, a stroke patient may have a drastically reduced capacity for acquiring new information, but can recall events from his earlier life. There seems to be agreement that in human, the time scale – i.e. the retention period – of short-term memory in on the order of seconds to minutes, whereas long-term memory lasts from minutes to hour to years, up to an entire lifetime. (It is interesting to note that the distinction between short-term and long-term memory is clearly based on the concept of memory as a storage system.)

Of course, it makes perfect sense to distinguish these two kinds of memory at a behavioral level. However, there is clearly a frame-of-reference issue: one must be particularly careful when postulating the existence of different underlying mechanisms, as the ones proposed by e.g. Atkinson and Shiffrin (1968); Ashcraft (1994); Baddeley (1998, 2003).

This point can be illustrated by the study presented in this chapter. On a short time scale, i.e. from one trial to the next, the behavior of the AMouse robot depends on the previous trial: we observed, especially in the first phase of the experiment, that the robot tends to turn again in the same arm after receiving a reward, or to turn in the opposite arm after receiving a punishment. From an observer perspective, it is reasonable to attribute some short-term memory to the robot – its ability to adapt its behavior depending on the outcome of the immediate previous trial, as if the robot would "remember" what happened during the previous trial. On a longer time scale, the robot starts displaying a new behavior: after several trials, it always turns into the correct arm, as indicated by the tactile cue. Once again, it is reasonable to attribute some long-term memory to the robot – its ability to adapt its behavior after several repetitions using the information provided by the tactile cue, as if the robot would then always "remember" that the cue indicates the correct arm.

Nevertheless, this distinction does not imply the existence of two different memory systems! There is obviously only one single mechanism underlying the behaviors of the AMouse robot, namely cross modal synaptic connections modulated by a simple Hebbian-like learning process.

It is interesting to note at this point some striking similarities between the neural network proposed in this thesis and the "network memory" model developed by the brain researcher Joaquín Fuster. Based on physiological and neuroimaging evidence, Fuster (1997) suggests that memories are distributed and self-organizing networks of interconnected neurons that are associated through synaptic connections with Hebbian-like plasticity. Similar to our present argument, he concludes that "evidence from microelectrode and imaging studies is forcing us to re-evaluate the neural basis of short- and long-term memory, and to seriously question their structural separateness" (Fuster, 1997, p. 458), a separation also brought into question by Glenberg (1997).

7.5.3 Alternative Approaches to Memory

We have shown so far that the classical notion of memory as storage is problematic. Let us now turn to some alternatives available for conceptualizing memory. There is indeed a growing number of approaches in memory research departing from the storehouse metaphor and focusing on embodiment and system-environment interaction. However, they are harder to understand, more difficult to describe and communicate, and require probably a different way of thinking.

Memory, according to Ashby, "is not an objective something that a system either does or does not posses; it is a concept that the observer invokes to fill in the gap caused when part of the system is unobservable" (1956, p. 117). Similarly, many other researchers view memory as a set of skills relating perception and action (Broca, 1861; Clancey, 1991b; MacLeod, 1997). Clearly, focus has shifted to what purpose memory serves in an agent interacting with its environment (Glenberg, 1997). Likewise, the ecological approach (Bartlett, 1932; Neisser, 1978) suggests that "memory might not be some location inside the agent but rather manifests itself in global changes in the entire system, leading to different interactions with the environment, interactions that we, as observers, might describe by invoking the memory concept" (Pfeifer and Scheier, 1999b, p. 520).

Interestingly enough, the experiments of this chapter are in line with many of these notions. First, we showed that trying to locate "memory" as an isolated set of elements inside the AMouse robot is doomed to failure. Of course, the robot's history is partly represented in the neural network because through Hebbian learning, the synaptic strengths have been changed. However, the external environment was shown to play as important a role as the synapses in shaping the behavior of the robot – if the colored wall is removed from the environment, the AMouse is not able to solve the delayed task anymore. Therefore, if any "memory" is to be attributed to the robot, it should include the environment as well.

Second, the homogeneous neural network has by definition no dedicated structures. For instance, the same mechanisms that lead to the collision avoidance reflex (see Section 7.3.5) are also involved in the process of learning to solve the task. This agrees with Glenberg's notion that memory is "in service of perception and action in a three-dimensional environment, and that memory is embodied to facilitate interaction with the environment" (Glenberg, 1997, p. 1), as well as with Fuster's claim that "the same cortical systems that serve us to perceive and move in the world serve us to remember it" (Fuster, 1997, p. 451).

In summary, it makes perfect sense to attribute – from an observer perspective – something like memory to the AMouse robot. However, it is important to realize that there is no corresponding representation inside the robot's brain. In other words, the internal mechanisms providing the memory function can only be fully understood by considering the whole dynamic interaction between the agent and the environment.

7.6 Conclusion

This chapter provides an additional case study where coherent, seemingly goal-directed behaviors are observed with a robot endowed with the neural architecture proposed in this thesis. The environment is a T-shaped maze, where a tactile cue, placed at the junction, indicates the arm in which the reward is located. We showed that after several repetitions of the experiment, where the side of the reward is randomly reassigned, the robot learns to solve the delayed reward task, i.e. to turn on each trial into to correct arm.

We also discussed how the robot, instead of possessing an explicit working memory of previous events and actions, exploits its interaction with the environment, taking advantage of environmental structures entirely neutral with respect to the task. Metaphorically speaking, the presented study illustrates how the memory can partly be off-loaded into the environment.

Finally, this work provides an original and concrete study supporting recent concepts in memory research, developed as alternatives to the common but problematic view of memory as stored information.

Chapter 8

Discussion

This concluding chapter summarizes the results obtained throughout the thesis, discusses the general principles that can be drawn from the behaviors observed in the preceding chapters, and considers future perspectives of the proposed approach to the study of intelligence.

8.1 Summary

Chapter 2 introduced a generic neural architecture for robotic agents, which was used throughout the thesis to explore artificial systems endowed with arbitrary, self-organizing dynamics. We described how the sensorimotor signals of the robot are represented by populations of artificial neurons, and how all the modalities are systematically and reciprocally coupled to each other through synaptic connections with Hebbian-like plasticity. With the help of a thought experiment, we illustrated how the interaction between the neural dynamics of the network and the physical dynamics of an embodied agent can spontaneously generate structures in the neural architecture and thus produce coherent, potentially interesting behaviors.

Chapter 3 provided a first series of experiments demonstrating how the proposed neural architecture could be successfully implemented on a real robot to produce various seemingly intentional behaviors. The mobile robot, equipped with a camera and whiskers, was observed – depending on the activity of a particular tactile neuron – to either avoid tactile contact with objects, or to approach and follow any object placed in its visual field until the whiskers got in contact with it.

Chapter 4 presented a series of experiments inspired by studies on insect navigation. The behaviors of the robot, situated in a desert-like environment, were shown to closely match two distinct navigation strategies observed in natural desert ants: path integration and visual landmark homing. In contrast to the models found in the current literature, which all suggest different and specific underlying neural structures, the results obtained show that the two homing strategies can be produced with one single, homogeneous neural architecture.

Inspired by research on cricket phonotaxis and by the theoretical framework of forward models, experiments were conducted in Chapter 5 to highlight the role of the homogeneous connectivity in the neural network. The results showed that synaptic couplings allowing neural activity to flow from sensor modalities to motor modalities could generate reflex-like behaviors, whereas the role of synaptic couplings in the opposite direction could be identified as predictive reafference cancellation, i.e. as forward models. This study therefore suggests that the same neural mechanism could be underlying both reflexes and predictive reafference cancellation observed in biological systems.

A fourth series of experiments with the proposed neural architecture was conducted in Chapter 6, showing that it could successfully be applied on a robot with a more complex body dynamics – a four-legged running robotic dog – to generate coherent behavior. Despite the complex physical dynamics of the robot and the high level of noise in its sensory signals, a coherent and stable self-generated behavior could be observed. Moreover, a closer investigation revealed that the robot “discovered” a strategy to turn that is more robust than the intuitive strategy commonly used to control the rotation of legged robots.

Finally, the potential of the proposed neural architecture – as well as its ability to scale up with more dynamic environments – was further explored in Chapter 7. Inspired by maze experiments used to study navigation and learning in rodents, we investigated the behavior observed with a robot engaged in a T-maze task, where a tactile cue indicated at the junction which arm of the maze contained the reward. After a number of trials where the positions of the cue and the reward were randomly reassigned, the robot was observed to spontaneously develop a strategy to solve the task – i.e. to consistently turn into the correct arm. Moreover, we demonstrated that the robot solved this delayed reward learning task without possessing any explicit working memory of previous events and actions. Rather, the robot was shown to exploit its interaction with the environment – including environmental structures entirely neutral with respect to the task. By illustrating how the memory can be off-loaded into the environment, this chapter provides a case study shedding new light on the neural basis of memory and supporting recent concepts in memory research developed as alternatives to the common but problematic view of memory as stored information.

8.2 Synthesis

As mentioned in the introductory chapter, the contribution of the work in this thesis is two-fold. On the one hand, it supports the approach of embodied AI: by investigating the behaviors of embodied autonomous agents, and by following the synthetic methodology of "understanding by building," the experiments described in this thesis provide concrete case studies illustrating many of the principles developed in this field, which are discussed in the next subsection 8.2.1.

On the other hand, this thesis explores a novel and complementary approach: by investigating the behaviors of artificial embodied systems endowed with an arbitrary, self-developing dynamics, it sheds new light on the possible underlying mechanisms of different natural phenomena, as well as on the understanding of natural and artificial intelligent behaviors. The principles drawn from this second facet of the work will be discussed in the subsection 8.2.2 that follows.

8.2.1 Embodiment

In recent years, the field of embodied AI has witnessed the continuous development of a set of heuristics, or *principles for autonomous agent design* (Brooks, 1991; Pfeifer, 1996; Brooks, 1999; Pfeifer and Scheier, 1999b; Lungarella, 2004; Pfeifer et al., 2005; Pfeifer and Bongard, 2007). These design principles not only provide guidance in actually designing and building systems, but also capture conceptual insights about intelligent (adaptive) behavior, thus providing a theoretical foundation for the understanding of intelligence.

First of all, the importance of **embodiment** for the understanding of behavior can hardly be overemphasized. This concept not only carries the meaning that intelligence needs a body to manifest itself through behavior. It also stresses the fact that natural or artificial behaviors we observe have to be conceived as continuous interactions between agents' brain, body and environment. Quite obviously, the series of experiments described in all chapters of this thesis clearly illustrate this point. They all show how the structures of a developing neural architecture – and thus the behaviors that follow – critically depend: on the morphology of the agent (i.e. its body), including a priori unrelated sensory modalities (Chapters 4 and 7) and material properties (Chapters 3, 6 and 7); on the environment, including irrelevant structures thereof (Chapters 4 and 7); and on the continuous interaction between the agent and the environment (in particular Chapters 5 and 7). Moreover, since all behaviors observed throughout this thesis were produced by robots equipped with the very same neural architecture, they obviously demonstrate how behavioral diversity –

an essential component of intelligent behavior – does not require complex or hierarchical structures in the brain, but can be produced by even a simple neural structure allowing a rich interaction to take place between the agent and its environment.

Furthermore, the way all experiments were conducted (see Chapter 2) illustrates how the behaviors we observed are to be conceived as sensory-motor processes. Structures in the neural architecture and in the observed behaviors were shaped by correlations of sensory-motor activity – correlations produced by (initially random) motor activity. In other words, the agents did not passively absorb sensory stimulation, but actively shaped their sensory experience. This point was particularly well illustrated in Chapter 7, which showed how the final behavior of the robot was produced by sensorimotor regularities produced by the robot itself while actively interacting in the T-maze environment. This concept is closely related to the **principle of sensory-motor coordination**, supposed to be of crucial importance for the development of perception and cognition (Pfeifer and Scheier, 1999a; Thelen and Smith, 1994; Lungarella, 2004). For instance, it has been shown that an agent, by actively interacting with its environment, generates sensorimotor correlations that can be exploited to significantly facilitate subsequent processing and to reduce the complexity of the task, such as discrimination or categorization (Edelman, 1987; Beer, 1996; Nolfi and Parisi, 1999; Nolfi, 2002; Fitzpatrick and Metta, 2002; te Boekhorst et al., 2003). In addition, Chapter 5 demonstrated the arbitrariness, at least at the neural level, of any hierarchical distinction between sensory and motor modalities: they constitute an inseparable whole, and arbitrary classifications (such as assuming that neural activity flow mainly from sensory to motor modalities) might lead to artificial complications (such as the artificial need of complex predictive regulation). Similarly, neurobiological studies in primate vision have shown the difficulty and artificiality of studying separately the sensory system and the motor system, since the two systems are tightly integrated (Douglas et al., 1993).

One of the most important aspects of the neural architecture is, as first discussed in Chapters 2 and 3, the absence of predefined structure in the way how the sensors and motors are coupled together. Rather, structures in the neural cross-modal couplings and in the observed behaviors are spontaneously produced by the tight coupling between sensory and motor activity, both internally – through the neural network – as well as externally – through motion of the agent in its environment. This generation of structures illustrates the recent information-theoretical concept of **information self-structuring**, which states that for embodied agents, sensory-motor coordinated behavior generates statistical regularities in the sensory input that can be exploited to form cross-modal associations, which in turn can promote concept formation and other forms of high-level cognition (Lungarella and

Sporns, 2005, 2006).

The work of this thesis, whose essence is to explore behaviors that are spontaneously produced by a self-organizing dynamics, obviously corroborates the **principle of emergence**, in the sense that systems should be designed for emergence. (The controversial term of *emergence* is used here is a very pragmatic way, namely as not being programmed explicitly for the observed behavior.) One of the reasons underlying this principle is that systems designed for emergence – where the designer commitments are further removed from the actual behavior – tend to be more adaptive and more robust (Steels, 1991; Lungarella, 2004; Pfeifer et al., 2005). Several experiments of this thesis illustrate this latter point: for instance, Chapter 6 described how the turning strategy produced by the quadruped robotic dog is more efficient and more robust than commonly used control strategies; also, Chapters 4 and 5 discussed how the behaviors produced by the systems are more adaptive than currently existing models explicitly designed for the observed behaviors.

A closely related issue is the **frame-of-reference principle** (Clancey, 1989, 1991a), which conceptualizes the relation between the designer, the observed agent and the environment: internal mechanisms, as well as their complexity, cannot be deduced from the observed behavior, especially since behavior is always the result of the system-environment interaction. Clearly, the series of experiments presented in the previous chapters all illustrate this point: using the very same internal mechanisms but producing a variety of behaviors, ranging from insect navigation to predictive cancellation to delayed reward learning, they showed that the observed behaviors do not depend only on internal mechanisms, but – even more importantly – on the interaction between the agent and its environment. Also, the comparison between the behaviors produced by the same robotic agent in Chapters 3 and 7 demonstrates that behavioral complexity cannot be attributed to the agent alone, but to the agent-environment interaction (Simon, 1969).

The absence of any centralized control or hierarchy in the proposed neural architecture (Chapter 2) relates to the **principle of parallel, loosely coupled processes**, which is at the core of the embodied approach to cognitive science. It postulates that behavior does not require an explicit process that controls all the others. Indeed, the earlier chapters illustrated how the neural dynamics of the homogeneous network– i.e. the propagation of neural activity – can be self-regulated through coupling to an embodied agent (Chapters 2 and 3). In addition, the later chapters suggested how the proposed architecture – consisting of parallel, loosely coupled processes, namely the populations of neurons corresponding to all modalities – could feature an intrinsic stability (Chapters 5 and 6) and scale up with increasing complexity of the agent's sensorimotor apparatus (Chapters 5 and 7). Moreover, Chapter 7 illustrated how *the environment* could be exploited to couple the processes and coordinate their interaction. This is similar to how insect walking can be partly coordinated through the interaction of the legs with the real world: if one leg is lifted, the force on all the other legs changes instantaneously, which can be exploited for coordination (Cruse et al., 2002).

The cross-modal correlations of neural activity, which are at the core of the proposed neural architecture, are often produced by the partial overlap of sensory modalities based on different physical processes (e.g. camera and whiskers both detecting objects in Chapter 3). The importance of this partial overlap of functionality is captured by the **redundancy principle**. Interestingly, the importance of some redundancy was also recognized in motor systems (Chapters 4 and 6), in the neural coupling (Chapters 5 and 7), and even in the environment (Chapters 4 and 7).

Related to this – and somehow complementary – is the **principle of cheap design** (Iida, 2005a). Again, the experiments described in this thesis can illustrate the two aspects of this principle. First, good designs are parsimonious, such as the simple, passive and elastic legs that confer an inherent stability and behavioral diversity to the quadruped running robot (Chapter 6). Second, they should exploit the givens: the physics of the agent-environment (e.g. the elasticity and ground friction of the legs in Chapter 6, or simply the fact that self-rotation produces a shift of the perceived environment as in Chapters 3, 5 and 7), and the constraints of the ecological niche (e.g. the robot exploiting the sensorimotor experience produced by its constrained motion within the T-maze in Chapter 7).

In the same way, these chapters suggest some kind of balance, or task distribution between morphology (Chapters 4, 6 and 7), materials (Chapters 3 and 6) and even the environment (Chapter 7). Also, there is almost no preprocessing on either sensory and motor sides in the proposed neural architecture. This fact confers thus a certain balance to the complexity of the sensory, motor and neural systems of the agents. Furthermore, the progression from Chapter 3 to Chapter 7 suggests a certain relation between the complexity of the agent-environment pair and the complexity of the observed behavior. Taken together, these issues are the constituent parts of the **principle of ecological balance**, to which they provide an evident support.

Embodied AI emphasizes the importance of studying intelligent-like phenomena under multiple time scales. The three time scale perspectives that are typically incorporated in explanations of natural and artificial behaviors, and which are captured in the **time perspectives principle**, are: (a) the short-term ("here and now") perspective, (b) the ontogenetic (learning and development) perspective, and (c) the phylogenetic (evolutionary) perspective. The work of this thesis is, by its definition, primarily concerned on how the ontogenetic perspective – the (self-)development of structures in the neural architecture –

influences phenomena – behaviors – observed at the short-term perspective. In particular, Appendix 5.A explicitly discussed how two apparently distinct phenomena at the short-term perspective could nevertheless be explained within a single framework when considered from an ontogenetic perspective. Additionally, Chapter 7 presented and discussed an example of behavior that resulted from the *reciprocal* interaction between short-term and ontogenetic perspectives – i.e. how the immediate behavior of the agent, shaped by the development of the neural structures, could in turn allow a novel behavioral strategy to be learned on the long run. Interestingly, however, this thesis never touched upon the third time scale, the evolutionary perspective. The main reason why the essence of the proposed approach – i.e. observing intelligent-like self-organizing phenomena without purpose – can apparently not be yet directly translated into a phylogenetic perspective is probably due to limitations of the currently available technology. The phylogenetic view is indeed characterized by the phenomenon of self-replication. Self-replication, however, cannot be achieved without materials that grow or basic elements that assemble – something that remains, even today, a substantial challenge. On the other hand, one could justifiably argue that evolutionary processes can be simulated. Yet, there are a few issues that make simulated phylogeny still incompatible with the approach proposed in this thesis. First, simulated self-replication is by and large based on artificial selection, i.e. on the measure of some arbitrary fitness value of the agents – exactly the kind of designer bias providing a functional goal to the system that our approach tries to avoid. Second, evolutionary simulations become extraordinarily expensive computationally if the designer does not define an arbitrary process for the development of the simulated agents that reduces the complexity of the parameter space. There is thus a high risk of biasing the system with some arbitrary assumptions about how things are thought to be (e.g. only allowing synaptic coupling to grow from sensory to motor areas – remember however the discussion of Chapter 5) rather than exploring and understanding how things could be[1]. Third, the richness of the physics obtained in simulated worlds remains incomparably inferior to the one of the real world. Yet, it is precisely this richness in the physical dynamics of the world – flexible materials (Chapters 3 and 6), complex reaction forces (Chapter 6), or even a priori irrelevant visual structures (Chapter 7), to mention but a few examples – that plays an essential role in shaping various interesting behaviors.

This leads us to the last principle we would like to mention in this section: the **synthetic methodology principle**. This methodology of "understanding by building" lies at the heart of embodied AI. As hopefully convincingly shown in this thesis, this approach of building

---

[1] In this respect, genetic regulatory networks (Eggenberger, 1997; Bongard, 2002) look like a promising direction.

artificial systems to understand natural phenomena can be conceptually quite prolific – both to abstract general principles from the obtained results as well as to rethink our conception of the world around us. In particular, we discussed in Chapters 3, 6 and 7 how the study of systems interacting with the *real* world, with all its intrinsic richness and complexity – as opposed to virtual simulated worlds (even though the latter can still be a precious tool to deepen our understanding of particular observations) – allowed us to explore phenomena from an interesting, sometimes even surprising perspective.

8.2.2 Exploring Self-Organization

Before discussing further principles that can be drawn from the specific approach of the thesis, let us briefly summarize – now that we benefit from a global overview of obtained results – the essential qualities of the proposed neural architecture. First, it turned out to be able to generate, under appropriate conditions, quite a wide variety of a priori unrelated behaviors, including homing strategies observed in insects (Chapter 4), taxis with predictive reafference cancellation (Chapter 5), or solving a T-maze task (Chapter 7). Second, despite its homogeneous connectivity, the network and its dynamics could be nevertheless fully accounted for by an analysis of individual components, as proclaimed in Chapter 2. For instance, the role of the learned cross-modal correlations, once graphically represented, could always be fairly easily understood. Third, the progression through the successive chapters suggests the ability for the neural architecture to scale up – at least to a certain extent – with increasing complexity, such as random perturbations (Chapter 5), richer body dynamics (Chapter 6), or additional sensorimotor modalities and a more dynamical environment (Chapter 7). In particular, the complexity of the internal neural dynamics was shown to be spontaneously regulated by structures produced through the coupling of the network to an *embodied* system. For instance, Chapters 5 and 7 illustrated how additional sensory modalities could be added to the system without perturbing it. Finally, the neural architecture was also shown to be mostly independent from the free parameters of the models. Chapter 2 demonstrated for example how the dynamics of the neural network is in general invariant to the calibration of the sensorimotor signals. Also, the experience gained from conducting the various experiments strongly suggests that the qualitative aspects of the observed behaviors do not critically depend on specific choices for the remaining free parameters (such as the learning rate or the range of the sensorimotor signals). For instance, it can easily be shown that a smaller learning rate of the neural network would not change the qualitative observed behaviors, but only slow down the motions of the robot, or increase the time until particular behaviors are univocally discerned(such as the robot learning the correct strategy in the T-maze).

Let us turn now to the general principles that can be drawn from the particular approach proposed in this work. The first one, of course, concerns the **exploratory methodology** advocated already from the introduction and followed throughout the other chapters of the thesis. By exploring artificial embodied systems endowed with an arbitrary, self-developing dynamics, by trying to find conditions producing interesting behaviors, and by understanding how those are produced, it is possible to discover novel, even unsuspected perspectives. Offering an alternative to currently existing approaches that try either to achieve given behaviors (thus relying implicitly on the intelligence of the designer) or to find an underlying principle orchestrating intelligent-like behaviors (thus looking for an explicit purpose), this exploratory methodology allows not only to discern some of our preconceptions restraining our view, but also to discover new – and often simpler – principles possibly underlying natural behaviors, as well as common properties shared by a priori unrelated phenomena. What is more, the various artificial embodied systems developed in the field of embodied AI – i.e. robots – lend themselves wonderfully to such exploratory endeavor.

This work also stresses the importance, when studying embodied agents, to consider the importance of **multi-modal interactions**, rather than focusing solely on unimodal (and thus mostly sensory) processing. On the one hand, this allows taking advantage – as already discussed previously – of structures, such as cross-modal correlations, which are provided "for free" by the embodiment of the system. On the other hand, the multiple possible *combinations* of cross-modal interactions automatically provide a rich dynamical repertoire to the system, out of which many behaviors can potentially be produced – which are then interpreted as "new" or "learned" behaviors, as shown respectively in Chapters 3 and 7. In other words, the behavioral diversity, which is essential for adaptive systems, does not need to be explicitly taken care of, but can be spontaneously achieved when multiple sensory and motor modalities are let to interact with each other.

Several experiments also illustrated the importance of not excluding from the start factors that seem irrelevant. For instance, Chapter 5 showed the benefit of leaving reciprocal synaptic coupling in the neural architecture, even though they did not play any significant role in the other experiments. Chapter 6 illustrated the importance of representing multiple motor parameters, so that the agent could discover a more robust motor strategy for turning. Also, Chapter 4 demonstrated the importance, in the observed homing behaviors, of two sensory modalities a priori not related at all to navigation. Similarly, Chapter 7 showed the essential role that could be played, in a learning task, by both a sensory modality and some structure in the environment that were entirely unrelated to the task. This principle, which can be summarized as **leaving room for emergence**, is in many respects closely related to the principle of emergence discussed in the previous subsection.

A further interesting characteristic of this work is that it is concerned with embodied agents without any predefined behavior, which develop cross-modal coupling structures that are initially inexistent. Allegorically, the systems display an **expanding dynamics**. This sharply contrasts with the common view of learning as a phenomenon of *convergence* towards a few numbers of locally optimal solutions, where an initially unconstrained exploratory behavioral repertoire is typically progressively *reduced* towards an ultimate ("optimal") behavior. On the other hand, it is in line with more recent views that stress the importance of development, such as the developmental mechanism of freezing and freeing (Bernstein, 1967), where degrees of freedom in a neuro-musculo-skeletal system are progressively released, allowing an increasingly complex dynamics to take place (Sporns and Edelman, 1993; Berthouze and Lungarella, 2004).

Finally, another issue the work of this thesis wants to bring to light is the question whether a theory of intelligent behavior might ever crystallize in a **mathematical** form. Does the amazing success the mathematical language has had in physics to describe inanimate matter imply that it is the appropriate formalism to describe animated behavior too? This doubt can be accentuated by the following examples. The essential aspect of Chapter 4 isn't so much the two observed behaviors – which could individually be described mathematically – rather than the fact that both *closely match* distinct homing strategies displayed by a *single* natural organism, such as the ant Cataglyphis – a similarity that clearly lies beyond the realm of any mathematical formalism. Also, Chapter 7 discussed the shortcomings of an information theoretical view on the observed phenomenon: the robot could only learn to solve the delayed reward task in the presence of a structure in the environment (the red wall), which was entirely *neutral* to the task – neither interfering with the robot, nor carrying any information whatsoever about the position of the cue or the reward. This apparent paradox evokes an amusing riddle[2] that logic nuts like to challenge each other with, whose solution requires the occurrence of an event that desperately seems to carry

---

[2] A malicious demon has captured several very wise logicians, and locked them in a room they all want to escape. Most of the logicians have been marked on their forehead by a red dot. On one wall of the room, there is a door, leading to the exit, which the demon briefly opens every hour. Everyone who gets through this door shall be asked by the demon whether he has a red dot on his forehead or not. A correct answer leads to freedom, whereas a wrong one is punished by immediate death. To increase the dramatic side of the story – or simply to make this riddle a riddle – the logicians are strictly forbidden to communicate with each other. Before starting the game, as the logicians are desperately leering at each other's forehead, the demon says loud in the room – with a mischievous smile on his face: "I see that at least one of you has got a red dot on his forehead!" Will the logicians be able to all escape safely?

The amusing fact about this riddle is that the last phrase of the demon, which obviously does not give any new information to any of the logicians (they all see other logicians wandering around with some red dots on their forehead), is the key for them – and the person challenged with the riddle – to find a solution.

no information. Both cases, even though some would prefer to consider them as purely anecdotal, nevertheless divulge some intrinsic explanatory limits of an information-based mathematical formalism.

8.3 Outlook

The contribution of this work is two-fold. Conceptually, it supports the view of adaptive, intelligent-like behaviors as self-organizing phenomena; it suggests a methodology to investigate principles that can be gained or strengthened from this perspective, and tries to identify the preconceptions that bias our view and constrain the understanding of the world around us. Technically, this work proposes a concrete neural architecture for the exploration of self-organizing behaviors with embodied artificial agents, and provides a series of illustrating case studies.

Let us first discuss possible future perspectives regarding the latter, more technical aspect. Evidently, the series of experiments conducted in this thesis only investigated some possible initial steps towards a promising research direction. Because of its predominantly explorative nature, this inceptive contribution offers many ways of being extended.

Beside the obvious investigation of systems with all sorts of sensorimotor morphologies, the interaction of multiple agents – including the interaction of agents at different levels of development, i.e. with different observable "skills" – promises interesting complementary perspectives on the self-organization of proto-communication (Steels, 2003; see also Quinn, 2001; Oudeyer, 2005), ranging from more implicit stigmergic interactions through the environment to more explicit "communication" if the agents dispose of a particular (e.g. acoustic) sensorimotor apparatus.

Whiskers – whose study provided the hotbed of the work of this thesis (Bovet et al., 2004; Fend et al., 2004, 2006) – turned out to be a prolific sensory modality to investigate, especially since they allow the exploration of the world from a perspective we, humans, are not used to, and where research can thus less rely on our intuition. Similarly, the exploration of robots engaged in environments that radically differ from our natural ecological niche – for instance, fish-like robots swimming in water, or light-weighted robots exploiting passive aerodynamics – might most likely bring further interesting aspects to light.

On the way towards these exciting, but technologically challenging directions, there are still many questions whose investigation with currently existing hardware could be fruitful. These include the role of different sensory and motor representations, such as population coding (Georgopoulos et al., 1986) or more complex receptive fields (Hubel and Wiesel, 2005), as well as the study of agents placed is varying environments, looking thus into how a self-organizing dynamics can cope with progressively changing conditions – and so into the observable *transfer of learning* (Perkins and Salomon, 1992).

As a final point, let us discuss the outlook for the conceptual contribution of this thesis. By showing that the purpose of natural and artificial organisms might only exist in the eyes of an observer, and by illustrating the significant progress that can be achieved through a synthetic methodology aware of this fact, this thesis opens a way not only to rethink preconceived ideas that restrain our understanding, but also to explore the flourishing opportunities offered by this new perspective.

The particular neural architecture explored in the chapters of this thesis is therefore just one possible approach out of many. It is true that finding a common neural mechanism underlying a wide range of behaviors is certainly intriguing. Nevertheless, there are undoubtedly many other paths heading for the same direction: beside "robots with self-developing brains," why not also explore some "self-developing robots without brain"?

Artificial intelligence, as it is conceived today, is still far from being reached – and may simply never be. Nevertheless, the way to this complex endeavor – similar to the quest of physics to discover the ultimate law governing the inanimate universe around us – is sprinkled with conceptual gems whose harvest might teach us more than what we ever could expect. Further exciting days still await to be lived.

Bibliography

Alexander, M. (1990). Three uses for springs in legged locomotion. *Int. J. Robotics Research*, 9(2):53–61.

Almassy, N., Edelman, G. M., and Sporns, O. (1998). Behavioral constraints in the development of neuronal properties: A cortical model embedded in a real world device. *Cerebral Cortex*, 8:346–361.

Andel, D. and Wehner, R. (2004). Path integration in desert ants, Cataglyphis: how to make a homing ant run away from home. *Proc. R. Soc. Lond.*, 271:1485–1489.

Andry, P., Gaussier, P., and Nadel, J. (2002). From visuo-motor development to low-level imitation. In *Proceedings of the Second International Workshop on Epigenetics Robotics*, pages 7–15.

Arabzadeh, E., Panzeri, S., and Diamond, M. E. (2006). Deciphering the spike train of a sensory neuron: Counts and temporal patterns in the rat whisker pathway. *J. Neurosci.*, 26:9216–9226.

Asada, M., MacDorman, K. F., Ishiguro, H., and Kuniyoshi, Y. (2001). Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Robotics and Autonomous Systems*, 37:185–193.

Ashby, W. R. (1947). Principles of the self-organizing dynamic system. *Journal of General Psychology*, 37:125–128.

Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman and Hall, London.

Ashcraft, M. H. (1994). *Human memory and cognition*. HarperCollins College Publishers, New York, 2nd edition.

Atkinson, R. C. and Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In Spence, K. W., editor, *The psychology of learning and motivation: Advances in research and theory*, volume 2, pages 89–195. Academic Press, New York.

Baddeley, A. (1986). *Working Memory*. Oxford University Press, Oxford.

Baddeley, A. (1998). *Human memory: theory and practice*. Allyn and Bacon, Boston, MA, rev. edition.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4(10):829–839.

Barron, J., Fleet, D., and Beauchemin, S. (1994). Performance of optical flow techniques. *International Journal of Computer Vision*, 12:43–77.

Bartlett, F. C. (1932). *Remembering*. Cambridge University Press, Cambridge, UK.

Bässler, U. and Büschges, A. (1998). Pattern generation for stick insect walking movements – multisensory control of a locomotor program. *Brain Research Reviews*, 27:65–88.

Bastian, J. (1995). Pyramidal-cell plasticity in weakly electric fish: a mechanism for attenuating responses to reafferent electrosensory inputs. *J. Comp. Physiol. A*, 176(1):63–78.

Bastian, J. (1999). Plasticity of feedback inputs in the apteronotid electrosensory system. *Journal of Experimental Biology*, 202(10):1327–1337.

Bays, P. M., Flanagan, J. R., and Wolpert, D. M. (2006). Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biology*, 4(2):e28:1–4.

Bays, P. M., Wolpert, D. M., and Flanagan, J. R. (2005). Perception of the consequences of self-action is temporally tuned and event driven. *Current Biology*, 15(12):1125–1128.

Beer, R. D. (1996). Toward the evolution of dynamical neural networks for minimally cognitive behavior. In Maes, P., Mataric, M., Meyer, J.-A., Pollack, J., and Wilson, S., editors, *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pages 421–429, Cambridge, Mass. MIT Press. A Bradford Book.

Bell, C., Bodznick, D., Montgomery, J., and Bastian, J. (1997). The generation and subtraction of sensory expectations within cerebellum-like structures. *Brain Behav. Evol.*, 50:17–31.

Bell, C. C. (1982). Properties of a modifiable efference copy in an electric fish. *Journal of Neurophysiology*, 47(6):1043–1056.

Bell, C. C. (2001). Memory-based expectations in electrosensory systems. *Current Opinion in Neurobiology*, 11:481–487.

Bernstein, N. (1967). *The Co-ordination and Regulation of Movements*. Pergamon, London.

Berthouze, L. and Lungarella, M. (2004). Motor skill acquisition under environmental perturbations: on the necessity of alternate freezing and freeing of degrees of freedom. *Adaptive Behavior*, 12(1):47–63.

Bhushan, N. and Shadmehr, R. (1999). Evidence for a forward dynamics model in human adaptive motor control. In Kearns, M. S., Solla, S. A., and Cohn, D. A., editors, *Advances in Neural Information Processing Systems*, volume 11, pages 3–9. MIT Press, Cambridge, Mass.

Blakemore, S.-J., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, 11(11):R11–R16.

Böhm, H., Schildberger, K., and Huber, F. (1991). Visual and acoustic course control in the cricket Gryllus Bimaculatus. *Journal of Experimental Biology*, 159:235–248.

Bongard, J. (2002). Evolving modular genetic regulatory networks. In *Proceedings of the IEEE 2002 Congress on Evolutionary Computation (CEC2002)*, pages 305–311, Cambridge, Mass. MIT Press.

Borst, A. and Egelhaaf, M. (1993). Detecting visual motion: Theory and models. In Miles, F. A. and Wallman, J., editors, *Visual Motion and its Role in the Stabilization of Gaze*, pages 3–27. Elsevier Science.

Bovet, S. (2006). Emergence of insect navigation strategies from homogeneous sensorimotor coupling. In *Proceedings of the 9th International Conference on Intelligent Autonomous Systems (IAS-9)*, pages 525–533, Tokyo.

Bovet, S., Fend, M., and Pfeifer, R. (2004). Simulating whisker sensors – on the role of material properties for morphology, behavior and evolution. In Schaal, S., Ijspeert, A., Billard, A., Vijayakumar, S., Hallam, J., and Meyer, J.-A., editors, *From animals to animats 8: Proceedings of the Eighth International Conference on Simulation of Adaptive Behavior*, pages 122–129. MIT Press.

Bovet, S. and Pfeifer, R. (2005a). Emergence of coherent behaviors from homogeneous sensorimotor coupling. In *Proceedings of the 12th International Conference on Advanced Robotics (ICAR)*, pages 324–330, Seattle.

Bovet, S. and Pfeifer, R. (2005b). Emergence of delayed reward learning from sensorimotor coordination. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 841–846, Edmonton.

Braitenberg, V. (1984). *Vehicles. Experiments in synthetic psychology*. MIT Press, Cambridge, Mass.

Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin de la Société Anatomique*, 6:330–357.

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

Brooks, R. A. (1991). New approaches to robotics. *Science*, 307:1227–1232.

Brooks, R. A. (1999). *Cambrian intelligence: The early history of the new AI*. MIT Press, Cambridge, Mass. A Bradford Book.

Brünnert, U., Kelber, A., and Zeil, J. (1994). Ground-nesting bees determine the location of their nest relative to a landmark by other than angular size cues. *J. Comp. Physiol. A*, 175(3):363–369.

Bursen, H. A. (1978). *Dismantling the memory machine: A philosophical investigation of machine theories of memory*. Reidel, Dordrecht; London.

Büschges, A. and Wolf, H. (1999). Phase-dependent presynaptic modulation of mechanosensory signals in the locust flight system. *Journal of Neurophysiology*, 81(2):959–962.

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., and Bonabeau, E. (2001). *Self-Organization in Biological Systems*. Princeton University Press, Princeton, New Jersey.

Cartwright, B. A. and Collett, T. S. (1983). Landmark learning in bees. *Journal of Comparative Physiology*, 151:521–543.

Clancey, W. J. (1989). The frame of reference problem in cognitive modeling. In *Proceedings The Eleventh Annual Conference of the Cognitive Science Society*, pages 107–114, Hillsdale, New Jersey. Lawrence Erlbaum.

Clancey, W. J. (1991a). The frame of reference problem in the design of intelligent machines. In van Lehn, K., editor, *Architectures for intelligence: The twentysecond Carnegie symposium on cognition*, pages 357–423. Lawrence Erlbaum, Hillsdale, New Jersey.

Clancey, W. J. (1991b). Review of Rosenfield's "The Invention of Memory". *Artificial Intelligence*, 50(2):241–284.

Clancey, W. J. (1997). *Human memory: Theory and practice*. Psychology Press, Hove, UK, revised edition.

Collett, M. and Collett, T. S. (2000). How do insects use path integration for their navigation? *Biological Cybernetics*, 83:245–259.

Collett, T. S. (1992). Landmark learning and guidance in insects. *Philosophical Transactions: Biological Sciences*, 337(1281):295–303.

Collins, S., Ruina, A., Tedrake, R., and Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science*, 307(5712):1082–1085.

Collins, S. H., Wisse, M., and Ruina, A. (2001). A three-dimensional passive-dynamic walking robot with two legs and knees. *International Journal of Robotics Research*, 20(7):607–615.

Connell, J. H. and Mahadevan, S., editors (1993). *Robot learning*, Boston. Kluwer Academic Publishers.

Cruse, H., Dean, J., Dürr, V., Kindermann, T., Schmitz, J., and Schumm, M. (2002). A decentralized, biologically based network for autonomous control of (hexapod) walking. In Ayers, J. and Davis, J., editors, *Neurotechnology for biomimetic robots*, pages 383–400, Cambridge, Mass. MIT Press.

Cruse, H., Dürr, V., and Schmitz, J. (2007). Insect walking is based on a decentralized architecture revealing a simple and robust controller. *Phil. Trans. R. Soc. A*, 265:221–250.

Crusio, W. E. and Schwegler, H. (2005). Learning spatial orientation tasks in the radial-maze and structural variation in the hippocampus in inbred mice. *Behavioral and Brain Functions*, 1(3).

Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human-robot interaction. *Phil. Trans. R. Soc. B*, 362:679–704.

Dautenhahn, K. and Nehaniv, C. L., editors (2002). *Imitation in Animals and Artifacts*. MIT Press, Cambridge, Mass. A Bradford Book.

Davidson, P. R. and Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: a short review. *J. Neural Eng.*, 2:S313–S319.

Deneubourg, J.-L., Goss, S., Franks, N., Sendova-Franks, A., Detrain, C., and Chrétien, L. (1991). The dynamics of collective sorting: Robot-like ants and ant-like robots. In Meyer, J.-A. and Wilson, S., editors, *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, pages 356–363, Cambridge, Mass. MIT Press. A Bradford Book.

D'Hooge, R. and De Deyn, P. P. (2001). Applications of the morris water maze in the study of learning and memory. *Brain Research Reviews*, 36(1):60–90.

Dickinson, M. H. (2000). How animals move: An integrative view. *Science*, 288(5463):100–106.

Douglas, R. J., Martin, K. A., and Nelson, J. C. (1993). The neurobiology of primate vision. *Baillière's Clinical Neurology*, 2(2):191–225.

Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Computation*, 12:219–245.

Doya, K., Kimura, H., and Kawato, M. (2001). Neural mechanisms of learning and control. *Control Systems Magazine, IEEE*, 21(4):42–54.

Duhamel, J. R., Colby, C. L., and Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040):90–92.

Edelman, G. M. (1987). *Neural Darwinism: The Theory of Neural Group Selection*. Basic Books, New York.

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

Eggenberger, P. (1997). Evolving morphologies of simulated 3d organisms based on differential gene expression. In Husbands, P. and Harvey, I., editors, *Proceedings of the Fourth European Conference on Artificial Life*, pages 205–213, Cambridge, Mass. MIT Press.

El Manira, A., DiCaprio, R. A., Cattaert, D., and Clarac, F. (1991). Monosynaptic interjoint reflexes and their central modulation during fictive locomotion in crayfish. *European Journal of Neuroscience*, 3:1219–1231.

Esch, H. E. and Burns, J. E. (1995). Honeybees use optic flow to measure the distance of a food source. *Naturwissenschaften*, 82(1):38–40.

Etienne, A. S. and Jeffrey, K. J. (2004). Path integration in mammals. *Hippocampus*, 14:180–192.

Fend, M., Bovet, S., and Hafner, V. V. (2004). The Artificial Mouse - a robot with whiskers and vision. In *Proceedings of the 35th International Symposium on Robotics (ISR)*, Paris.

Fend, M., Bovet, S., and Pfeifer, R. (2006). On the influence of morphology of tactile sensors for behavior and control. *Robotics and Autonomous Systems*, 54(8):686–695.

Fend, M., Bovet, S., Pfeifer, R., and Yokoi, H. (2003). An active artificial whisker array for texture discrimination. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 1044–1049, Las Vegas.

Fermüller, C., Shulman, D., and Aloimonos, Y. (2001). The statistics of optical flow. *Computer Vision and Image Understanding*, 82:1–32.

Fitzpatrick, P. and Metta, G. (2002). Towards manipulation-driven vision. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Lausanne, Switzerland.

Flanagan, J. R., King, S., Wolpert, D. M., and Johansson, R. S. (2001). Sensorimotor prediction and memory in object manipulation. *Canadian JOurnal of Experimental Psychology*, 55(2):89–97.

Franz, M. O. and Mallot, H. A. (2000). Biomimetic robot navigation. *Robotics and Autonomous Systems*, 30:133–153.

Franz, M. O., Schölkopf, B., Mallot, H. A., and Bülthoff, H. H. (1998). Where did I take the snapshot? Scene-based homing by image matching. *Biol. Cybern.*, 79:191–202.

Fuster, J. M. (1997). Network memory. *Trends in Neurosciences*, 20(10):451–459.

Gebhardt, M. and Honegger, H.-W. (2001). Physiological characterisation of antennal mechanosensory descending interneurons in an insect (gryllus bimaculatus, gryllus campestris) brain. *Journal of Experimental Biology*, 204:2265–2275.

Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20:1–55.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, 14:477–485.

Götz, K. G. (1975). The optomotor equilibrium of the Drosophila navigation system. *Journal of Comparative Physiology A*, 99(3):187–210.

Gray Walter, W. (1950). An imitation of life. *Scientific American*, 182(5):42–45.

Gray Walter, W. (1951). A machine that learns. *Scientific American*, 185(2):60–63.

Grillner, S. (1996). Neural networks for vertebrate locomotion. *Scientific American*, 274(1):64–69.

Grüsser, O.-J. (1995). On the history of the ideas of efference copy and reafference. *Clio Med.*, 33:35–55.

Hafner, V. V., Kunz, H., and Pfeifer, R. (2002). An investigation into obstacle avoidance as an ‘emergent’ behaviour from two different perspectives. In *Proc. of the EPSRC/BBSRC Int. Workshop on Biol.-Inspired Robotics: The Legacy of W. Grey Walter*, pages 166–173, Bristol.

Hafner, V. V. and Möller, R. (2001). Learning of visual navigation strategies. In Quoy, M., Gaussier, P., and Wyatt, J., editors, *Proc. of the Eur. Workshop on Learning Robots (EWLR-9)*, pages 47–56, Prague.

Hara, F. and Pfeifer, R. (2000). On the relation among morphology, material and control in morpho-functional machines. In Meyer, Berthouze, Floreano, Roitblat, and Wilson, editors, *From Animals to Animats: 6. Proceedings of the sixth International Conference on Simulation of Adaptive Behavior*, pages 33–40.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

Haruno, M., Wolpert, D. M., and Kawato, M. (2001). MOSAIC model for sensorimotor learning and control. *Neural Computation*, 13:2201–2220.

Hebb, D. O. (1949). *The organization of behavior*. John Wiley, New York.

Heran, H. and Wanke, L. (1952). Beobachtungen über die Entfernungsmeldung der Sammelbienen. *Z. vergl. Physiol.*, 34:383–393.

Hipp, J., Arabzadeh, E., Diamond, M. E., Kayser, C., Conrad, J., Zorzin, E., and König, P. (2006). Texture signals in whisker vibrations. *J. Neurophysiol.*, 95:1792–1799.

Hodges, H. (1996). Maze procedures: the radial-arm and water maze compared. *Brain Res Cogn Brain Res*, 3:167–181.

Horn, B. K. P. and Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17:185–203.

Horn, B. K. P. and Schunck, B. G. (1994). Determining optical flow: a retrospective. In Bobrow, D. G., editor, *Artificial intelligence in perspective*, pages 81–87, Cambridge, Mass. MIT Press.

Hosoda, K., Tada, Y., and Asada, M. (2006). Anthropomorphic robotic soft fingertip with randomly distributed receptors. *Robotics and Autonomous Systems*, 54(2):104–109.

Hubel, D. H. and Wiesel, T. N. (2005). *Brain and visual perception: the story of a 25-year collaboration*. Oxford University Press, Oxford.

Iida, F. (2005a). *Cheap design and behavioral diversity for adaptive autonomous robots*. PhD dissertation, University of Zurich, Faculty of Mathematics and Science.

Iida, F. (2005b). Cheap design approach to adaptive behavior: Walking and sensing through body dynamics. *International Symposium on Adaptive Motion of Animals and Machines*.

Iida, F., Gómez, G. J., and Pfeifer, R. (2005). Exploiting body dynamics for controlling a running quadruped robot. In *Proc. of the 12th Int. Conf. on Advanced Robotics (ICAR05)*, pages 229–235, Seattle.

Iida, F. and Lambrinos, D. (2000). Navigation in an autonomous flying robot by using a biologically inspired visual odometer. *Sensor Fusion and Decentralized Control in Robotic System III, Photonics East, Proceeding of SPIE*, 4196:86–97.

Iida, F. and Pfeifer, R. (2004). "Cheap" rapid locomotion of a quadruped robot: self-stabilization of bounding gait. In Groen, F., Amato, N., Bonarini, A., Yoshida, E., and Krose, B., editors, *Intelligent Autonomous Systems 8*, pages 642–649. IOS Press, Amsterdam.

Iida, F. and Pfeifer, R. (2005). Structuring sensory information through body dynamics. Technical report, IROS05 Workshop on Morphology, Control and Passive Dynamics.

Iida, F., Pfeifer, R., Steels, L., and Kuniyoshi, Y., editors (2004). *Embodied Artificial Intelligence*, Heidelberg. Springer Verlag.

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

Ishiguro, A., Shimizu, M., and Kawakatsu, T. (2004). Don't try to control everything!: An emergent morphology control of a modular robot. In *Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 981–985, Sendai, Japan.

Ishiguro, A., Shimizu, M., and Kawakatsu, T. (2006). A modular robot that exhibits amoebic locomotion. *Robotics and Autonomous Systems*, 54(8):641–650.

Johansson, R. S. and Cole, K. J. (1992). Sensory-motor coordination during grasping and manipulative actions. *Current Opinion in Neurobiology*, 2(6):815–823.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285.

Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (1991). *Principles of neural science*. Elsevier, New York.

Kaneko, K. and Tsuda, I. (2001). *Complex systems: chaos and beyond*. Springer, Berlin.

Kaplan, F. and Oudeyer, P.-Y. (2004). Maximizing learning progress: An internal reward system for development. In Iida, F., Pfeifer, R., Steels, L., and Kuniyoshi, Y., editors, *Embodied Artificial Intelligence, International Seminar, Dagstuhl Castle, Germany, July 7-11, 2003, Revised Papers*, pages 259–270, Berlin Heidelberg. Springer-Verlag.

Kauffman, S. (1993). *Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, Oxford.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(7):718–727.

Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., and Yoshioka, T. (2003). Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. In Prablanc, C., Pélisson, D., and Rossetti, Y., editors, *Progress in Brain Research*, volume 142, pages 171–188. Elsevier Science.

Kelso, J. A. S. (1995). *Dynamic patterns: the self-organization of brain and behavior*. MIT Press, Cambridge, Mass. A Bradford Book.

Kim, D. (2004). Evolving internal memory for T-maze tasks in noisy environments. *Connection Science*, 16(3):183–210.

Koriat, A. and Goldsmith, M. (1996). Memory metaphors and the laboratory/real-life controversy: Correspondence versus storehouse views of memory. *Behavioral and Brain Sciences*, 19:167–188.

Kosko, B. (1986). Differential hebbian learning. *AIP Conference Proceedings*, 151(1):277–282.

Kuniyoshi, Y. and Sangawa, S. (2006). Early motor development from partially ordered neural-body dynamics: experiments with a cortico-spinal-musculo-skeletal model. *Biological Cybernetics*, 95(6):589–605.

Kuniyoshi, Y. and Suzuki, S. (2004). Dynamic emergence and adaptation of behavior through embodiment as coupled chaotic field. In *Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2042–2049, Sendai, Japan.

Kuniyoshi, Y., Yorozu, Y., Ohmura, Y., Terada, K., Otani, T., Nagakubo, A., and Yamamoto, T. (2004). From humanoid embodiment to theory of mind. In Iida, F., Pfeifer, R., Steels, L., and Kuniyoshi, Y., editors, *Embodied Artificial Intelligence, International Seminar, Dagstuhl Castle, Germany, July 7-11, 2003, Revised Papers*, pages 202–218, Berlin Heidelberg. Springer-Verlag.

Lakoff, G. and Núñez, R. E. (2000). *Where mathematics comes from : how the embodied mind brings mathematics into being*. Basic Books, New York.

Lambrinos, D. (1999). *Navigation in biorobotic agents*. PhD thesis, University of Zurich, Department of Computer Science.

Lambrinos, D., Möller, R., Labhart, T., Pfeifer, R., and Wehner, R. (2000). A mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems, special issue: Biomimetic Robots*, 30:39–64.

Langton, C. G., editor (1995). *Artificial life : an overview*. Bradford Books series in complex adaptive systems. MIT Press, Cambridge, Mass.

Lappe, M., editor (2000). *Neuronal Processing of Optic Flow*. Academic Press, San Diego.

Levi, R. and Camhi, J. M. (2000). Population vector coding by the giant interneurons of the cockroach. *The Journal of Neuroscience*, 20(10):3822–3829.

Lichtensteiger, L. (2005). *Bodies That Think Quickly and Learn Fast: On the Interdependence of Morphology and Control for Intelligent Behavior*. Shaker, Aachen, Germany.

Lipson, H. and Pollack, J. B. (2000). Automatic design and manufacture of artificial life forms. *Nature*, 406:974–978.

Lungarella, M. (2004). *Exploring Principles Towards a Developmental Theory of Embodied Artificial Intelligence*. Unpublished PhD thesis, University of Zurich, Switzerland.

Lungarella, M., Hafner, V., Pfeifer, R., and Yokoi, H. (2002a). An artificial whisker sensor for robotics. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2931–2936.

Lungarella, M., Hafner, V., Pfeifer, R., and Yokoi, H. (2002b). Whisking: an unexplored sensory modality. In *Proceedings of the Seventh International Conference on the Simulation of Adaptive Behavior (SAB'02)*, pages 58–59. MIT Press.

Lungarella, M., Pegors, T., Bulwinkle, D., and Sporns, O. (2005). Methods for quantifying the informational structure of sensory and motor data. *Neuroinformatics*, 3(3):243–262.

Lungarella, M. and Sporns, O. (2005). Information self-structuring: Key principle for learning and development. In *Proc. of the 4th Int. Conf. on Development and Learning*, pages 25–30.

Lungarella, M. and Sporns, O. (2006). Mapping information flow in sensorimotor networks. *PLoS Comput Biol*, 2(10):e144.

MacLeod, C. M. (1997). Is memory caught in the mesh? *Behavioral and Brain Sciences*, 20:30. Peer commentary on Glenberg (1997).

Maris, M. and te Boekhorst, R. (1996). Exploiting physical constraints: Heap formation through behavioral error in a group of robots. In Asada, M., editor, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1655–1660.

Matarić, M. and Cliff, D. (1996). Challenges in evolving controllers for physical robots. *Robotics and Autonomous Systems*, 19(1):67–83.

Maturana, H. R. and Varela, F. J. (1992). *The Tree of Knowledge. The Biological Roots of Human Understanding*. Shambhala USA, revised edition.

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

McKinstry, J. L., Edelman, G. M., and Krichmar, J. L. (2006). A cerebellar model for predictive motor control tested in a brain-based device. *PNAS*, 103(9):3387–3392.

McMahon, T. A. (1984). *Muscles, reflexes, and locomotion*. Princeton University Press, Princeton, New Jersey.

McMahon, T. A. and Cheng, G. C. (1990). The mechanics of running: how does stiffness couple with speed? *J. Biomech.*, 23(Suppl 1):65–78.

Merfeld, D. M., Zupan, L., and Peterka, R. J. (1999). Humans use internal models to estimate gravity and linear acceleration. *Nature*, 398:615–618.

Miall, R. C. and Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, 9:1265–1279.

Minsky, M. L. (1961). Steps toward artificial intelligence. *Proceedings of the Institute of Radio Engineers*, 49:8–30.

Mittelstaedt, H. (1983). The role of multimodel convergence in homing by path integration. *Fortschr. Zool.*, 28:197–212.

Möller, R. (2000). Insect visual homing strategies in a robot with analog processing. *Biol. Cybern.*, 83(3):231–243.

Möller, R., Lambrinos, D., Pfeifer, R., Labhart, T., and Wehner, R. (1998). Modeling ant navigation with an autonomous agent. In Pfeifer, R., Blumberg, B., Meyer, J., and Wilson, S., editors, *From Animals to Animats 5, Proceedings of SAB98*, pages 185–194. MIT Press, Cambridge, Mass.

Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods*, 11(1):47–60.

Müller, M. and Wehner, R. (1988). Path integration in desert ants, cataglyphis fortis. *Proc. Natl. Acad. Sci. USA*, 85:5287–5290.

Müller, M. and Wehner, R. (1994). The hidden spiral: systematic search and path integration in desert ants, Cataglyphis fortis. *J. Comp. Physiol. A*, 175:525–530.

Neisser, U. (1978). Memory: What are the important questions? In Gruneberg, M. M., Morris, P., and Sykes, R., editors, *Practical aspects of memory*, pages 3–24. Academic Press, London.

Nicolis, G. and Prigogine, I. (1977). *Self-organization in nonequilibrium systems : from dissipative structures to order through fluctuations*. Wiley, New York.

Nolfi, S. (2002). Power and the limits of reactive agents. *Neurocomputing*, 49:119–145.

Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press, Cambridge, Mass.

Nolfi, S. and Parisi, D. (1999). Exploiting the power of sensory-motor coordination. *Lecture Notes in Computer Science*, 1674:173–182.

O'Keefe, J. (1979). A review of the hippocampal place cells. *Progress in Neurobiology*, 13(4):419–439.

O'Keefe, J. and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1):171–5.

Olsson, L., Nehaniv, C. L., and Polani, D. (2006). From unknown sensors and actuators to actions grounded in sensorimotor perceptions. *Connection Science, Special Issue on Developmental Robotics*, 18(2):243–262.

Olton, D. S. (1987). The radial arm maze as a tool in behavioral pharmacology. *Physiol. Behav.*, 40(6):793–7.

Oudeyer, P.-Y. (2005). The self-organization of speech sounds. *Journal of Theoretical Biology*, 233:435–449.

Pasternak, T. and Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, 6:97–107.

Perkins, D. N. and Salomon, G. (1992). Transfer of learning. In *International Encyclopedia of Education (2nd Edition)*, Oxford, UK. Pergamon Press.

Pfeifer, R. (1996). Building "fungus eaters": Design principles of autonomous agents. In Maes, P., Mataric, M., Meyer, J.-A., Pollack, J., and Wilson, S., editors, *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pages 3–12, Cambridge, Mass. MIT Press. A Bradford Book.

Pfeifer, R. and Bongard, J. (2007). *How the body shapes the way we think: a new view of intelligence*. MIT Press, Cambridge, Mass.ss.

Pfeifer, R., Iida, F., and Bongard, J. (2005). New robotics: Design principles for intelligent systems. *Artificial Life*, 11:99–120.

Pfeifer, R. and Scheier, C. (1999a). *Understanding Intelligence*. MIT Press, Cambridge, Mass.

Pfeifer, R. and Scheier, C. (1999b). *Understanding Intelligence*. MIT Press, Cambridge, Mass.ss.

Pollack, G. S. (1998). Neural processing of acoustic signals. In Popper, A. N., Hoy, R. R., and Fay, R. R., editors, *Comparative hearing: insects*, pages 139–196. Springer.

Poulet, J. F. and Hedwig, B. (2002). A corollary discharge maintains auditory sensitivity during sound production. *Nature*, 418:872–876.

Poulet, J. F. and Hedwig, B. (2003). A corollary discharge mechanism modulates central auditory processing in singing crickets. *J Neurophysiol*, 89:1528–1540.

Pylyshyn, Z. W., editor (1987). *The Robots dilemma : the frame problem in artificial intelligence*. Ablex Publishing Corporation, Norwood, NJ.

Quinn, M. (2001). Evolving communication without dedicated communication channels. In Kelemen, J. and Sosk, P., editors, *Advances in Artificial Life: Proceedings of the 6th European Conference (ECAL 2001)*, Lectures Notes in Computer Science, pages 357–366, Prague. Springer.

Reeve, R. R. and Webb, B. (2003a). New neural circuits for robot phonotaxis. *Phil. Trans. R. Soc. Lond. A*, 361:2245–2266.

Reeve, R. R. and Webb, B. (2003b). Reafferent or redundant: Integration of phonotaxis and optomotor behavior in crickets and robots. *Adaptive Behavior*, 11(3):137–158.

Reichardt, W. (1969). Movement perception in insects. In Reichardt, W., editor, *Processing of optical data by organisms and machines*, pages 465–493. Academic Press, New York.

Ronacher, B., Gallizzi, K., Wohlgemuth, S., and Wehner, R. (2000). Lateral optic flow does not influence distance estimation in the desert ant Cataglyphis fortis. *J. Exp. Biol.*, 203(7):1113–1121.

Ronacher, B. and Wehner, R. (1995). Desert ants Cataglyphis fortis use self-induced optic flow to measure distances travelled. *J. Comp. Physiol. A*, 177(1):21–27.

Rosenfield, I. (1992). *The invention of memory: A new view of the brain*. Basic Books, New York, 2nd edition.

Rosheim, M. E. (1996). Leonardo's lost robot. *Achademia Leonardi Vinci*, IX:99–110.

Russo, P., Webb, B., Reeve, R., Arena, P., and Patané, L. (2005). A cricket-inspired neural network for feedforward compensation and multisensory integration. *IEEE Conference on Decision and Control and European Control Conference*.

Rylatt, R. M. and Czarnecki, C. A. (2000). Embedding connectionist autonomous agents in time: The 'road sign problem'. *Neural Processing Letters*, 12(2):145–158.

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., and Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry*, 162:2384–2386.

Simon, H. A. (1969). *The sciences of the artificial*. MIT Press, Cambridge, Mass., 2nd edition.

Smith, L. B. and Thelen, E., editors (1993). *A dynamic systems approach to development: Applications*. Bradford Books series in cognitive psychology. MIT Press, Cambridge, Mass.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psych.*, 43:482–489.

Sporns, O. and Edelman, G. M. (1993). Solving bernstein's problem: A proposal for the development of coordinated movement by selection. *Child Development*, 64(4):960–981.

Srinivasan, M. V., Chahl, J. S., Weber, K., Venkatesh, S., Nagle, M., and Zhang, S. (1999). Robot navigation inspired by principles of insect vision. *Robotics and Autonomous Systems*, 26(2):203–216.

Srinivasan, M. V., Zhang, S., Altwein, M., and Tautz, J. (2000). Honeybee navigation: Nature and calibration of the "odometer". *Science*, 287(5454):851–853.

Steels, L. (1991). Towards a theory of emergent functionality. In Meyer, J.-A. and Wilson, S., editors, *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, pages 451–461, Cambridge, Mass. MIT Press. A Bradford Book.

Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, 7(7):308–312.

Steels, L. (2004). The autotelic principle. In Iida, F., Pfeifer, R., Steels, L., and Kuniyoshi, Y., editors, *Embodied Artificial Intelligence, International Seminar, Dagstuhl Castle, Germany, July 7-11, 2003, Revised Papers*, pages 231–242, Berlin Heidelberg. Springer-Verlag.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: an introduction*. MIT Press, Cambridge, Mass. A Bradford Book.

te Boekhorst, R., Lungarella, M., and Pfeifer, R. (2003). Dimensionality reduction through sensory-motor coordination. In Kaynak, O., Alpaydin, E., Oja, E., and Xu, L., editors, *Proceedings of the Joint International Conference ICANN/ICONIP*, pages 496–503. LNCS 2714.

Thelen, E. and Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Bradford Books series in cognitive psychology. MIT Press, Cambridge, Mass.

Thiélin-Bescond, M. and Beugnon, G. (2005). Vision-independent odometry in the ant Cataglyphis cursor. *Naturwissenschaften*, 92(4):193–197.

Thompson, A. (1996). Silicon evolution. In et al., J. R. K., editor, *Proceedings of Genetic Programming*, pages 444–452, Cambridge, Mass. MIT Press.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. MIT Press, Cambridge, Mass.

von Holst, E. and Mittlestaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, 37:464–476.

Webb, B. (1995). Using robots to model animals: a cricket test. *Robotics and Autonomous Systems*, 16:117–134.

Webb, B. (2001). Can robots make good models of biological behaviour? *Behavioral and Brain Sciences*, 24:1033–1050.

Webb, B. (2002). Robots in invertebrate neuroscience. *Nature*, 417:359–363.

Webb, B. (2004). Neural mechanisms for prediction: do insects have forward models? *Trends in Neurosciences*, 27(5):278–282.

Webb, B. and Scutt, T. (2000). A simple latency-dependent spiking-neuron model of cricket phonotaxis. *Biological Cybernetics*, 82(3):247–269.

Weber, K., Venkatesh, S., and Srinivasan, M. V. (1996). Insect inspired behaviours for the autonomous control of mobile robots. *Proceedings of the 13th International Conference on Pattern Recognition*, 1:156–160.

Wehner, R. (1992). Homing in arthropods. In Papi, F., editor, *Animal Homing*, pages 45–144. Chapman and Hall, London.

Wehner, R. (1994). The polarization-vision project: Championing organismic biology. In Schildberger, K. and Elsner, N., editors, *Neural basis of Behavioural Adaptations*, pages 103–143. Gustav Fischer Verlag, Stuttgart.

Wehner, R. (2003). Desert ant navigation: how miniature brains solve complex tasks. *J. Comp. Physiol. A*, 189(8):579–588.

Wehner, R., Michel, B., and Antonsen, P. (1996). Visual navigation in insects: Coupling of egocentric and geocentric information. *Journal of Experimental Biology*, 199:129–140.

Wehner, R. and Räber, F. (1979). Visual spatial memory in desert ants, Cataglyphis bicolor (Hymenoptera: Formicidae). *Cellular and Molecular Life Sciences*, 35(12):1569–1571.

Wehner, R. and Srinivasan, M. V. (1981). Searching behaviour of desert ants, genus Cataglyphis (Formicidae, Hymenoptera). *J. Comp. Physiol.*, 142:315–338.

Wehner, R. and Srinivasan, M. V. (2003). Path integration in insects. In Jeffery, K., editor, *The neurobiology of spatial behaviour*, pages 9–30. Oxford University Press, Oxford.

Wehner, R. and Wehner, S. (1986). Path integration in desert ants: Approaching a long-standing puzzle in insect navigation. *Monitor Zoologico Italiano*, 20(3):309–31.

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., and Thelen, E. (2001). Autonomous mental development by robots and animals. *Science*, 291(5504):599–600.

Wittlinger, M., Wehner, R., and Wolf, H. (2006). The ant odometer: Stepping on stilts and stumps. *Science*, 312(5782):1965–1967.

Wohlgemuth, S., Ronacher, B., and Wehner, R. (2001). Ant odometry in the third dimension. *Nature*, 411(6839):795–8.

Wolpert, D. M. and Flanagan, J. R. (2001). Motor prediction. *Current Biology*, 11(18):R729–R732.

Wolpert, D. M. and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, 3:1212–1217.

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882.

Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11:1317–1329.

Wolpert, D. M., Miall, R. C., and Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9):338–347.

Yamamoto, T. and Kuniyoshi, Y. (2001). Harnessing the robot's body dynamics: a global dynamics approach. In *Proc. 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 518–525, Maui.

Lebenslauf

Name	Bovet
Vornamen	Simon Igal
Geburtsdatum	4. Mai 1978
Heimatort	Fleurier NE

Ausbildung

1996	Maturité type C, Gymnase cantonal de Neuchâtel
1996 – 2001	Studium in Physik, ETH Zürich
2001	Diplom als Physiker, ETH Zürich
Diplomarbeit	“Implementation and Learning of Computer Games in Neural Networks”
2002 – 2007	Doktorand und Forschungsassistent am Labor für Künstliche Intelligenz, Institut für Informatik, Universität Zürich