Action-Driven Perception –
Neural Architectures Based On Sensorimotor Principles

Dissertation

zur Erlangung des akademischen Grades
Dr. rer. nat.
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Universität Hamburg

eingereicht beim Fach-Promotionsausschuss Informatik von

Dr. med. Jens Philipp Kleesiek

aus Aachen

Hamburg, 2012

Gutachter:	Prof. Dr. Andreas K. Engel
	Prof. Dr. Stefan Wermter
Vorsitzender der Prüfungskommission:	Prof. Dr. Jianwei Zhang
Tag der Disputation:	13. August 2012

Für Anke und Linus.

Abstract

The active nature of perception and the intimate relation of action and cognition has been emphasized in philosophy and cognitive science for a long time. However, most of the current approaches do not consider the fundamental role of action for perception. Inspired by theories rooted in the research field of embodied cognition we have designed artificial neural architectures for the learning of sensorimotor laws. All our models have in common that the agent actually needs to act to perceive. This core principle is exploited for the design of a series of computational studies, including simulations and real-world robot experiments. In a first experiment, a virtual robot learns to navigate towards a target region. For this purpose, it learns sensorimotor laws and visual features simultaneously, using the world as an outside memory. The control laws are trained using a two-layer network consisting of a feature (sensory) layer that feeds into an action (reinforcement learning) layer. The prediction error modulates the learning of both layers. In a second experiment, we introduce a novel bio-inspired neural architecture that combines reinforcement learning and Sigma-Pi neurons. In a simulation we verify that a virtual agent successfully learns to reach for an object while discovering invariant hand-object relations simultaneously. Again, the prediction error of the action layer is used to modulate all the weights in the network. In a third experiment we extend a recurrent architecture with an adaptive learning regime and use this algorithm for an object categorization task with a real humanoid robot. Based on self-organized dynamic multi-modal sensory perceptions, the robot is able to 'feel' different objects and discriminate them with a very low error rate. All these experiments are inspired by the same sensorimotor design principles. Further, they are united by the idea that *actively* acquired sensorimotor knowledge enhances perception and results in goal-directed behavior.

Zusammenfassung

In der Philosophie und in den Kognitionswissenschaften wird schon seit längerer Zeit auf die besonders enge Verknüpfung, die Handlungen und kognitive Prozesse haben, hingewiesen. Leider berücksichtigen die meisten der gegenwärtigen Studien aus dem Bereich der Robotik diesen fundamentalen Einfluss von Handlungen auf die Wahrnehmung nicht. Inspiriert durch Theorien, die ihren Ursprung in einem Forschungsfeld haben, das unter dem Begriff des *Embodiments* zusammengefasst wird, einer These nach der Intelligenz die physikalische Interaktion des Körpers voraussetzt, haben wir verschiedene künstliche neuronale Netzwerkarchitekturen entwickelt, die in der Lage sind, sensomotorische Zusammenhänge zu erlernen. Allen unseren Modellen ist gemein, dass der Agent handeln muss, um überhaupt etwas wahrzunehmen. Dieses Kernprinzip nutzen wir für verschiedene Computerexperimente aus, die Simulationen sowie Studien mit echten Robotern umfassen.

Die erste Studie befasst sich mit der Navigation zu einer Zielregion. Ein virtueller Roboter erlernt sensomotorische Gesetzmäßigkeiten und extrahiert dabei gleichzeitig visuelle Merkmale aus seiner Umwelt. Hierfür ist der Agent mit einem zwei-schichtigen künstlichen neuronalen Netz ausgerüstet, das aus einer sensorischen und einer Handlungs-Schicht besteht. Der Vorhersagefehler der Handlungs-Schicht, realisiert durch verstärkendes Lernen, dient hierbei nicht nur zur Anpassung der Synapsen dieser Schicht, sondern moduliert gleichzeitig auch noch die Synapsen der sensorischen Neuronen.

In einem zweiten Experiment stellen wir eine neu entwickelte bio-inspirierte Netzwerkarchitektur vor, die verstärkendes Lernen mit Sigma-Pi Neuronen verbindet. Es wird in einer Simulation gezeigt, dass ein virtueller Agent mit Hilfe dieser Architektur in der Lage ist, invariante Situationen zu erkennen. Gleichzeitig erlernt er auch noch das erfolgreiche Greifen nach Objekten. Auch in diesem Fall beeinflusst der Vorhersagefehler der Handlungs-Schicht alle synaptischen Gewichte des Netzwerks.

In der dritten Studie erlernt ein echter humanoider Roboter, Bauklötze durch multisensorische Wahrnehmung zu kategorisieren. Zu diesem Zweck haben wir den Algorithmus einer speziellen rekurrenten Netzwerkarchitektur um eine adaptive Lernregel erweitert. Das rekurrente Netz speichert und gruppiert die multisensorischen Eindrücke, die durch die Interaktion mit den Objekten entstehen. Hierdurch ist der Roboter später in der Lage, verschiedene Objekte zu 'erfühlen' und erfolgreich voneinander zu diskriminieren.

Alle drei Studien sind durch die selben sensomotorischen Design-Prinzipien motiviert. Außerdem verbindet sie die Idee, dass *aktiv* erworbene sensomotorische Zusammenhänge die Wahrnehmung erweitern und dadurch zu zielgerichtetem und erfolgreichem Handeln führen können.

Contents

Abbreviations

AI	artificial intelligence
AIP	anterior intraparietal area
API	application programming interface
BPTT	back propagation through time
CTRNN	continuous time recurrent neural network
CV	computer vision
DH	dynamical hypothesis
DST	dynamical systems theory
EC	embodied cognition
fMRI	functional magnetic resonance imaging
FOR	frame of reference
GOFAI	good old-fashioned artificial intelligence
IM	ideomotor theory
LIP	lateral intraparietal area
MSE	mean squared error
NN	neural network
PB	parametric bias
PMA	premotor area
PRR	parietal reach region
RF	receptive field
RL	reinforcement learning
RNN	recurrent neural network
RNNPB	recurrent neural network with parametric bias
RT	reaction time
SLAM	simultaneous localization and mapping
SMCs	sensory motor contingencies
SOM	self-organizing map
STD	standard deviation
TEC	theory of event coding
VE	virtual environment
VIP	ventral intraparietal area
VR	virtual reality

List of Figures

List of Tables

We must perceive in order to move, but we must also move in order to perceive.

James J. Gibson

Chapter 1
Introduction

This thesis is the result of interdisciplinary work in the field of computer science and neuroscience. The main focus of the presented projects is on artificial neural networks that exploit an *action-driven perception* paradigm for the learning of *sensorimotor laws*.

To clarify the meaning of *action-driven* perception and to embed the research theme in the field of embodied cognition (EC), results from philosophy, psychology, neuroscience, cybernetics, computer science and robotics are presented and integrated in Ch. 2. Insights of historic as well as state-of-the-art approaches are critically discussed, pointing out shortcomings and inconsistencies of the current theories. This survey tries to establish a coherent picture of this developing research field, which has yet to define itself.

Establishing a methodology in this way allows one to understand the motivation and inspiration that led to the design and implementation of the presented experiments. Furthermore, this knowledge is important for the critical discussion of the computational models (Ch. 6).

1.1 Sensorimotor Laws

Within this thesis the term *sensorimotor laws* reflects the intertwined relationship of action and perception. Especially, those regularities that are acquired during learning, connecting motor commands to sensory impression in a lawful manner (and *vice versa*), are summarized with this expression. The usage of the term *law* was chosen in homage to Helmholtz (1867), who also used this expression to characterize the learned relationship between actions and perceptions (*cf.* Ch. 2.2.1). While *enacting* our world we can rely on this relationship and use it for goal-directed behaviors. In the first two experiments (Ch. 3 and 4) we[1] examine two-layer neural architectures that combine reinforcement learning (RL) with sensory feature extraction. During learning, actions shape the receptive fields (RFs) of the sensory layer, establishing the *sensorimotor laws* that allow successful mastery of the given task. Later on, confronted with a sensory stimulus the agent can immediately link it to an appropriate action. In the third experiment (Ch. 5) a real-world robot interacts with objects, thereby eliciting multi-modal sensory impressions. These sensory sequences are learned by a recurrent neural network (RNN) that self-organizes them into clusters. In this experiment, those clusters reflect the *sensorimotor laws*, which can be called on during object categorization.

1.2 Research Agenda and Sensorimotor Principles

The idea of learning and exploiting *sensorimotor laws* for goal-directed behavior unites all the experiments that will be presented in the following chapters. In general, the purpose of our three studies is to capture the intertwined relationship of action and perception while pursuing the research question: **is it possible to rely on the same sensorimotor design principles for the development of different artificial neural architectures and experiments?** To answer this question, differing connectionist neural architectures have been devised or refined and their potential has been evaluated in simulations and real-world robot experiments.

The scope of this thesis is not to prove that actions are fundamental for perception. This fact is taken for granted and theories and experiments supporting this notion are presented in Ch. 2. Instead, we demonstrate that paying attention to a common set of sensorimotor design principles during the development of artificial neural architectures, as well as during the planning of robotic experiments, is indeed sufficient to obtain robots that are able to master a given task successfully. Primarily, the following sensorimotor principles have been employed to realize the action-driven perception paradigm of our experiments:

- **perceiving is a way of acting** a.k.a.
  perception as a sensorimotor experience,
- **open channel perception** a.k.a.
  the world serves as an outside memory,
- **information self-structuring**.

Perceiving is a way of acting stresses the importance of organism-world interactions for perception. The theme *the world serves as an outside memory* describes a (constant) interaction between the world and the neural system that contributes to the cognitive process as a whole. *Information self-structuring* summarizes the

[1]Throughout the thesis the ‘scientific’ we is used, even if personal opinions and ideas are expressed.

principle that actions enable an agent to structure the (sensory) information, which is readily available in its environment. For further details and additional sensorimotor principles please refer to Ch. 2.

1.3 Scientific Contributions

The main scientific contributions of this thesis comprise the following achievements. The sensorimotor design principles described in the previous section can readily be found in all the experiments. Hence, the agent actually needs to act to perceive, mimicking the principle *perceiving is a way of acting*. First, we introduce an innovative navigation paradigm that is independent of a world model. Instead, *the world itself serves as an outside memory*. Second, we propose a novel bio-inspired neural architecture that combines reinforcement learning and Sigma-Pi neurons. This allows the given reaching task to be successfully mastered and invariant hand-object relations simultaneously to be discovered. Again, *the world itself serves as an outside memory*. Third, we extend a recurrent architecture with an adaptive learning regime, leading to a significantly reduced training time. This novel training method is used for an active object categorization task of a humanoid robot. Based on *information self-structuring* of multi-modal sensory perceptions, the robot is able to 'feel' different objects and discriminate them with a very low error rate. In addition, several future experiments are suggested that can be conducted based on the theoretical and methodological framework established within this thesis. Most parts of the thesis (text and figures) have been published. Please see appendix B for a complete list of the publications.

1.4 Structure of the Thesis

Ch. 2 gives an introduction to the fascinating research field of embodied cognition and points out thoughts and ideas that influenced the design of the neural architectures and experiments of this thesis. In Ch. 3 – 5 results of the computational studies, including simulations and real-world robot experiments, are described. The experiments comprise an entire sequence of related tasks. In the first experiment, a virtual robot learns to navigate towards a target region (Ch. 3). This is followed by a reaching study (Ch. 4) and a dynamic object recognition task where a real humanoid robot moves objects up and down and rotates them back and forth, while holding them in its hand (Ch. 5). In Ch. 6 the outcome of the individual experiments, as well as their joint relevance for the action-driven perception paradigm within the research field of embodied cognition, is discussed.

In addition to being a sequence of related tasks, the conducted experiments are also related by (at least) two additional meta levels. As stated above, they are all inspired and governed by the same sensorimotor design principles. Further, they also represent the evolutionary process of finding suitable artificial neural architectures and appropriate experiments for *action-driven* learning based on *sensorimotor principles*. In the first study (Ch. 3), the focus lay on the methodology of the artificial neural architecture, i.e. becoming acquainted with these types of two-layer neural networks. This paved the way for the development of the novel bio-inspired architecture presented in Ch. 4. Due to scaling issues of the feedforward networks employed in the first two studies and a known superior generalization potential of recurrent neural networks, we moved to these types of connectionist architectures for the final experiment (Ch. 5). Again, the knowledge gained previously helped the design of the experiment and the improvement of the learning algorithm.

If you desire to see,
learn how to act.
HEINZ VON FOERSTER

Chapter 2 Action-Driven Perception

2.1 Introduction

What is *action-driven perception*? The aim of this chapter is to buttress the fundamental role of action for perception and to introduce the concept of sensorimotor laws. For this purpose, the notions are embedded in the framework of *embodied cognition*.

Consider the sayings 'taking a perspective on a problem' or 'grasping a concept'. Taking these utterances literally immediately illustrates the entanglement of action and perception, their *enactive* nature. However, there is more to it. To explain the role of action and perception for cognition[1], we have to dig deeper into this "extremely active and diverse research area" (Anderson, 2003).

Trying to identify a common ground, we follow the categorization of Lakoff and Johnson (1999). They distinguish between first and second-generation cognitive science, entitling them *disembodied* and *embodied* mind, respectively. Like in Cartesian dualism, disembodied cognitive science maintains a mind-body dichotomy at its core, whereas second-generation cognitive science sees a necessity for an intimate interaction of the mind and the body.

Despite the seemingly trivial influence of action on perception, the conception of this relation in general, and of the embodiment paradigm in particular, is –within and across disciplines– far from uniform. Furthermore, it is an ongoing debate if EC resembles a "Copernican revolution" (Lindblom and Ziemke, 2006) or if it is rather a modest shift in emphasis. Also Froese (2010) is aware of this "ambiguity about whether the enactive paradigm entails another minor reformation or a major revolution of the cognitive sciences." As a potential explanation he spots

[1] The mental *action* or process of acquiring knowledge and understanding through thought, experience, and the *senses* (Oxford dictionary, `http://oxforddictionaries.com`).

the parallel and apparent unnoticed development of this paradigm in different disciplines, like for instance cybernetics and cognitive science.

Looking at the historic roots reveals that for a long time attempts have been made to "anchor research to the real world poles of sensing and acting"(Clark, 1995), resulting in the appearance of "world and body [...] as significant players in the cognitive arena"(ibid.). This, of course, raises the question "[h]ow could we ever have forgotten them?" (ibid.).

2.2 Historic Overview

The 18th century scientific revolution, which was triggered by Cartesian dualism, paved the way for scientific exploration and technological domination. Nevertheless, at the same time it established a mysterious relationship between observer and the observed (Froese, 2010), which has kept scientists busy until today.

Reducing the body to a mere *representation* in the mind is an inherent procedure of the empirical sciences, especially when considering the fact that experiments are by definition conducted in controlled environments. However, this reductionism does not suffice to explain e.g. *qualia*[2]. Different scientific traditions try to overcome this "explanatory gap" (Levine, 1983) with the tools they have at hand. A current trend that can be witnessed in the cognitive sciences, also referred to as the "pragmatic turn" (Engel, 2010, *cf.* Ch. 2.4.5), tries to resolve this "hard problem" (Chalmers, 1995) via an embodied approach.

In this section different schools of thought will be presented that are, according to our view, important forerunners for the contemporary EC paradigm. This enumeration, however, is not intended to be a comprehensive historical overview, which would be far beyond the scope of this thesis.

2.2.1 Von Helmholtz

Hermann von Helmholtz (1821 – 1894) was a *polymath* contributing to various fields of research, including physics and physiology. His far-reaching achievements are realized in a book review of *Helmholtz on Perception: Its Physiology and Development* (Warren and Warren, 1968), stating "that understanding of a surprising number of problems discussed by Helmholtz has been little advanced in the intervening decades" (Mueller, 1968). For the present discussion exactly those epistemological attainments that relate perception ("*Sinneswahrnehmungen*") to action ("*Bewegung*") by the means of sensorimotor[3] laws ("*Gesetze*") are of primary importance[4].

[2]Subjective conscious experiences; the *what it's like*. For a detailed account see Tye (2009).
[3]Von Helmholtz does not use the expression *sensorimotor*.
[4]Relevant excerpts of von Helmholtz' book *Handbuch der physiologischen Optik*, written in German, can be found in Appendix C. These are summarized in the following paragraphs. Please note that at the time of publication (1867) no unified grammatical and spelling norms existed. The quotations are taken 'as is', resembling the original typesetting.

Of utter significance for cognition, he notes, are voluntary movements, explicitly contrasting them to merely passive observations.[5] Sensation and action are linked by a causal nexus, a law of causation, which does not develop if we do not practice it and constantly exercise our skills. This relationship is dependent on the physical properties, the nature ("*der Natur*") of our body, the world and the interplay of both. He suggests that our perceptions ("*Vorstellungen*") have to be 'hands-on' ("*praktisch*") and he realizes that they, instead of being images ("*Abbilder*"), must be signs[6] ("*Zeichen*"). An image, he continues, is supposed to share some kind of equality with its original source. A sign, on the other hand, does not have to meet this requirement. It only has to be constant, i.e. the same object, given identical conditions, has to elicit the same sign. Thus, the relationship ("*Geschehen*") is captured by a law, which, once learned, helps us to identify the object. The dynamic process of sensation is further stressed by the following quote taken from *The Facts of Perception* (von Helmholtz and Kahl, 1971 [1878]):

> "Let us assume that the man at first finds himself to be just one object in a region of stationary objects. As long as he initiates no motor impulses, his sensations will remain unchanged. However, if he makes some movement (if he moves his eyes or his hands, for example, or moves forward), his sensations will change. And if he returns (in memory or by another movement) to his initial state, all his sensations will again be the same as they were earlier. [. . .]
> It is easy to see that by moving our fingers over an object, we can learn the sequences in which impressions of it present themselves and that these sequences are unchanging, regardless which finger we use. It is thus that our knowledge of the spatial arrangement of objects is attained. Judgments concerning their size result from observations of the congruence of our hand with parts or points of an object's surface, or from the congruence of the retina with parts or points of the retinal image. A strange consequence, characteristic of the ideas in the minds of individuals with at least some experience, follows from the fact that the perceived spatial ordering of things originates in the sequences in which the qualities of sensations are presented by our moving sense organs: the objects in the space around us appear to possess the

[5] He conjectures that if objects were only moved passively, by an external force, through our field of vision, we would not be able to learn 'seeing' (*cf.* experiments by Held and Hein, Ch. 2.4.1).

[6] *Zeichen* could also be translated with symbol. But it does not have the meaning of symbolic representation here, it is rather a signal from the outside world, which is able to capture the laws of the real world ("die Abbildung der Gesetzmäßsigkeit in den Vorgängen der wirklichen Welt"). Peirce (*cf.* Pragmatism, Ch. 2.2.2) "define[s] a sign as anything which is so determined by something else, called its Object, and so determines an effect upon a person, which effect [he] call[s] its interpretant, that the latter is thereby mediately determined by the former" (1998). Hence, "[a] sign [. . .] is something which stands to somebody for something in some respect or capacity" (Peirce, 1950 [1897]). For a further distinction of signs please also see Millikan (2004).

> qualities of our sensations. They appear to be red or green, cold or warm, to have an odour or a taste, and so on. Yet these qualities of sensations belong only to our nervous system and do not extend at all into the space around us. Even when we know this, however, the illusion does not cease, for it is the primary and fundamental truth. The illusion is quite simply the sensations which are given to us in spatial order to begin with."

Once we have learnt the sensorimotor laws, we can use actions ("*Handlungen*") to trigger specific sensations. For instance, we are able to guide our gaze purposefully over an object that we already know. Thus, von Helmholtz concludes that the holistic perception of an object's shape is identical to the sequence of motor commands needed to examine it with our eyes[7], i.e. its sensorimotor laws. Furthermore, the lively imagination of these laws sufficiently constitutes our perception of objects[8].

But how do we acquire such knowledge? For this purpose, von Helmholtz vividly describes the interaction of infants with (toy) objects (*cf.* development of sensorimotor skills, Ch. 2.4.1). They touch the toys, look at them for hours, stick them into their mouth, pound them onto the ground, repeating this procedure day by day, over and over again. Hence, von Helmholtz infers that the sensorimotor skills have to be exercised, they are not connate.

Several concepts proposed by von Helmholtz will *re*appear in the following sections and, as it will be seen, are still state-of-the-art in the research field of embodied cognition. Needless to say, they had a significant impact on the performed experiments, e.g. the dynamic nature of action-driven perception, voluntary movements and lifelong exercising of the obtained sensorimotor laws can readily be found (*cf.* experiments, Ch. 3 – 5).

2.2.2 Pragmatism

Pragmatism as a philosophical notion originated in the United States around the 1870's. The most important advocates of the 'classical' pragmatists were Charles Sanders Peirce (1839 – 1914), William James (1842 – 1910) and John Dewey (1859 – 1952). In essence, pragmatism links practice and theory. Theoretical knowledge obtained from experiments is in turn used to redefine the experiments and to clarify the practical consequences of the hypothesis. This implies that if a theory is making correct predictions, i.e. it is practical, it ought to be true. On the other hand, impractical ideas, which do not reflect observations, have to be rejected.

All three of the traditional protagonists have emphasized the active nature of perception and the intertwined relationship of action and cognition. John Dewey

[7] "Hier bewährt sich also in der That die Gesammtauffassung der Körperform gleich als die Regel für die Vorstellung, nach welcher man die beiden Blicklinien zu führen hat, [. . .]"(von Helmholtz and König, 1896).

[8] "[Der Begriff eines Objekts wird] nur durch die lebendige Vorstellug des Gesetzes, nach dem seine perspektivischen Bilder einander folgen, zusammengehalten"(ibid.).

aptly summarizes this relationship and points out the unconditional necessity of action for perception:

> "Upon analysis, we find that we begin not with a sensory stimulus. but with a sensori-motor coordination, the optical-ocular, and that in a certain sense it is the movement which is primary, and the sensation which is secondary, the movement of body, head and eye muscles determining the quality of what is experienced. In other words, the real beginning is with the act of seeing; it is looking, and not a sensation of light. The sensory quale gives the value of the act, just as the movement furnishes its mechanism and control, but both sensation and movement lie inside, not outside the act." (Dewey, 1896)[9]

As noted above, these 'action-guided' concepts have been rediscovered in contemporary research projects summarized under the term embodied cognition, including disciplines like robotics, psychology and more recently also neuroscience (Engel, 2010). The contributions of William James to the *ideomotor theory* (IM) are discussed in the next section.

2.2.3 Ideomotor Theory

Ideomotor[10] theory is a framework for action planning, suggesting that actions are represented by their perceptual effects. Usage of the term in the literature is not uniform. Most often, induced actions that are triggered endogenously or exogenously by perceptual phenomena are meant (Shin *et al.*, 2010). This perceptual impact on actions is often expressed by the term ideomotor *action*. Nevertheless, the opposite direction, ideomotor *perception*, where a perception is influenced by an action is conceivable as well.

Historically, IM originated in the 19$^{\text{th}}$ century from two different roots[11], a German and a British one. Carpenter (1813 – 1885), belonging to the British root, originally coined the term 'ideomotor' (Carpenter, 1852). Accompanied by Laycock (1812 – 1876), they sought to explain ideomotor phenomena by means of cerebral reflex actions. The older German root was pursued by Herbart (1776 – 1841), Lotze (1817 – 1881) and Harleß (1820 – 1862). They regarded the ideomotor principle as a core mechanism for all human intentional behaviour. Both roots coalesced in James' *magnum opus The Principles of Psychology* (1890).

During learning, actions are linked to their perceptual effects. According to James, this association can also be reversed (James, 1950 [1890], p. 526). A known (previously learned) sensation may evoke a corresponding action. This makes IM akin to the notion of forward and reverse models of computational motor control (Wolpert and Ghahramani, 2004). Indeed, ideomotor theory supports the sensory

[9] The article can be obtained from `http://psychclassics.yorku.ca/Dewey/reflex.htm`.
[10] Not reflex but motivated by an idea (Merriam Webster, `http://www.merriam-webster.com`).
[11] For a detailed review of the historical roots please refer to Stock and Stock (2004).

prediction of a performed or imagined action, as well as the selection of an action leading to a certain anticipated perceptual experience.

The artificial neural architectures presented in Ch. 3 and 4 contain similar properties. The combination of reinforcement learning and a sensory layer allows for triggering of suitable actions based on sensations, once the relations of actions and perceptions have been learned, i.e. the *sensorimotor laws* relevant for the given task have been acquired.

IM fell into oblivion with the advent of *behaviorism* and has only recently reappeared. Its contemporary successor, the theory of event coding (TEC), will be presented briefly in Ch. 2.4.6 .

2.2.4 Phenomenology

Phenomenology was founded in the early 20$^{\text{th}}$ century by the mathematician and philosopher Edmund Husserl (1859 – 1938). In a nutshell, phenomenology studies the structures of conscious experience from a first-person perspective. In this process, *intentionality* is the central structure of an experience. It is defined by Husserl as directedness of experience towards things in the world, the property of consciousness that it is a consciousness of or about something. This notion can also be inverted. That is to say, phenomenology can be used to identify conditions of conscious experiences, like perception, (embodied) action, bodily awareness, thought, memory, imagination, emotion, desire, social activity, *etc.* , and to use them to give experiences their intentionality. Husserl intended phenomenology to be a method of philosophical inquiry, which replaces the rationalist bias dominating Western thought since Plato with reflective attentiveness, thereby revealing the "lived experience" of the individuals (Husserl, 1970).

Two of Husserl's students, Martin Heidegger (1889 – 1976) and Maurice Merleau-Ponty (1908 – 1961), are important for the present discussion, which emphasizes that action is mandatory for perception and cognition.

Heidegger (1975; 1977) developed the concept of 'being-in-the-world' ("*In-der-Welt-Sein*") to abolish Cartesian dualism, the dichotomy of mind and body, and to establish a foundation for the intentionality concept proposed by his mentor. He realized that a cognitive agent has to be *situated* in its world and that a practical understanding of the world arises by virtue of its own body, i.e. it needs to be *embodied*. The embodied cognitive agent and the world are merged to form a holistic structure comprising all components of the situation, leading to a concept denoted as *Bewandtnisganzheit* (Heidegger, 1975).

Similarly, Merleau-Ponty comprehends the perceptual world as inseparably intertwined with the body of the cognitive agent:

> "Visible and mobile, my body is a thing among things; it's caught in the fabric of the world, and its cohesion is that of a thing. But, because it moves itself and sees, it holds things in a circle around itself. Things are an annex or prolongation of itself; they are incrusted into

> its flesh, they are part of its full definition; the world is made of the same stuff as the body" (Merleau-Ponty, 1974, p. 284)

This coalescence of body and world, Merleau-Ponty suggests, leads to cognition and self-awareness:

> "The world is [...] the natural setting of, and field for, all my thoughts and all my explicit perceptions. Truth does not 'inhabit' only 'the inner man', or more accurately, there is no inner man, man is in the world, and only in the world does he know himself." (Merleau-Ponty, 1962, Preface)

Further, this view of embodied cognition is grounded in sensorimotor activity. He emphasizes the fundamental role of action for perception; without action there is no perception, there is no cognition:

> "Since all the movements of the organism are always conditioned by external influences, one can, if one wishes, readily treat behaviour as an effect of the milieu. But in the same way, since all the stimulations which the organism receives have in turn been possible only by its preceding movements which have culminated in exposing the receptor organ to external influences, one could also say that behavior is the first cause of all stimulations. Thus the form of the excitant is created by the organism itself." (Merleau-Ponty, 1963, p. 13)

Due to the striking resemblance to the thoughts of von Helmholtz (*cf.* Ch. 2.2.1) and Dewey (*cf.* Ch. 2.2.2), it is not surprising that the experiments presented in Ch. 3 – 5 are also reminiscent of the thoughts presented in this section.

2.2.5 Cybernetics

The research field of cybernetics was founded by Norbert Wiener (1894 – 1964) who published a book with the very same title in the 1950s[12]. Originally, as defined in Wiener's book, cybernetics studied control and communication in animals and machines. Since then, the research field has been expanded to an interdisciplinary science investigating the structure of information flow in regulatory systems in general, ranging from "stars to brains"[13].

Like the cognitive sciences, this research field can be subdivided into two phases. Traditional first-generation cybernetics was strongly influenced by the work of William Ross Ashby (1903 – 1972). According to Froese (2010), he can be regarded as the 'culprit' who triggered a series of historical developments that caused the transformation to second-order cybernetics. Von Foerster (1911 – 2002), being one of the leading scientists of the second phase, realized that

[12] For a detailed review of the early American cybernetics tradition please refer to Dupuy (2009).
[13] This expression is adopted from `http://en.wikipedia.org/wiki/Cybernetics`.

> “a brain is required to write a theory of a brain. From this follows that a theory of the brain, that has any aspirations for completeness, has to account for the writing of this theory. And even more fascinating, the writer of this theory has to account for her or himself. Translated into the domain of cybernetics; the cybernetician, by entering his own domain, has to account for his or her own activity. Cybernetics then becomes cybernetics of cybernetics, or second-order cybernetics.”(von Foerster, 2003)

Hence, the researcher cannot find out how a system, e.g. the brain, works from the outside. He will always be affected and, more importantly, affect the functioning of the system himself (*cf. anthropomorphic bias*, Ch. 2.4.9).

Second-order cybernetics and the distinct scientific traditions *computationalism* and *connectionism*, which developed in parallel, jointly culminated in the enactive paradigm[14] (Froese, 2010).

Aware of the importance of actions for perception, von Foerster (1984) stated that “[i]f you desire to see, learn how to act”. This proclamation can be seen as the red thread of this thesis (*cf.* experiments, Ch. 3 – 5).

2.2.6 Symbol Grounding Problem

According to *computationalism* and *Good Old-Fashioned Artificial Intelligence*[15] (GOFAI), cognition (intelligence) is merely a matter of abstract symbol manipulation. For this purpose, a set of symbols and syntactic rules that relate them to each other is defined (Newell and Simon, 1976). This is no problem as long as a human interpreter is involved who is capable of relating the result of the symbol manipulation to the outside world, i.e. the symbol processing is grounded in the experience of the human interpreter. But how does meaning arise? This cannot be mediated by an external observer because this would lead to an infinite regress, just like the Rosetta stone would never have been deciphered if none of the symbols (words) on it had been known.

The philosophical debate on how concepts and ideas obtain their meaning and how they are grounded was fueled by Searle’s (1980) *Chinese room argument*[16]. Briefly, in this *Gedankenexperiment* a solely English speaking person is located in a sealed room, surrounded by Chinese speakers. The only contact to the outside world is mediated via a slot through which papers containing symbols, presumably Chinese, are handed to the inhabitant of the room. An instruction manual, written in English, tells him which symbols to write on the paper in response to the incoming messages, not knowing the meaning behind them. Nevertheless, to the Chinese speakers outside, the ‘room’ seems to be able to understand and speak Chinese. Interpreting the thought experiment suggests that a system merely

[14] Froese (2009) considers Hume (1711 – 1776) as yet another forerunner of the enactive paradigm.

[15] This term was coined by the American philosopher John Haugeland (1985).

[16] For a recent debate on a potential logical hole in this argument, please see (Shaffer, 2009; Nute, 2011).

operating on syntactic processes, i.e. algorithms (recipes for transforming input into output), cannot acquire any meaning or intentionality.

So, how can we design (artificial) cognitive agents, for instance robots, that are autonomous, in such a way that they are aware of the meaning of things and therefore independent of an external human interpreter? Harnad (1990) argues that this problem can be solved if the symbols are grounded in the sensorimotor system. In this way, internal manipulations, as well as sensorimotor activities, are constrained by the same laws.

2.2.7 Gibson's Ecological Theory of Perception

James Jerome Gibson (1904 – 1979), an American psychologist, was the founder of ecological psychology. According to his account, perception is an active process that requires an organism to move around in its environment. He realized that, in contrast to passive observations, motion provides a much deeper source of information. "The changes come from the locomotion, and the nonchanges come from the rigid layout of environmental surfaces"(Gibson, 1979, p. 73). In addition, sensory information typically does not come to the organism by itself, but instead it must be actively acquired. This is achieved by the perceptual system in conjunction with the whole body.

> "Each perceptual system orients itself in appropriate ways for the pickup of environmental information, and depends on the general orienting system of the whole body. Head movements, ear movements and eye movements are part and parcel of the perceptual system they serve. [...] They serve to explore the information available in sound, mechanical contact, chemical contact and light."(Gibson, 1966)

Consequently, "we must perceive in order to move, but we must also move in order to perceive"(Gibson, 1979, p. 223). Gibson (1966) further points out that the stimulation, which is received by the organism, is structured. The *ambient optic array*, i.e. the light that converges on a point of observation, has a structure due to diffusion and reflections caused by the surrounding environmental surfaces. This information can readily be used to discover *invariances* in the visual stimulation, namely those features that stay constant during movement-induced transformations. In the opinion of Gibson, learning these invariances in combination with their causing counterparts of the body and the environment contributes, at least partially, to the understanding of vision.

Due to the *richness* of information included in visual stimulation *per se*, Gibson believed, it does not need to be processed any further. Alternatively, he takes up an idea previously proposed by Donald Hebb (1904 – 1985) in his famous book *The Organization of Behavior: A Neuropsychological Theory* (1949). Hebb proposed that the brain resonates or reverberates to stimulation[17]. In a similar manner, Gibson concludes:

[17] The idea of reverberation is also taken up by Humphrey (2006).

> "Instead of postulating that the brain constructs information from the input of a sensory nerve, we can suppose that the centers of the nervous system, including the brain, resonate to information. [...] The 'resonating' or 'tuning' of a system suggests the analogy of a radio receiver. This model is inadequate because there would have to be a little man to twiddle the knobs. A perceiver is a *self-tuning* system. What makes it resonating to the interesting broadcasts that are available instead of to all the trash that fills the air? The answer might be that the pickup of information is *reinforcing*." (Gibson, 1966)

The neural architectures presented within this thesis (*cf.* Ch. 3 – 5) can be interpreted in the light of these thoughts. The first two experiments use reinforcement learning to 'carve out' invariant perceptual stimuli. The third experiment aims at action-driven object perception containing a recurrent neural network with parametric bias (PB) units at its core. The values of the PB units emerge unsupervised ('self-tuned') and can be regarded as fixed-points of this dynamical system that 'resonate' to object-specific perceptual stimuli.

Affordances

Within his framework of ecological psychology, Gibson is especially well known for his theory of *affordances*[18] (1977; 1979). "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill"(ibid., p. 127). Thus, affordances are defined as 'action possibilities' that are available in the environment. They are dependent on the (evolutionary) historical interaction and effected by the needs and properties of an organism. As a result, the recognition of affordances influences how an organism perceives its world.

Previously, in the early twentieth century, Jakob von Uexküll (1980 [1920]) had already suggested a comparable concept, according to which objects have a functional coloring (*funktionale Tönung*). Additionally, Gibson was aware of the *demanding character* of an object as introduced by Koffka (1935) in his *Principles of Gestalt Psychology*. However, he pointed out a crucial difference.

> "The affordance of something does not change as the need of the observer changes. An affordance is not bestowed upon an object by a need of an observer and his act of perceiving it. The object offers what it does because it is what it is." (Gibson, 1979, pp. 138-139)

This quote emphasizes again the core matter of Gibson's theory, the inseparability of action and perception, clearly qualifying it as an antecedent of the embodied cognition framework.

There are several studies in the computer vision (CV) and robotics community that make use of the affordance concept. For instance, Stark *et al.* (2008) proposed

[18] The noun was made up by Gibson, derived from the verb *to afford*, meaning 'to make available' and 'to provide naturally or inevitably' (`http://www.merriam-webster.com`).

using *affordance cues*, i.e. distinct visual features that suggest performing an action in a specific way, to facilitate the detection of (novel) functional objects categories. In another experiment, Ugur and Sahin (2010) presented a mobile robot equipped with range sensing ability that, by interacting with its environment, is able to learn to perceive a specific (traversability) affordance. Weiller *et al.* (2010) used a robot to model adaptive goal-directed navigational behavior. It learned the reflexive and action-based affordances of its environment in an unsupervised way, based on acquired knowledge that relates motor actions to sensory outcome.

2.3 *Disembodied* Cognitive Science

The intellectual origins of classical first-generation cognitive science, or the *disembodied* mind (Lakoff and Johnson, 1999), date back to the mid-1950s. The still valid core suppositions that characterize this classical cognitivist thinking were influenced by the work of Newell and Simon (1972), Fodor (1981) and others. For a comprehensive historical review, please refer to Boden (2006).

The central hypothesis of this very interdisciplinary field assumes that cognition can be understood best as computation over mental representational structures[19]. Subjects are conceived as passive recipients of stimulation[20], which is algorithmically processed in their brains. All 'actions' are located within the brain – the cause of the inputs and the effect of the outputs on the world are irrelevant for the computational process and, thus, also irrelevant for the understanding of cognition. In that sense, cognition, like computational processes, can be regarded as (methodologically) solipsistic (Fodor, 1980), i.e. mental states solely depend on the input and other internal states without the need to take the physical world in which the organism is embedded into account. Holyoak (1999) shares this opinion and states that "[t]he central focus of psychology concerns the information processing that intervenes between sensory inputs and motoric outputs".

On the other hand, as a proponent of *externalism*, Hurley questions this view by summarizing its core assertion.

> "If perception is input from the world to the mind and action is output from the mind to the world, then the mind as distinct from the world is what the input is to and the output is from. So, despite the web of causal relations between organisms and environments, we suppose the mind must be in a separate place, within some boundary that sets it apart from the world." (Hurley, 1998, pp. 1 – 2)

Classical cognitive science, as described above, has many proponents, see e.g. Thagard (2005; 2010), and still coexists with the *embodied* second-generation cognitive science that is presented in the remainder of this chapter.

[19] While there is much agreement within traditional cognitive science on the computational nature of cognition, the exact composition of the representational structures is heavily debated.

[20] This is actually the case for many carefully controlled studies in psychology and neuroscience.

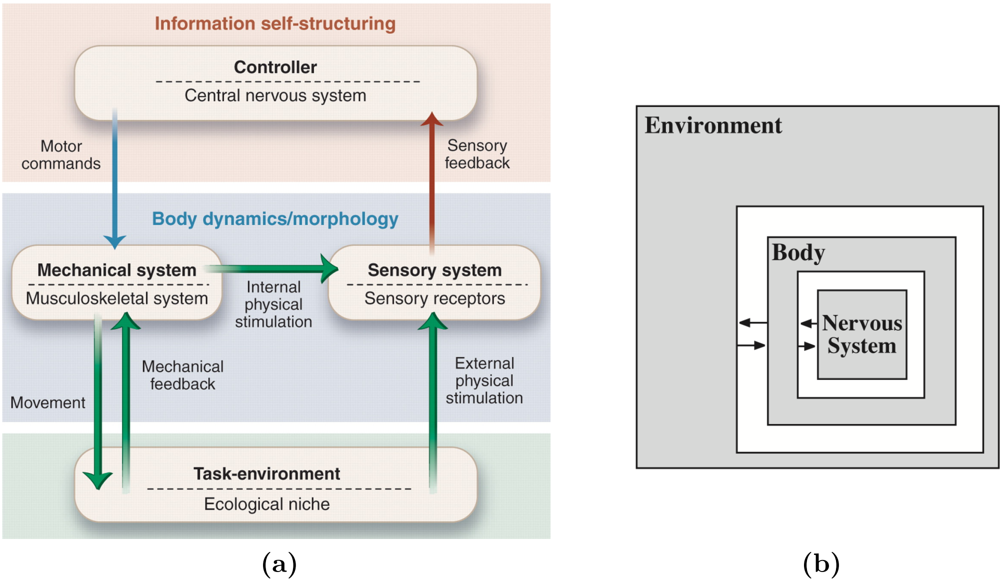


Figure 2.1: Interaction of the brain, the body and the world. (a) Embodiment involves the interplay of the brain and the environment, mediated by the body. Drawing taken from Pfeifer *et al.* (2007). (b) Embedded view of a situated and embodied agent proposed by Beer (2003). All subsystems are in constant dynamic interaction.

2.4 *Embodied* Cognitive Science

Embodied cognitive science[21] postulates that the body is not a mere container for the brain. Instead, it is an integral part and cognition arises from "a dynamic dance in which body, perception and world guide each other's steps" (Shapiro, 2011, p. 61). Rather than being a computer that simply processes information, the brain is now conceived as a controller, embedded in the body (embodied) and the environment (situated), guiding movements (enacting) that lead to information (Fig. 2.1). In essence, cognitive processes are grounded in their sensorimotor experiences (Lakoff and Johnson, 1980; Feldman and Narayanan, 2004; Barsalou, 2008).

Unfortunately, the term *embodiment* is not defined uniformly in the literature. Rohrer (2007) describes a dozen different uses and also Shapiro notes incongruities:

> "Claims about the meaning of embodiment [. . .] are far from uniform in the commitments they entail. More troubling still is that the claims often step far beyond the evidence or argument made in their support." (Shapiro, 2011)

To get a deeper understanding of the multiple facets of embodied cognition and its fundamental role for action-driven perception, experimental studies and different theories from various disciplines, along with their interpretations, will be reviewed.

[21] Whereas in principle, embodied cognitive science can be distinguished from both other forms of situated cognition (Clark, 2008; Rupert, 2009; Shapiro, 2011), i.e. embedded and extended cognition, these views are considered jointly in the present discussion.

2.4.1 Experimental Studies

Various experimental studies in psychology, neuroscience and related fields exist that support the notion of embodied cognition. Here we will focus on insights from infant development, the concept of efference copy and two classical studies that emphasize the fundamental role of action for perception.

Infant Development

The neuroanatomic pathways linking action and perception have recently been reviewed by Guillery (2005). During early post-natal development, dorsal root afferents feed into lower premotor and motor centers of the brainstem and may already innervate the thalamus as well. In contrast, thalamocortical connections are still immature at this time and, thus, are of subordinate importance for this first stage in the development of sensorimotor contingencies[22]. At this stage the infant learns that, for instance, certain hand movements produce particular proprioceptive, tactile, and visual responses. As development proceeds, more and more primary cortical areas become involved. The maturation of the intrinsic circuitry is thereby mediated by the previously established lower sensorimotor contingencies. Subsequently, higher cortical components become part of the circuitry, now comprising learned associations between different sensory and motor areas. The cortex has an impact on the thalamus and also sends an *efference copy* (see below) to lower brainstem areas. During this development of the sensorimotor contingencies, the higher order corticothalamic circuits maintain their plasticity for a longer time period than the 'lower' first order circuits do.

The psychologist Jean Piaget (1896 – 1980) emphasized the role of sensorimotor activity for the cognitive development, which can be subdivided into a number of developmental stages. Like von Helmholtz (*cf.* Ch. 2.2.1), he described that young infants repeatedly perform the same actions (Piaget, 1952). It has also been observed that newborns, while awake, spend up to 20 % of their time touching their face (Korner and Kraemer, 1972). This behavior, in analogy to vocal babbling, has been termed motor or body babbling (Meltzoff *et al.*, 1997) or visual-proprioceptive calibration (Rochat and Hespos, 1997; Rochat, 1998). Through babbling, intermodal redundancies, temporal contingencies and spatial congruences can be acquired. It also allows self to be discriminated from external (environmental) stimulation and movement (Rochat and Hespos, 1997; Rochat, 1998). This is possible because infants do not only passively observe, but actively bring in their motor apparatus to obtain a body representation and to establish a relation between actions and sensory consequences (Bushnell and Boudreau, 1993). While their sensorimotor skills mature, their action skills develop in parallel, enhancing jointly their cognitive capacities (Rosenbaum *et al.*, 2001).

[22]Guillery adopts this expression from O'Regan and Noë (2001, *cf.* Ch. 2.4.4), but uses it in a much narrower sense than intended by the original authors. Namely, to denote "perception [that is] closely related to activity in the motor pathways" (Guillery, 2005).

Efference Copy

Together with the acquisition of a body representation, infants develop a sense of agency, i.e. (in this context) knowing that they are responsible for the generation of an action[23]. This sense of agency becomes impaired when an organism is confronted with a sensorimotor mismatch, becoming evident if the predicted sensory feedback is not in agreement with the actual sensory feedback. Research on the underlying neural mechanisms thus deals with the question: how can an organism distinguish sensory changes resulting from its own movements from those which are not caused by its own actions?

It was probably Descartes who first documented the fact that passive movements of the eye result in the impression of the world 'moving', whereas voluntary movements do not (Grüsser, 1995). Purkyně and von Helmholtz suggested that this might be due to an internally copied motor command interfering with the sensory input; the knowledge of the voluntary eye movement already predicts the upcoming sensory change. This idea was then captured by von Uexküll and Mach in the form of feedback diagrams (ibid.). The underlying concept was eventually formalized, independently and simultaneously, by von Holst and Mittelstaedt (1950) and Sperry (1950). Von Holst and Mittelstaedt denoted it as *efference copy*, a copy of the efferent motor command, that is sent to the sensory system. Comparing it to the reafference signal, i.e. the real sensory feedback, allows ambiguities present in sensory information to be resolved that in turn can be exploited to control behavior. On the other hand, Sperry focused more on the anticipatory prediction of input, denoting his concept *corollary discharge*. These motor-to-sensory signals may influence the sensory processing stream at various levels[24].

The fundamental question, not addressed by either, is how the (copy of) the motor command can be transformed so that it can be readily compared to sensory information. In technical terms the underlying principle is often denoted as *forward models* (Jordan and Rumelhart, 1992; Kawato, 1999; Davidson and Wolpert, 2005). Here, the problem is solved implicitly by the terminology used. Usually, the term efference copy refers to the input to the forward model, also known as the predictor, whereas the resulting prediction is considered as the corollary discharge[25].

In general, the research associated with the notion of efference copy stresses the importance that action-driven predictions of sensory outcome have for perception. This, for instance, has been shown for saccades (Sommer and Wurtz, 2006) and reaching movements (Desmurget and Grafton, 2000).

Two Classical Studies

Next to introducing the concept of corollary discharge, Nobel Prize winner Sperry repeatedly emphasized the role of the motor system for sensory processing.

[23] For a more detailed account of the 'sense of agency', please refer to David *et al.* (2008).
[24] For a recent review, please see Crapse and Sommer (2008).
[25] This term is actually rarely used in the technical literature.

> "An analysis of our current thinking will show that it tends to suffer generally from a failure to view mental activities in their proper relation, or even in any relation, to motor behaviour. The remedy lies in further insight into the relationship between the sensori-associative functions of the brain on the one hand and its motor activity on the other. In order to achieve this insight, our present one-sided preoccupation with the sensory avenues to the study of mental processes will need to be supplemented by increased attention to the motor patterns, and especially to what can be inferred from these regarding the nature of the associative and sensory functions." (Sperry, 1952, p. 296)

Since Sperry suggested paying more attention to the motor system when investigating mental processes, many studies have been conducted considering this advice. In the following, a brief overview of some classical, as well as some recent experimental findings will be given.

The importance of self-actuated movements for the development of the visual system has been shown by Held and Hein (1963) in their famous experiment. Two kittens were harnessed to the same carousel, one of them being able to actively walk on the circular path constrained by the attachment, the other one sitting in a basket and thus only exposed to indirect passive stimulation. Due to this setup, both animals received the same visual stimulation, but only the actively moving one had 'access' to the motor commands. As a result, only the self-moving kitten developed normal depth perception and an unimpaired paw-eye coordination.

Further support for the essential role of action for perception has been obtained in sensory substitution experiments. In sensory substitution a (possibly impaired) modality is replaced by another 'sense'. Bach-y-Rita (1972) performed groundbreaking experiments where visual information captured from a video camera was transformed in tactile stimulation that was applied subsequently on the back of blind and normal sighted subjects. With some training, blind subjects managed to use this transformed stimulation for purposeful and goal-directed behavior. They were actually able to perceive objects in the external space instead of merely feeling them on their skin (Bach-y-Rita and Kercel, 2003), providing they were allowed to actively manipulate the camera and thus take the associated dynamics into account (Bach-y-Rita, 1972; Bach-y-Rita, 2004).

2.4.2 Enactivism – Varela, Thompson & Rosch

Many themes from the seminal work *The Embodied Mind* by Varela *et al.* (1991) have become central dogmas in the field of embodied cognition. In this book the authors query the pure computational subject matter of classical cognitive science, stress the importance of a Gibsonian (*cf.* Ch. 2.2.7) approach to cognition and challenge the concept of representations for explaining cognition. Most importantly, they emphasize that perception and cognition are inevitably linked with action.

> "By using the term *embodied* we mean to highlight two points: first, that cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological, and cultural context. By using the term *action* we mean to emphasize once again that sensory and motor processes, perception and action, are fundamentally inseparable in lived cognition. Indeed, they are not merely contingently linked in individuals; they have also evolved together." (Varela *et al.*, 1991, pp. 172–173)

To define the meaning of *embodied action*, they call for a new approach which they term *enactive* and ask "[...] how the perceiver can guide his actions in his local situation", thereby identifying the laws linking sensory and motor systems (ibid., p. 173). Instead of accepting a pre-given world, the world is now conceived as "perceiver-dependent". The sensorimotor interaction of organisms with their environment, which "appears to be filled with regularities" resulting from past experiences (Maturana and Varela, 1992), is required to *bring forth*[26] a world (Varela *et al.*, 1991). As highlighted by Varela *et al.*, this approach to perception is rooted in the work of Merleau-Ponty (*cf.* Ch. 2.2.4). For a more detailed review of enactivism, please see the special issue of the Journal *Phenomenology and the Cognitive Sciences* with an informative introduction by Torrance (2005).

2.4.3 The Extended Mind – Andy Clark

Andy Clark contributed to the discussion of *embodied cognition* with several books and essays. At the core of his deliberations is the question of how the body and the brain symbiotically simplify 'work', which traditionally is believed to be exclusively handled by the brain. According to his view, mental activity includes not only the brain, but also the body, the world and the interplay of both. Clark tries to identify the constituents, comprising likewise objects and properties not located within the brain, which *extend the mind* and thus asks the question: "Where does the mind stop and the rest of the world begin?" (Clark and Chalmers, 1998).

To demarcate work in embodied cognition from standard cognitive science he defines six ingredients that subserve to capture the general spirit of an embodied approach: *Nontrivial Causal Spread*, *Principle of Ecological Assembly*, *Open Channel Perception*, *Information Self-Structuring*, *Perception as Sensorimotor Experience* and *Dynamic-Computational Complementarity* (Clark, 2008). Lists also covering six (why exactly six?[27]) topics have been reported in the literature by Wilson (2002), Ziemke (2003b) and Engel (2010). Surprisingly, these four different listings are rather distinct than showing a common overlap.

[26] This term is adapted from the hermeneutic tradition of Martin Heidegger and his student Hans Gadamer.

[27] This question has also been raised by Rolf Pfeifer in his talk "The Emergence of Cognition from the Interaction of Brain, Body, and Environment" at the 4th EUCogII Members Conference.

Because several of the themes put forward by Clark can be readily attributed to the neural computational architectures presented in this thesis, they deserve further elucidation.

Nontrivial Causal Spread refers to a phenomenon that at first glimpse seems to be caused by complicated internal mechanisms. A more careful look, however, reveals an (almost) passive explanation, e.g. bipedal robots based on passive-dynamic walkers (Collins *et al.*, 2005).

The *Principle of Ecological Assembly* is based on the observation that a human being "tends to recruit [. . .] whatever mix of problem-solving resources will yield an acceptable result with a minimum effort" (Clark, 2008). Studies investigating the costs associated with performing intended actions indeed reveal that the geometry of the world is combined with behavioral goals, as well as with the costs associated with achieving these goals (Proffitt, 2006).

Open Channel Perception assumes that the perceptual channel between the world and the neural system is constantly 'online'. Accordingly, the world serves as an *outside memory* (O'Regan, 1992; a similar idea was proposed by Dennett, 1991 and Minsky, 1988) which can be called upon, for instance, during navigation. In contrast to classical approaches that solve the navigation task using a world model, i.e. the robot has a map of the surrounding area available or builds it on the fly during exploration of the environment (SLAM, Thrun *et al.*, 2005), we only rely on the world as an external memory to simultaneously learn sensorimotor laws and visual features essential for navigation (*cf.* experiment Ch. 3).

The principal of *Information Self-Structuring* uses "the presence of an active, self-controlled, sensing body [that] allows an agent to create or elicit appropriate inputs, generating good data" (Clark, 2008). This underpins the central statement of this thesis: robots (agents) should rely on *active perception* to structure the information readily available in their environment. Use of the interplay of action and perception can be found in all of the presented experiments. Adequate actions for a given situation, e.g. reflecting the affordances of an object (*cf.* Gibson Ch. 2.2.7), can significantly enhance the perceptual experience and take away ambiguities that might be present otherwise. This can, for instance, be exploited for object classification (*cf.* experiment Ch. 5).

The theme *Perception as Sensorimotor Experience* also runs like a red thread throughout this work. It is closely related to the theory of *sensorimotor contingencies* (SMCs) proposed by O'Regan and Noë, 2001 (*cf.* Ch. 2.4.4). For Clark (2008), the SMCs account of perceptual experience suggests that the locus of perceiving is not restricted to the brain only but involves cycles of organism-world interactions.

Dynamic-Computational Complementarity aims at retaining explanatory concepts, like computation and representation, from canonical cognitive science, because in Clark's view these are vital for understanding certain aspects of cognition.

Together with Robert Wilson, Clark defines the notion of *wide computationalism* (Wilson and Clark, 2006). This concept considers cognitive processes to be computational and, at the same time, asserts that they (may) comprise *constituents* which extend outside the cranium. This leads inevitably to the question of whether an (external) process has merely a *causal* influence on cognition or if it indeed is a constituent of it. When trying to clarify one or the other conception, it has to be heeded that "the dispute over constitution at times threatens to dissolve into nothing more than a linguistic issue" (Shapiro, 2011). Nevertheless, there are proponents for both views. To substantiate the claim that an external event is actually constituent for cognition, Clark and Chalmers (1998) introduced what has become known as the *parity principle* (Clark, 2008):

> "If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process. Cognitive processes ain't (all) in the head!"

The meaning of this principle is figuratively clarified by the famous example of Otto and Inga (Clark and Chalmers, 1998). In short, both protagonists want to see an exhibition at the Museum of Modern Art. Inga has normal capacities for memory and, hence, she can simply call upon those to obtain the address. In contrast, Otto suffers from Alzheimer's disease and therefore has to consult his notebook to retrieve the same information, i.e. "his notebook plays the role usually played by a biological memory" (Clark, 2008, p. 227). In this example, parity states that the location of the memory is irrelevant but what matters is how it is integrated and used.

Critics of extended cognition state several objections. One of them being the *coupling-constitution fallacy* (Adams and Aizawa, 2008; 2009; 2010). Schematically, the fallacy becomes evident when the observation of a given process X – that is somehow causally coupled to another process Y – leads to the conclusion that X is actually part of Y (Adams and Aizawa, 2009, p. 81). As a reply, it is often noted that constituents of a cognitive process are not (necessarily) themselves engaged in this cognitive process. This, of course, can easily lead to the previously mentioned verbal dispute. Another frequent argument against the constitution hypothesis is that their supporters have not yet managed to define the *marks of the cognitive* (Adams and Aizawa, 2010). "[W]hile transcranial cognition may be both a logical and nomological possibility, no case has been made for its current existence" (Adams and Aizawa, 2001). This stance is also referred to as *contingent intracranialism*. Further, it is argued that extended cognitive systems are not well-formed and consequently not amenable to methods from the classical sciences (Adams and Aizawa, 2010).

> "Tools do not constitute a natural kind; tools are, after all, artifacts. It is for this reason that, a would-be brain-tool science would have to cover more than just a multiplicity of causal processes. It would have to cover a genuine motley. A brain-tool science would not have

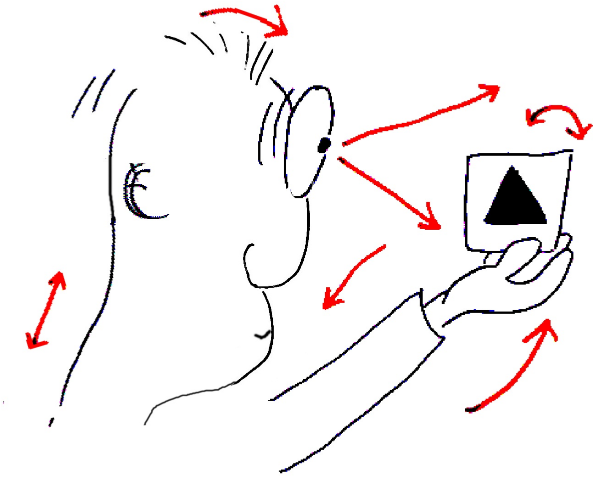
> to cover a mere disjunction of things; it would have to cover an open disjunction."

The second explanatory concept from cognitive science Clark holds on to is the notion of representations. In his view, representations are "local and action-oriented rather than objective and action-independent" (Clark, 1997a, p. 149). Representations are not an "objective world model", but they are "already geared toward the production of appropriate action" (ibid., p. 152). He distinguishes *weak* and *strong* internal representations (Clark, 1997b). The weak instance merely has the function of carrying information about an object that is in contact with the sensory organs and "control[s] immediate environmental interactions" (ibid., p. 464). According to Ward and Ward (2009), weak representations are those that neuroscientists would consider to be the receptive field of a neuron. On the other hand, strong representations refer to internal states that are used *offline* for planning or mental simulation of action.

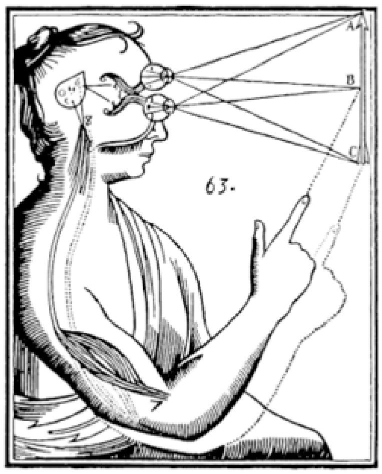
Considering the symbiotic division of tasks between the brain and the body, as well as the parity principle of the extended mind hypothesis, one is immediately tempted to establish a link to tool use experiments performed with primates (Berti and Frassinetti, 2000; Johnson-Frey, 2003; Maravita *et al.*, 2003; Maravita and Iriki, 2004; Holmes *et al.*, 2004). These studies show that tools, e.g. a rake, are readily incorporated in the body representation of the primate brain. Retrieving food with the help of the tool extends the visual receptive field of so-called *distal-type* neurons, coding previously only for the hand, to include the entire length of the rake. The receptive field of another type of neurons, termed the *proximal-type*, also expands to match the newly acquired space within reach, i.e. the peripersonal space. However, this only happens when the monkey actively makes use of the rake and not if it only holds the tool passively in its hand; again stressing the fundamental role of action for perception. Recently, tools have been developed that are taped into the brain of a rat. In order to achieve the best possible match between the intention of the animal and the action performed with the artifact, reinforcement learning is used to actively (and in addition to the internal natural adjustment) adapt the parameters of the artificial body extension (DiGiovanna *et al.*, 2009; Sanchez *et al.*, 2009).

2.4.4 Sensorimotor Contingencies – O'Regan and Noë

In contrast to many traditional beliefs that rest on the idea that the brain stores an internal representation of the world, sometimes referred to as the *orthodox internalist* view, O'Regan and Noë (2001) propose a theory where the world itself serves as an external memory (*cf. Open Channel Perception*, Ch. 2.4.3). The perceptual experience is a result of the learned mastery of what they call sensorimotor contingencies. These SMCs comprise actions, physical properties of the environment and characteristics stemming from the sensory systems. This approach naturally accounts for the differences in the perceived quality of sensory experience across modalities (e.g. *seeing* or *touching*) and stresses the necessity of



(a)



(b)

Figure 2.2: Sensorimotor contingencies illustration and the Cartesian theater (a) The spongeman (name given by O'Regan) exercising sensorimotor skills. According to the sensorimotor contingencies theory of O'Regan and Noë, this is the only 'ingredient' necessary to give rise to a perceptual experience (see text for details). Drawing courtesy of Kevin O'Regan[28]. (b) Illustration of Descartes demonstrating the dualism of mind and body, later termed the 'Cartesian theater' by Dennett (1991). The epiphysis looks at an internal cinema screen to yield the perceptual experience of seeing.

the interplay between actions and perception which has already been suggested by von Helmholtz (*cf.* Ch. 2.2.1).

We seem to have a rich, detailed impression of any given ordinary visual scene. This is despite experimental findings about the blind spot, change blindness, saccadic smears and an impoverished peripheral (color) vision. So where does this paramount resolution come from that we are obviously experiencing? A trivial, albeit not very convincing, explanation is called the *grand illusion* view, stating that this rich and detailed visual experience is simply an illusion our brain generates. This claim, as well as the orthodox internalist view, is denied by O'Regan and Noë. Instead, they suggest that "seeing is a way of acting" (Fig. 2.2 a). Perceiving agents learn to acquire (motor) skills that allow to pick up all the details needed, which are readily available in their environment, to obtain a rich and detailed sensory impression. Thus, it is not internally generated neural activity which is responsible for perceptual presence, but rather the mastery and the access to our sensorimotor skills that enables us to see (O'Regan and Noë, 2001; Bompas and O'Regan, 2006), smell (Cooke and Myin, in press), hear (Aytekin *et al.*, 2008) and touch (Myin and O'Regan, 2009; O'Regan, 2011). It is important to note that the authors emphasize that their approach does not only explain (visual) cognition but at the same time it is supposed to account for (visual) consciousness (O'Regan and Noë, 2001; O'Regan, 2011).

[28]`http://nivea.psycho.univ-paris5.fr`

Despite stating that the active engagement in a sensory manipulation is sufficient to give rise to a perceptual experience, O'Regan (2011) does not deny that the brain is somehow involved in this process and he also admits the "usefulness of representations" (ibid., p. 68). The words *representation*, *activate* and *generate*, however, should be used with caution (ibid.).

Critics of this account put forward several arguments. There must be something more than just the "knowledge of the relevant sensorimotor contingencies" (O'Regan and Noë, 2001, p. 943) for perception. Simply by stating that there is no *explanatory gap* (Levine, 1983) does not compensate for the lack of understanding how and why we *feel* (perceive) the way we do (see also comments of e.g. Kurthen, Oberauer, Manzotti and Sandini, on O'Regan and Noë, 2001). This objection is reminiscent of the *Cartesian theater* analogy suggested by Dennett (1991). Descartes assumed that the 'mind', materialized in the pineal gland, looks at a cinema screen in order to 'see' (Fig. 2.2 b). Of course, this leads to an infinite regress. Still, assuming that the knowledge of sensorimotor contingencies are constituents of the perceptual experience and "there is nothing more to it" (personal communication with Kevin O'Regan) does not provide a satisfying explanation. Another tripping hazard of the SMCs account lies in the level of interpretation. If the claim that perceptual experience is to be equated with skillful sensorimotor activity is taken literally, as is suggested for instance by the statement "seeing involves testing the changes that occur through eye, body, and attention movements" (O'Regan and Noë, 2001, p. 947), we would actually need to be engaged actively in the process of perceiving each time. On the other hand, if only the *potential* to perform certain actions is required, e.g. as suggested by "seeing a stationary object consists in the knowledge that if you were to move your eye slightly leftwards, the object would shift one way on your retina, but if you were to move your eye rightwards, the object would shift the other way" (ibid., p. 949), then a *brain in a vat*[29] could do the job as well. This, however, contradicts the hypothesis that "it is not the brain process that generates the feel, but the mode of interaction that constitutes the feel" (O'Regan, 2011, p. 113).

Only few computational studies[30], either with real or virtual robots, exist that explicitly relate their implementation to the SMCs framework (Philipona *et al.*, 2003; Philipona *et al.*, 2004; Olsson *et al.*, 2006; Aytekin *et al.*, 2008; Maye and Engel, 2011). Interestingly, there are more studies that do not explicitly establish this relation but do qualify at least to the same extend (Fitzpatrick and Metta, 2003; Metta and Fitzpatrick, 2003; Roy, 2005a; Roy, 2005b; Bovet and Pfeifer, 2005; Natale *et al.*, 2005; Maillard *et al.*, 2005; Hoffmann, 2007). Because these experiments picture a general access to sensorimotor learning they will be discussed in the cognitive robotics section (Ch. 2.4.9).

[29] Well known thought experiment: the brain is located in a vat sending exactly the same motor commands and receiving exactly the same sensory stimuli as if it were in a skull. If this is the only way for the brain to interact with its environment, then it is not possible for it to tell whether it is located in a skull or a vat.

[30] We are aware of.

Philipona *et al.* (2003; 2004) addressed the question of whether it is possible for an agent with an unknown body to infer that the body is located in a 3-D Euclidean space. Indeed, their algorithm detected the dimensionality of the world via deducing sensorimotor laws that link actions to perceptions. Further, this method enabled a virtual rat to control its head in a rigid fashion, despite an initially unknown body and the unknown structure of the world it was embedded in.

Using an approach rooted in information theory, Olsson *et al.* (2006) showed that it is possible for a real robot to learn a model of its own sensors and actuators. Initially, the robot performed random movements (*cf.* motor babbling, Ch. 2.4.1) to create an informational map of its sensors, which was subsequently exploited to acquire knowledge about the effects that different actions have on those. In the end, the robot was able to perform basic visually guided movements.

Aytekin *et al.* (2008) proposed a computational method for learning of auditory space based solely on acoustic inputs and their relation to motor outputs. By means of manifold learning methods it was demonstrated that organisms like humans and bats can learn to localize sound sources. This requires neither any *a priori* neural representation of their head-related transfer function nor any prior experience with auditory spatial information.

In another experiment, sensorimotor contingencies have been realized using conditional probabilities that were represented as a set of Markov models (Maye and Engel, 2011). Implemented on a real robot platform, the experience-dependent probability distributions allowed two different objects to be distinguished. Depending on the object, a characterizing SMC was triggered, which in turn led to an action that signaled the classification result.

2.4.5 Directive Minds – Andreas K. Engel

As previously noted, Engel (2010) describes the rethinking that is currently taking place in the cognitive sciences as a *pragmatic turn* (*cf.* Ch. 2.2) away from the static representation-centered framework towards an action-centered paradigm that is concerned with understanding of the intimate relation between action and cognition. At first, this 'action-driven' paradigm found its way into robotics and has only recently begun to have a notable impact on research in cognitive psychology and neuroscience.

Several challenges have to be met by the emerging 'pragmatic neuroscience' framework that aims at explaining how brains function as 'vehicles of world-making', rather then just being world-mirroring devices (*cf.* enactivism, Ch. 2.4.2). The experimenter, aware of the dynamic nature of cognition, has to maintain a holistic perspective during design, carry-out and analysis of an experiment. Not the neural activation patterns in relation to stimuli alone are of primary importance, instead the action the subject is currently engaged in, as well as the environment they are embedded in, have to be focused on. Further, the cognitive system as a whole, including interactions within and across neural assemblies of several subsystems (e.g. top-down influences), has to be investigated.

Because neural activation patterns do not carry 'images' or symbolic representations of the external world, but rather "support the capacity of structuring situations through action", Engel (ibid.) suggests getting rid of the "cognitivist burden" carried on with the word *representation* and replace it with the expression *directive*. He stresses that directive is not just a different term for 'action-oriented representation', which is merely located in the head. Instead, the dynamics of directives extend beyond the brain through the entire cognitive system, including body and environment (*cf.* extended mind, Ch. 2.4.3 and SMCs, Ch. 2.4.4). Another characteristic of directives, which can readily be attributed to the artificial neural architecture developed within this thesis (*cf.* Ch. 3 and 4), is based on the description that "directives are immediately related to action selection. Activating directives directly controls the respective action" (ibid.). Further, "[t]he concept of an object corresponds to 'nothing but' the set of possible actions relating to this object; there is no context-neutral 'description' above and beyond the directives" (ibid., *cf.* affordances, Ch. 2.2.7).

Directives are not isolated neural activation patterns. They are the dynamic interactions between highly distributed neural populations across numerous brain regions (ibid.). These distributed processes comprising activity in sensory, motor, limbic and memory regions have to be coordinated via a dynamic 'binding principle'. A potential mechanism that makes a (temporal) coordination possible is called *neural synchrony* (Engel *et al.*, 1992; Singer and Gray, 1995; Engel and Singer, 2001; Engel *et al.*, 2001; Fries, 2005). Synchrony is often associated with oscillatory activity that, at least over larger distances, enables neuronal communication. However, experimental evidence relating neural synchrony to the generation of actions is thus far only supported by few studies (Engel, 2010).

It is arguable whether the introduction of a new word (directives) helps to resolve the ambiguities and (logical) problems the notion of *representations* causes (Press, 2008). In any case, due to a pervasive relation of the *directive minds* to the theory of sensorimotor contingencies this stance is, of course, susceptible to similar critical objections (Ch. 2.4.4).

2.4.6 The Theory of Event Coding

The historic roots of the theory of event coding (Hommel *et al.*, 2001; Hommel, 2009) have been presented in Ch. 2.2.3. TEC is a general framework explaining how perceived and produced events (stimuli and responses) are represented in common domain and how this *common coding* generates action and perception. This is in contrast to traditional views that regard sensory, motor and cognitive processes as distinct entities and, hence, require a transduction of modal (sensory) codes to amodal symbols, which in turn are processed to generate actions (Newell and Simon, 1972). Unfortunately, TEC is underspecified at present and some of its core concepts, such as an event, have been defined only loosely. This is realized by the authors, admitting that "TEC's main mission at this point is to stimulate deliberations and discussions about basic principles of perception/action architectures" (Hommel *et al.*, 2001, p. 914). For a detailed recent review comprising multiple facets of ideomotor theory in general, and TEC in particular, please see Shin *et al.* (2010).

2.4.7 Dynamical Systems Theory

The behavior, how a system changes over time, can be formalized using dynamical systems theory (DST), an area of applied mathematics. Usually differential equations are used to describe how the dynamical system changes over time, but also a recurrent neural network can be considered (*cf.* experiment 3, Ch. 5). Dynamical systems have some intriguing properties. They are *self-organizing*, also referred to as *emergent*, they can be coupled and can have fixed-points functioning as attractor states. Due to the general description – a system that changes over time – almost any system is a dynamical system and so is the brain. This was already realized by Turing (1950):

> "The nervous system is certainly not a discrete-state machine. A small error in the information about the size of a nervous impulse impinging on a neuron, may make a large difference to the size of the outgoing impulse. It may be argued that, this being so, one cannot expect to be able to mimic the behaviour of the nervous system with a discrete-state system."

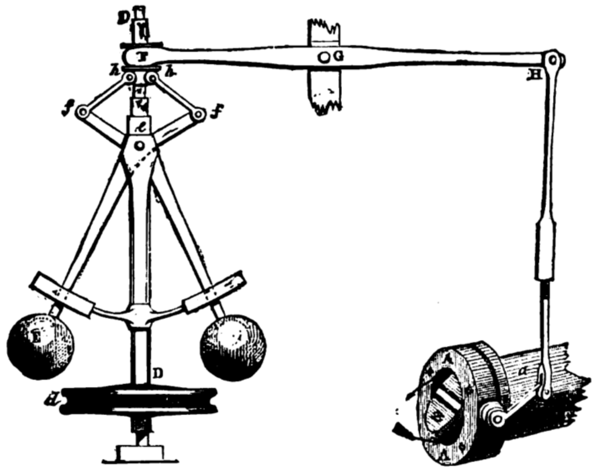
This reasoning is actually shared by the proponents of DST. They argue that "cognition should be described in terms of agent-environment dynamics rather than in terms of computation and representation" (*Radical Embodied Cognition*, Chemero, 2009). What could be more appropriate than describing a dynamical system with DST? The continuous flow of information among brain areas and the environment establishes a complex and inseparable dynamical system. Hence, there is no need to draw artificial boundaries between action and cognition. Yet, it should be considered that "the foundational aspects of a 'dynamicist' cognitive science should be clarified before it can really offer itself as a novel candidate paradigm" (Spivey, 2007). In the following paragraphs several hallmarks of DST that are important for the embodied cognition 'movement' will be introduced.

Dynamical Hypothesis – van Gelder

The *dynamical hypothesis* proposed by van Gelder (1998) focuses on two elements. The first part specifies the *nature* of cognitive agents, stating that they are dynamical systems. The second part, denoted as the *knowledge* hypothesis, claims "that we can and should understand cognition dynamically" (ibid., p. 619).

To strengthen his hypothesis, van Gelder (1995) presents Watt's centrifugal governor[31] as an example (Fig. 2.3). Prior to this invention steam engines were not

[31]Named after its Scottish inventor James Watt (1736 – 1819).



Governor and Throttle-Valve.

Figure 2.3: Watt's centrifugal governor. Example utilized by van Gelder to strengthen his dynamical hypothesis[32].

able to maintain a constant output speed, leading to a fluctuating performance that possibly affected the quality of the product (e.g. in cotton mills). The governor regulates the flow of steam, and thus the output speed, via a throttle valve. Once the flywheel slows down, the valve opening increases, compensating for the loss of power. However, if the speed of the flywheel is too high, the opening of the throttle-valve decreases, thereby slowing down the flywheel.

Van Gelder (ibid.) offers two possible solutions for how this system could function. The first is rooted in cybernetics (*cf.* Ch. 2.2.5). Measuring the actual value of the flywheel speed and comparing it to the desired value allows an error to be calculated, which in turn can be used to adjust the throttle-valve and thus regulate the output speed (proportional controller). This computational process happens in discrete time steps and successively. Further, it requires a representation and a memory. In contrast, the second explanation is based on the *embodiment* of the device. During rotation of the flywheel, a centrifugal force causes the flyballs to rise, simultaneously altering the opening of the throttle valve. This mechanism can be formalized using differential equations. In the continuous interaction of the flywheel and the throttle valve there is no need to maintain a representation. According to van Gelder (ibid.), for two reasons – continuity and lack of representations – the second dynamical explanation has to be preferred over the computational one. Despite being a dedicated anti-representationalist, he admits that there might be a situation where representations could be useful for the understanding of some dynamical systems (ibid.; 1998).

The reasons presented by van Gelder for preferring one approach over the other are certainly debatable. The governor is a dynamical system, but is it also a cognitive system (comparable to the brain)? What do the two offered solutions tell us about the functioning of cognitive systems in general? The computational explanation

[32]Drawing adopted from `http://commons.wikimedia.org`.

seems to be the more general one, a feature often requested for cognitive systems because it does not contain a detailed description of how certain parts of the 'body' relate to each other (e.g. the differential equation contains the arm angle as a function of engine speed). On the other hand, being a coupled system suggests that all parts, despite the fact that they are causally linked, are at the same time constituents of the (cognitive?) process.

Dynamic Field Theory – Thelen

The computational modeling study of Thelen *et al.* (2001) is often mentioned as the prime example of successful application of DST in the field of embodied cognition.

> "To say that cognition is embodied means that it arises from bodily interactions with the world. From this point of view, cognition depends on the kinds of experiences that come from having a body with particular perceptual and motor capacities that are inseparably linked [...]" (ibid.)

They modeled Piaget's A-not-B error experiment (1954) using dynamic field theory (Amari, 1977). Facing an about six to seven month old infant with a simple reaching task elicits *perseverative* behaviour. On the table in front of the infant two identical containers, one denoted as box A the other as box B, are arranged. The experimenter presents a (toy) object to the toddler and places it beneath container A, so it is no longer visual. This results in reaching of the child toward container A. After repeating this several times the experimenter alters suddenly its strategy and places the object of desire, as the infant watches, under container B. Instead of reaching toward the new location B the toddler continuous to reach toward A, perserving the original activeness. This effect vanishes again at an age of about one year. According to Piaget this is due to a not yet matured concept of object permanence which develops accompanied by increasingly more complex representations. Interestingly, the infant succeeds in reaching toward B when for this action a change of its motor plan is necessary, e.g. it has to bypass an obstacle or springs are attached to its arms thereby increasing the resistance of the movement.

Thelen *et al.* (2001) deliver an explanation completely getting by without an object concept. "Indeed the cornerstone of our dynamic model is that 'knowing' is perceiving, moving, and remembering as they evolve over time, and that the error can be understood simply and completely in terms of these coupled processes." Thus, their model explains the complete behavioral range of the A-not-B error effect by combining evidence of the visual modality, the reaching and remembering in a single motor decision field that yields as an outcome 'knowing' or 'not-knowing' about the proper target location.

Dynamic Categorical Perception – Beer

A central dogma of the dynamicist view is that interactions between brain, body and world contribute to the emergence of cognition. In this context, the dynamical interaction between the brain and the body is referred to as *embodied*. At the same time, the body is embedded (*situated*) in the environment. All three components, each by itself a dynamical system, are jointly engaged in a coupled relationship (Fig. 2.1 b). "Because all of the individual components are described in the same mathematical language, it is much easier to approach questions involving their interaction" (Beer, 2003).

In his well-known study, Beer (ibid.) investigated categorical perception in a simulated agent. As the agent moved left and right it had to discriminate circles from diamonds falling down from above. If it encountered a circle it had to catch it, i.e. center itself beneath it, whereas if it was confronted with a diamond it was supposed to avoid it. To accomplish its goal the agent was equipped with seven sensors, each having a single line of sight. These sensors served as the input nodes to a continuous time recurrent neural network (CTRNN). In addition to the seven input nodes the network had five hidden nodes, as well as two output nodes that acted as motor neurons. Evolutionary algorithms were used to evolve the 'nervous system' of this special type of connectionist network. Based on performance in the task, a successful agent was chosen and subjected to further dynamical systems analysis. As it turned out, it was not possible to clearly identify neural correlates for either of the two categories. Instead, the agent had learned to actively scan the falling object. Thus, rather than actually representing externally defined category labels, the categories are reflected by the behaviors of the agent, which in turn is guided by complex brain, body and environment interactions. Similar work with CTRNNs and simple simulated agents, summarized with the term *minimal cognition*, has been conducted by Di Paolo and co-workers (Di Paolo, 2000; Izquierdo-Torres and Di Paolo, 2005; Fine *et al.*, 2007).

Next to acknowledging that connectionist recurrent neural networks can approximate arbitrary dynamical systems, Beer (in press) points out that

> "the connectionist framework emphasizes the network architecture, the learning algorithm, the training protocol and the intermediate distributed representations that are developed. In this sense, many connectionist models are still disembodied, unsituated, and computational (albeit distributed) in nature."

This, however, suggests that only the dynamic *interpretation*, not the dynamics itself, leads to an embodied and situated cognitive system and it confirms the objections expressed by Spivey (2007), namely, that the 'dynamicist' cognitive science paradigm has yet to be defined (see above).

2.4.8 Virtual Reality

The notion of *embodied interaction* was introduced to human-computer interaction early this century (Dourish, 2001). It aims at incorporating phenomenological concepts (Ch. 2.2.4) and emphasizes the importance of our body for the social and physical interaction with the world we are embedded in. This world, however, does not need to be the real physical world – it can also be a computer-generated *virtual environment* (VE). A variety of design and research projects are specifically concerned with embodied metaphors, tangible user interfaces and wearables in *virtual reality* (VR) systems. One of the primary goals of such systems is to fully *immerse* the user, leading to the experience of really *being there*. This conscious experience has also been denoted as *presence*, the subjective feeling of being in a virtual environment while being temporarily unaware of one's real location and the technology responsible for the virtual input to the senses. Sanchez-Vives and Slater (2005) stress the constitutive role of actions needed for the experience of presence to come into existence[33]. "The key to this approach is that the sense of 'being there' in a VE is grounded on the ability to 'do' there" (ibid.).

It has also been suggested that intuitive interaction methods for VR are rooted in embodied skills (Beckhaus and Kleesiek, 2011). In particular, it is important for an intuitive interaction device that it signals appropriate and inherent affordances (*cf.* Gibson, Ch. 2.2.7) that integrate well in the current setting. Next, motor skills and knowledge about the usage of the device either have to be known in advance or can be achieved rapidly, immediately leading to a tool that extends our body (Ch. 2.4.3). As a further potential neuroscientific explanation the authors refer to the concept of efference copy (Ch. 2.4.1). If voluntary movements performed with a user interface lead to virtual sensory feedback (e.g. visual flow) that is (almost) identical to the predicted sensory experience and, thus, matches the embodied concepts and expectations of the user, then the usage of this device is indeed intuitive.

2.4.9 Cognitive Robotics

Cognitive robotics is a broad research area comprising several sub-fields, e.g. developmental robotics, that aim at grounding cognition in sensorimotor experiences. One common goal is to build autonomous robots that can operate in complex, open-ended real-world scenarios, possibly interacting with humans. Another frequent objective is to use robots to evaluate neuroscientific hypotheses. Ideally, this should result in a fruitful cross-fertilization where the implementation of principles and constraints derived from animal and human cognition lead to better performing robots and the confirmation of hypotheses.

Robots are embodied. Their body, however, is usually drastically different to the human body. Taking embodiment claims literally could, thus, suggest that robots, given their current technical development, cannot be good models for

[33] Premotor cortex is part of the network proposed for the experience of presence. For a detailed review of the underlying neural mechanisms, please see Jäncke *et al.* (2009).

human cognition because "to conceive of the world as a human being does require having a body like a human being" (Shapiro, 2011, p. 71)[34]. Another criticism can be summarized by the term *anthropomorphic bias*. It is very hard, if not impossible, to imagine how the body of a robot 'feels' and how, due to these morphological differences, the world is conceived from a robot's perspective. Yet, this is hardly ever considered when devising algorithms for robots.

On the other hand, if a robot explores the possible actions by itself and learns the sensory consequences, both objections can be relaxed. Self-acquired *sensorimotor laws*, relating actuators and sensors, are shaped by the embodiment, i.e. the morphological configuration, of the robot and the physical properties of the world it is embedded in. Further, using a state-of-the-art humanoid robot leads to the best possible result one can currently obtain (assuming that the allegation would be true). As a matter of fact, for most of our experiments we use the Nao robot[35], a humanoid with 25 degrees of freedom.

Historic Roots

Some of the first electronic autonomous robots were constructed by Walter (1953). Instead of viewing mental processes in terms of digital computation, Walter stressed the importance of purely analog circuits. According to his working hypothesis, the key element for the functioning of the brain is determined by the relay of the neurons (how they are 'wired up'). Despite having (rich) connections between only few brain cells, the analog machines demonstrated complex behaviors, e.g. phototaxis, which confirmed his belief.

Braitenberg vehicles are another form of simple agents based on embodied cognition (Braitenberg, 1984). They display complex and dynamic behavior that emerges from sensorimotor interaction between the agent and its environment, without any need for an internal representation. In the simplest form, sensors are directly connected to actuators. Depending on the connection, i.e. excitatory or inhibitory, different (goal-directed) movements result. Yet, the agent operates purely reactively without any information processing.

Another pioneer in the field of cognitive robotics is Rodney Brooks. He created situated and embodied robots that "experience the world directly – their actions are part of a dynamic with the world, and the actions have immediate feedback on the robots' own sensations" (Brooks, 1991a). To implement these behavior-guided robots, Brooks proposed a *subsumption* architecture. Instead of establishing a hierarchy of different modules, each performing computations on a representation as suggested in GOFAI, he arranged the different components in parallel, directly linking sensing with acting. In this way, the traditional sequence of sensing, building up a representation, planning and acting no longer exist.

[34] This is not the opinion of Shapiro. He uses this extreme statement to clarify the concept of embodiment. This has also been done in a similar manner by Pezzulo *et al.* (2011).

[35] Aldebaran Robotics

> "Each activity producing layer connects perception to action directly. It is only the observer of the Creature who imputes a central representation or central control. The Creature itself has none; it is a collection of competing behaviors. Out of the local chaos of their interactions there emerges, in the eye of an observer, a coherent pattern of behavior." (Brooks, 1991b)

According to Brooks, representations are – at least for very simple levels of intelligence – not necessary, because the robot demonstrator "uses the world as its own model"(ibid.). This, of course, shows a striking resemblance to the claim that the world serves as an outside memory (*cf.* SMCs, Ch. 2.4.4). Furthermore, the statement of Brooks points out another type of *anthropomorphic bias*, which holds true for the work of Braitenberg, Walter and others. It might not be the robot that is intelligent, but the human observer who attributes, based on own experiences and expectations, intelligence to the behavior of the machine. Is the, albeit complex, behavior we observe, really a trace of *cognitive* processing[36]?

Selected Publications

In this section selected contemporary publications from the vastly growing field of cognitive robotics are presented. Due to the myriad of publications, no complete list can be given; rather some, according to our view, representative publications are considered that highlight different possible approaches for modeling embodied cognition on robotic systems, thereby exploiting the interaction with the environment for the acquisition of sensorimotor laws.

Pfeifer and Scheier (1997) presented 'sensory-motor coordination agents' that were located in an arena with small and big cylinders. In order to accomplish their goal – picking up the small cylinders and ignoring the big ones – they needed to be able to discriminate between the two kinds of objects. To tell apart the two classes by visual means only is rather difficult. However, exploiting a built-in reflex, which caused them to circle around an object, facilitated the classification. While *actively* exploring a cylinder, the object diameter led to a distinct angular velocity that could subsequently be used for discrimination.

In another study, Bovet and Pfeifer (2005) proposed a robot control architecture for learning delayed rewards without memory. An artificial mouse, equipped with sensors for vision, touch and reward was supposed to find food (i.e. an electric reward) in a T-maze. At the entrance to the maze a tactile cue signaled on which side the reward will be. However, the agent did not know this. For each sensor and motor modality (e.g. intensity values of the camera or movement direction) there was a set of neurons representing the current value and another set of neurons indicating the change of these values. Additionally, there was also a neuron whose activity reflected the reward. All groups of neurons were interconnected and a simple Hebbian learning mechanism enforced the connections that were simultaneously active, i.e. the connections were strengthened based on

[36] *cf.* definition of cognition in Ch. 2.1

correlations. In this task, the world was used as an external memory (*cf.* SMCs, Ch. 2.4.4). After a few trials the sensory impressions, which are readily available in the environment, the movements and the reward displayed correlated activity, leading to the desired behavior.

Roy (2005a) proposed a computational architecture in which the meaning of spoken words and sentences is grounded in multi-modal sensorimotor experiences of a robot. For instance, words that describe a color can be associated with the expected sensory experience. Similarly, affordances (*cf.* Ch. 2.2.7) like reachable or graspable are grounded in motor and sensor primitives, as well as in an expected sensory outcome of a specific action chosen by a planner. For example, given a word like ‘ball’ will result in a set of expected action outcomes the robot can perform with the object.

Lungarella *et al.* (2005; 2006) investigated how different information theoretic and statistical measures capture the information structure between sensor and motor channels of a robot. In their experiments, they showed how actions of an embodied agent lead to a structuring of the sensory input. The robot, while performing saliency-based attentional behaviors, decreases the entropy of the sensory information and increases the statistical dependencies between the sensors. They demonstrated that the information flow in sensorimotor networks is spatially and temporally specific and that it can be affected by learning and by changes in body morphology.

Pitti *et al.* (2009) offered a computational model that monitors and simulates the concordance between voluntary actions and their consequences. In their study, body representations were encoded as spatio-temporal patterns in a spiking-neural network. While a head-neck-eye robot interacted with its environment the connections of the network were modulated via spike-timing-dependent plasticity. Over time, the most congruent sensorimotor pairs were reinforced, anticipating the ongoing sensorimotor activity and predicting the next state of the embodied agent (*cf.* efference copy, Ch. 2.4.1). This led to a self-organization of sensorimotor maps, enabling the robot to fixate salient objects as well as making saccades to new locations. Further analysis of the neural activities of the network revealed that the agent was able to discriminate between self-generated and externally caused movements.

Hoffmann (2007) presented a mobile robot that combines action-based object recognition and sensorimotor anticipation, i.e. expected sensory effects resulting from a simulated motor command, for obstacle avoidance in a navigation task. Initially, the robot learned to predict, using a multilayer perceptron, what effect its movements have on the visual input. Later, this knowledge was exploited for obstacle avoidance via a ‘mental’ simulation of movements. In turn, the outcome of this simulation was used to avoid inappropriate actions.

Butko and Movellan (2010) proposed a generative model that allows autonomous agents to learn intentional looking behavior without access to anything beyond their own sensorimotor experiences. This was realized using a Bayesian framework that combines two different sources of uncertainty. The robot only had access to its own motor commands and to the sensory information from its camera. Yet, this was sufficient to discover how sensors and actuators relate and to use these sensorimotor skills for purposeful looking at visual targets.

Natale *et al.* (2005) reported on a developmental sequence of a humanoid robot that is able to learn about its body, the interaction with its environment and eventually is able to exploit a learned affordance (graspable) for the search of objects which offer the same affordance (*cf.* Ch. 2.2.7). Equipped with a basic set of sensor and motor competencies, as well as a sophisticated visual attention system, the robot was able to improve its sensorimotor coordination via the interaction with its environment. First, it learned the weight of its arm and to recognize its hand. Then it started reaching and grasping for objects and, if successful, it acquired physical properties of the grasped objects. For this purpose, the robot learned an internal model of its hand, which allowed it to be localized in the visual scene and, at the same time, an inverse model that allowed an intended position in space to be reached towards.

In a series of articles, Fitzpatrick and co-workers demonstrated a robot that is able to acquire visual experience through simple experimental manipulations (Fitzpatrick and Metta, 2002; Fitzpatrick and Metta, 2003; Fitzpatrick *et al.*, 2003; Metta and Fitzpatrick, 2003). While actively probing its environment, it accumulated knowledge about which parts are physically connected and, thus, move together and which parts are (more or less) independent entities that can be moved separately. The authors stress that only the repeated 'playing' with objects enabled the robot to acquire a sensorimotor mapping that links actions to its visual consequences (*cf.* motor babbling Ch. 2.4.1). Next to this 'discovery mode', the robot also had a 'goal-directed mode', which was obtained by the inversion of the sensorimotor mapping. This allowed the robot to select an action that will lead to a particular visual change. After 'learning to act', the robot was able to segment, recognize and localize objects without any prior knowledge about their visual experience. Further, the agent was able to utilize this information to recognize the effect of an action performed by a human and to subsequently imitate its teacher.

2.5 Active Perception in Computer Vision

In the period from 1986 – 2010, about 2000 research papers have been published that are closely related to the topic of active vision perception in robotics (Chen *et al.*, 2011). The field emerged around 1988 with several influential articles coining expressions like *active vision* (Aloimonos *et al.*, 1988), *active perception* (Bajcsy, 1988), *smart sensing* (Burt, 1988) and *animate vision* (Ballard, 1991). Since then, these terms have been commonly used, sometimes even interchangeably, despite varying intentions pursued by the original authors. In her article on active perception Ruzena Bajcsy suggested for machine vision and robotics that

> "[...] it should be axiomatic that perception is not passive, but active.

> Perceptual activity is exploratory, probing, searching; percepts do not simply fall onto sensors as rain falls onto ground. We do not just see, we look." (Bajcsy, 1988)

Since this time the field has been prospering. Still, many of the current approaches do not follow these insights.

Usually, active perception in computer vision refers to a sensor that can be moved actively. This is done to narrow down the potential interpretations the sensor values permit. Therefore, the robot has to decide 'where to look', it needs the ability to actively place the sensor at different viewpoints through a planning strategy. This procedure is called *sensor planning* and one of the main research areas in active vision systems. In general, active vision research can be roughly subdivided into two approaches: model-based and model-free. Clearly, the model-free approaches, i.e. where the object model has yet to be learned, are more related to the methods for learning of sensorimotor laws discussed above.

Next to moving a sensor, it is also possible for a robot to influence the orientation of a camera by (voluntary) self-motion of its whole body or just parts of it. In this way, the agent is able to increase its field of view and orient itself to a region where it expects the most useful information.

It is often the case that a single perspective does not provide enough features that allow for an unambiguous identification of an object (Byun and Nagata, 1996). Two objects may indeed have all views in common with respect to a given feature set. Thus, the only way to distinguish these two objects is by considering a sequence of views (Roy *et al.*, 2000; 2005). For a review of sensor planning for active recognition, please refer to Roy *et al.* (2004).

For a human, gaze and attention are important to purposefully explore the environment. This is also mimicked in robots. Active gaze control allows the limitations imposed by a monocular camera with a relatively small field of view to be overcome. Frintrop and Jensfelt (2008b) extended the SLAM algorithm to include active gaze control. Based on a biologically motivated visual attention mechanism for salient feature detection their framework created and maintained a sparse set of landmarks. The landmarks were actively tracked to get a better pose estimation. Further, unknown regions of the environment were explored to get a better distribution of landmarks. An analysis revealed that an active camera control outperformed the passive approach (Frintrop and Jensfelt, 2008a).

Bjorkman and Kragic (2004) introduced an automatic real-time vision system consisting of a four-camera stereo head. Based on disparity, one of the binocular cameras is used to detect salient stimuli in the periphery. The other stereo camera is used for the foveation of the identified focus of attention. In this way the conflicting requirements – wide field of view *vs.* high resolution – can be overcome. Combining appearance and geometric models, the system was successfully tested in a realistic indoor environment (Rasolzadeh *et al.*, 2009).

In an algorithm for automatic *active segmentation* of a visual scene, Mishra *et al.* (2011) also relied on an attention mechanism. Initially, the artificial visual system fixates a salient point in the scene. Assuming that this fixation point belongs to an object (or part of it), their method starts to segment this object by finding the 'optimal' closed contour around the identified fixation point in polar space. To find this optimal contour, first, all visual cues of the scene are combined to generate a probabilistic boundary edge map. Next, this edge map is used to search for a connected set of boundary edge fragments. The segmentation is further refined via a feedback mechanism between identified regions and low-level visual cues. The primary difference between this algorithm and canonical approaches is that only one object at a time is segmented and not the entire scene. To segment further nested regions, the segmentation process is repeated for salient points lying inside the object of interest.

In contrast to the approaches presented in the section on cognitive robotics (Ch. 2.4.9), the methods summarized here do not explicitly include knowledge about the actions of the robot into their algorithms. There is (usually) no forward model that, based on the agent's motor commands, is utilized to predict a sensory change. Rather most methods in active vision benefit either from the dynamic sequences leading to the extraction of better features or from the active focusing on only specific elements currently present in the scene.

2.6 Synopsis

The overview of historic and state-of-the-art embodied theories has shown that the research field is still developing (Pezzulo *et al.*, 2011). At the current stage, embodied cognition should be understood as an agenda or research theme, rather than a well-defined theory or paradigm already in place (Engel, 2010; Shapiro, 2011). The meaning of various definitions and concepts is far from uniform, sometimes even threatened to dissolve into a merely verbal dispute. Hence, the identification of a common ground is far from being trivial. Considering the historic insights gathered by von Helmholtz and others (*cf.* Ch. 2.2), EC should rather be considered a minor reformation. However, knowing that these findings have largely been ignored by the first-generation cognitive science and GOFAI community, one is very tempted to accept that the changes are considered as a "Copernican revolution" by some researchers (Lindblom and Ziemke, 2006). Shapiro (2011) notes that the embodied view undoubtedly resolves some problems but at the same time adds new ones. Due to this dilemma he sees a *raison d'être* for both fields. Yet, other authors claim that all challenges can be met by simply augmenting the long-established computational-representational framework of classical cognitive science (Thagard, 2010).

It immediately becomes apparent that this discourse on embodied theories cannot be resolved within this thesis (if ever). However, knowledge of the various views helps to define the *sensorimotor design principles* that are of major importance for the scope of this work, stressing the fundamental role of action for perception. Various studies from different disciplines have been presented supporting this notion. As a matter of fact, in the experiments that will be presented within this thesis, the agent *must* act in order to perceive, mimicking the *perceiving is* *a way of acting* hypothesis put forward by the SMCs account (Ch. 2.4.4). Only when performing movements, the agent is able to learn sensorimotor laws that link action and perception. Further, to successfully master the given tasks the agent will need to rely on the *world as an outside memory*. Also the principle of *information self-structuring* (Ch. 2.4.3) can readily be found in the presented experiments.

Regarding the elusive notion of *representations* we follow the suggestion of Clark (*cf.* Ch. 2.4.3). If not otherwise noted, we refer to the “local and action-oriented” weak instance (Clark, 1997a), denoting a receptive field of a neuron.

Perception is something you do,
not something that happens to you.
Bruce Bridgeman

3 Chapter
Reward-Driven Learning of Sensorimotor Laws for Navigation

3.1 Introduction

A frequently re-occurring task for humans and robots is the autonomous navigation towards a target location. In robotics, this is often realized using a world model, i.e. the robot has a map of the surrounding area which allows it to do planning. If this map is not 'pre-given' it can e.g. be acquired using the SLAM algorithm (Thrun *et al.*, 2005). This has severe limitations. Whenever located in an unknown environment, the robot has to build up this map before it can perform goal-directed actions. Inspired by nautical navigation we propose a different approach that allows navigating to a target position, e.g. a harbor or a docking position, in an unknown environment, once the robot has learned what visual features it has to attend to and how those relate to its actions. Consider the nautical leading lights shown in Fig. 3.1. The lower sign (U) is a triangle pointing upward, the upper sign (O) a triangle pointing downward. They are placed in some distance from each other to cause a perspective offset and usually signal the entryway to a harbor. Moving the boat and simultaneously observing the leading lights immediately illustrates, at least for a human, the mode of operation of this navigation aid. If they are in line – the tips of the triangle point at each other – the tillerman has to maintain his current bearing, i.e. move forward. On the other hand, if the upper triangle is shifted to the left he has to bring the triangles in line again by moving to the right and *vice versa*. In this way, the harbor can be reached without knowing the point of the compass and the exact location.

To implement this novel robotic navigation paradigm we apply an algorithm

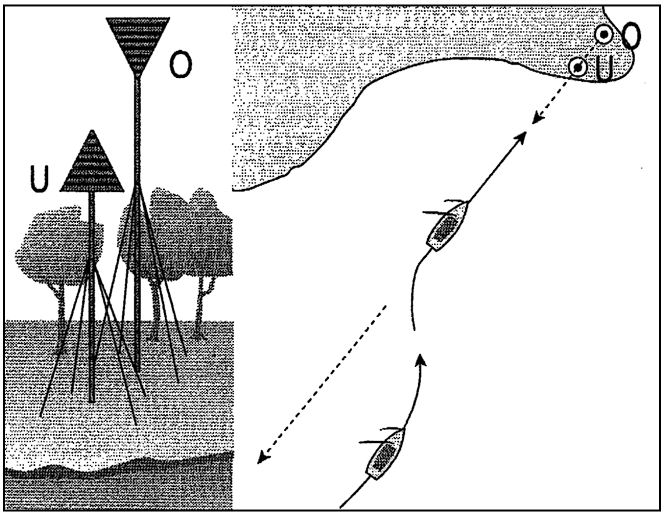


Figure 3.1: Leading lights for nautical navigation. The lower (U) leading light in conjunction with the upper (O) one can be exploited for navigation. Depending on the observed relation the tillerman can deduce how to navigate to reach the destination. Once the tips of the triangles point at each other he has to maintain the current bearing. For further details please see the text. Picture taken from Schult (2008).

that is capable of doing both, extracting task-relevant visual features as well as assigning adequate actions to those, all in a single-step procedure and within one united architecture. The network with winner-take-all-like layers considers goal-relevance of sensory input dimensions, and learns to neglect irrelevant parts of the input. To achieve this, the prediction error δ of the top layer (RL) is not only used to modulate learning of action weights that encode both, value function and action strategy (Q-values), it is also used to adapt the weights of the feature neurons of the lower layer, which are responsible for learning the action-relevant input manifold associated to a specific action. Thus, by *enacting* its world (*cf.* Ch. 2.4.2), the robot is able to identify the visual cues that are relevant for successful navigation. Previously, this type of artificial neural network has been successfully applied to learn action-relevant features of artificial stimuli (Weber and Triesch, 2009). Now we demonstrate for the first time its applicability to a realistic robot scenario.

3.1.1 Biological Inspiration

Evidence for long-term changes of sensorimotor neural representations has been obtained during habit learning in the rat striatum (Jog *et al.*, 1999). The striatum receives direct cortical input and is part of the basal ganglia. Doya (1999) proposed that unsupervised learning happens in the cortex and reinforcement learning in the basal ganglia. Accordingly, the cortex preprocesses data to yield a representation that is suitable for reinforcement learning by the basal ganglia (ibid.).

The seven deep brain nuclei of the basal ganglia are involved in a variety of crucial brain functions (for a recent review please see Chakravarthy *et al.*, 2010) and are tightly linked to the dopaminergic neuromodulatory system, which plays a fundamental role in predicting future rewards and punishment (Schultz *et al.*, 1997; Schultz, 1998). More precisely, the dopamine signal seems to represent the error between predicted future reward and actually received reward (Schultz *et al.*, 1997). This has a direct analogy to the temporal difference error, δ, in reinforcement learning models (Sutton and Barto, 1998), where this error is used to maximize future rewards and avoid punishment. The RL agent interacts with its environment, initially guided by trial and error, seeking to find a mapping between states and actions that will yield the maximal future reward. In other words, it tries to find an optimal motor strategy which is adequate for the given scenario.

However, it remains open how the relevant inputs from the cortex are determined, i.e. which features are read from the cortical activation pattern that are relevant for selecting actions and obtaining rewards. Experiments from Shuler and Bear suggest that RL also occurs in early sensory areas like the primary visual cortex of the rat (Shuler and Bear, 2006). This implies a link between RL and feature learning.

3.2 Scenario

Docking, i.e. navigating towards a predefined position, of a mobile robot is the initial problem that has to be solved before other applications, e.g. grasping, user interaction or recharging can be performed. Therefore, we modeled a general docking situation in a Webots (Michel, 2004) simulation environment (Fig. 3.2). A 3-D geometric shape with several beneficial attributes serves as a landmark, signaling the target region. First, depending on the perspective, it generates a different visual impression (*cf.* Fig. 3.1). From this, the algorithm needs to extract location-specific relevant features and assign them to an adequate action. Next, it can be preprocessed easily. The raw camera image is simply cropped and color-thresholded. After downsizing (32×10 pixels) and a grayscale conversion, it is then directly used as input to the network (Fig. 3.2 C).

In the simulation the robot performs four actions – moving forward, backward, right and left. In one trial a maximum of 25 steps are allowed for reaching the goal. The robot is randomly initialized in a trapezoidal region in front of the target. Two scenarios have been simulated. In the first experiment the robot is only initialized in close proximity to the target, so that the resolution of the visual input is optimal for the extraction of the visual features. It is a common practice to start with easier situations and then gradually move towards more and more difficult ones. Asada *et al.* coined the term "Learning from Easy Missions" (LEM) for this procedure (Asada *et al.*, 1996). Hence, in the second simulation the region is incrementally enlarged during learning to finally span a distance of up to 1.5 m. At a larger distance the robot camera is not able to discriminate the geometrical properties of the stimulus anymore.

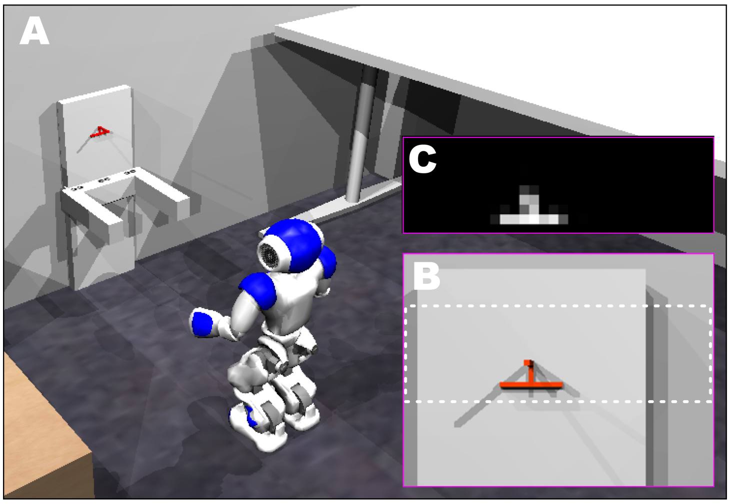


Figure 3.2: Nao robot in front of the docking position. (A) Webots simulation environment representing a domestic situation. (B) Camera view from the robot. The landmark (red) is located within a white dotted rectangle reflecting the region that serves as an input to the network. (C) Input image I after preprocessing.

3.3 Theory

3.3.1 Neural Architecture and Algorithm

For the implementation of our novel robotic navigation paradigm we chose a unified framework that is capable of extracting task-relevant visual features and at the same time learns adequate actions. The model is a two-layer feedforward network (schematically shown in Fig. 3.3) with full connectivity between adjacent layers. The input layer (320 neurons) holds a sensory vector I, representing a 32×10 pixel grayscale image (Fig. 3.2 C). A hidden feature layer (either 4 or 36 neurons) learns visual features within its weight matrix W and encodes this information in a state vector s, which is governed by a softmax activation. In turn, s is mapped via the action weights Q to the output layer (4 neurons) representing the currently selected action a. The choice for the number of hidden feature neurons has been influenced by pragmatic deliberations. To solve the navigation task of the first experiment successfully at least four different states are necessary. For the second, more challenging experiment the number of neurons was increased to 36, governed by the constraint that the algorithm should still be able to run in real-time. In this way, receptive fields are allowed to form that might not have been foreseen by the human designer.

The learning algorithm, which inherits the top-level structure of the SARSA algorithm (Sutton and Barto, 1998), can be summarized as follows (for details and a derivation of the gradient descent learning rule see below, Ch. 3.3.2). At the beginning of each trial, the agent is placed at a random position, with the constraint that the landmark indicating the docking position is within its field of view. The agent reads sensor values I to obtain the (internal) state activation s_j

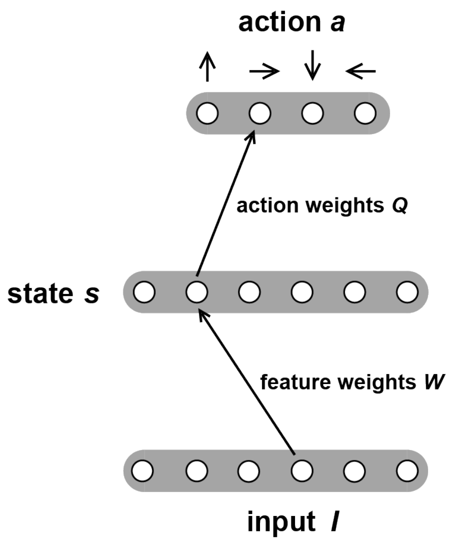


Figure 3.3: Schematic overview of the network architecture. Only one example connection between any two layers is shown.

of neuron j via softmax:

$$h_j = \sum_n W_{jn} I_n \ , \tag{3.1}$$

$$s_j = \frac{e^{\beta^s h_j}}{\sum_k e^{\beta^s h_k}} \tag{3.2}$$

We used a large $\beta^s = 100$ for a winner-take-all-like behavior. Next, an action a_i for neuron i is chosen stochastically (via softmax):

$$h_i = \sum_j Q_{ij} s_j \ , \tag{3.3}$$

$$\texttt{Prob}(a_i = 1) = \frac{e^{\beta^a h_i}}{\sum_k e^{\beta^a h_k}} \ . \tag{3.4}$$

During training we used $\beta^a = 2$ to make the agent explore. For testing $\beta^a = 100$ was chosen to exploit the learned skills. Based on the state activation and on the chosen action the value v is computed:

$$v = \sum_{k,l} Q_{kl} a_k s_l \ . \tag{3.5}$$

The time-discounted (discount factor $\gamma = 0.9$) future value v' and the current value v are used to determine the prediction error δ. A reward $r = 1$ is assigned if the goal position has been reached, otherwise $r = 0$.

$$\delta = r + \gamma v' - v \ . \tag{3.6}$$

Using a δ-modulated Hebbian rule with state s and action a as pre- and post-synaptic values, respectively, the action layer weights Q can be updated:

$$\Delta Q_{ij} \propto \delta\, a_i s_j \; . \tag{3.7}$$

In addition to the normal SARSA algorithm we use the δ signal not only to modulate the learning of the action layer weights, but at the same time to update the feature layer weights, even when no reward is given:

$$\Delta W_{jn} \propto \delta s_j I_n (Q_{ij} \; - \; \sum_k Q_{ik} s_k) \; . \tag{3.8}$$

In each phase of the learning algorithm the feature weights W are rectified to be positive and normalized to length 1, which ensures that a unit that wins for one data point will not also win for all others.

Through the softmax function (Eq. 3.2) the feature layer performs soft competitive learning. Further, the learning progress of the feature weights is modulated by the prediction error of the action layer. On average, only relevant visual stimuli for the given task will contribute to learning. Irrelevant sensory information is usually not correlated with successful actions and, hence, will not influence the learning progress of the agent.

3.3.2 Gradient Descent Learning

To get a better understanding of the learning rules (Eq. 3.7 and 3.8) and to justify their usage, they are derived by performing gradient descent on an energy function[1].

A SARSA agent interacts with its environment and updates its policy based on the actions taken. Due to this interactive process the algorithm is rated as an *on-policy* learning procedure in which the values v, that reflect the expected utility of taking an action a in state s, are estimated. The value function increases towards the goal state, i.e. in our task the final docking position where the agent receives its reward r.

The network parameters can be summarized with $\theta = (Q, W)$, where Q represents the action weights of the upper layer and W the feature weights of the lower layer (Fig. 3.3). Following Sutton and Barto (1998, Chapter 8), the values $v = v(\theta)$ will be adjusted to minimize the mean squared error (MSE) given by:

$$E(\theta) \;=\; \frac{1}{2} \sum_{s,a} P^{\pi}(s,a)(V^{\pi}(s,a) - v(s,a))^2 \; . \tag{3.9}$$

$V^{\pi}(s,a)$ is the "true" value given an action policy π and $v(s,a)$ is the current estimate of the value function. The difference of the two value functions, the prediction error δ (see Eq. 3.6), can be used to improve the estimate v. In practice, the correct value of V^{π} is unknown and we use the information of a better estimate

[1] This section follows Weber and Triesch (2009).

v' obtained in the next time step to approximate it:

$$V^\pi - v \;=\; r + \gamma v' - v \;=\; \delta \; . \tag{3.10}$$

Both, action and sensation, determine the probability distribution $P^\pi(s,a)$ that weighs the prediction error (Eq. 3.9). Due to the vast number of possible states it is usually not possible to obtain a prediction error $\delta = 0$ for all situations. An improved approximation of the value function for a certain state might come at the cost of a worse estimation for other states. This trade-off is defined by the probability distribution $P^\pi(s,a)$ reflecting the frequency with which states are faced by the agent while it interacts with its environment.

Computing the derivative of the energy function E (Eq. 3.9) with respect to the network parameters θ results in an online update regime:

$$\Delta\theta \;\propto\; -\frac{\partial E}{\partial \theta} \;=\; (V^\pi - v)\,\frac{\partial}{\partial \theta} v \;=\; \delta\,\frac{\partial}{\partial \theta} v \; . \tag{3.11}$$

First, we compute the partial derivative with respect to the action layer weights Q. Substituting $v = \sum_{k,l} Q_{kl} a_k s_l$, as given in Eq. 3.5, we yield:

$$\Delta Q_{ij} \;\propto\; -\frac{\partial E}{\partial Q_{ij}} \;=\; \delta\, a_i\, s_j \; . \tag{3.12}$$

To compute the partial derivative of the energy function E with respect to the feature weights W we need a differentiable transfer function on the feature layer. For this purpose, we employ a softmax function (Eq. 3.2), which displays winner-take-all-like behavior for large values of β. For a given activation of action unit i and taking into account that W_{jn} influences the activation of all feature layer units, we can state:

$$\begin{aligned} \Delta W_{jn} \;\;\propto\;\; -\frac{\partial E}{\partial W_{jn}} \;\;&=\;\; -\sum_k \frac{\partial E}{\partial s_k}\,\frac{\partial s_k}{\partial h_j}\,\frac{\partial h_j}{\partial W_{jn}} \\ &=\;\; \delta \sum_k Q_{ik} \frac{\partial s_k}{\partial h_j}\, I_n \; , \end{aligned} \tag{3.13}$$

Nguyen (2006) proposed the following identities for the softmax function:

$$\frac{\partial s_j}{\partial h_j} = s_j(1 - s_j) \tag{3.14}$$

and

$$\frac{\partial s_j}{\partial h_{k,k\neq j}} = -s_k s_j \; . \tag{3.15}$$

Utilizing those identities allows us to define the update rule for the action layer weights W as:

$$\begin{aligned} \Delta W_{jn} &\propto \delta Q_{ij} s_j (1 - s_j) I_n \; - \; \delta \sum_{k, k \neq j} Q_{ik} s_k \, s_j I_n \\ &= \delta Q_{ij} s_j I_n \; - \; \delta \sum_k Q_{ik} s_k \, s_j I_n \\ &= \delta s_j I_n (Q_{ij} \; - \; \sum_k Q_{ik} s_k) \; . \end{aligned} \tag{3.16}$$

The first term in brackets, Q_{ij}, is a residual of the backpropagation procedure. It encodes the strength of the effect of state neuron j on the output. Since all weights tend to be non-negative when positive rewards are given, one might interpret this factor as influencing learning speed, but not the final result. The second term is a summation of the action weights Q_{ik} weighted by the state activations s_k. It can be interpreted as a competitive decay term: if a strong activation of a feature neuron is paired with a large action weight, the learning process is slowed down. Eventually, when a winning feature neuron is found, i.e. only a single unit is active ($s_j = 1$; $s_{k,k \neq j} = 0$), the two terms cancel each other out and learning has converged.

In contrast to the learning rule for the action weights (Eq. 3.12), the update of the feature weights (Eq. 3.16) represents a non-local learning rule, because i) the action layer weights Q are involved and ii) it is summed over all activations of the feature layer. By omitting the non-local terms (aggregated in brackets in Eq. 3.8), we yield a purely local learning rule. This biologically more realistic approximation has been successfully applied to the first experimental scenario presented below. However, for the more difficult task, the modulatory effect on $\delta s_j I_n$ via the non-local terms has been included, mostly because of the endeavor to perform vanilla gradient descent.

3.4 Results

3.4.1 Experiment 1 – Simple Navigation Scenario

In the performed Webots simulations the Nao robot was trained to navigate towards the docking position solely based on visual input. In the first experiment it was placed in close proximity to the docking position and encountered visual input similar to the one shown in the top of Fig. 3.4. After reaching its goal position and receiving a reward about 25 times, the robot was already able to master the simple task successfully in 100 % of the trials. The bottom part of Fig. 3.4 shows the receptive fields of the hidden neurons after 100 training steps. The visual features relevant for determining its state and for performing effective navigation have been extracted and stored in the weights connecting the input with the hidden state neurons. In the receptive field depicted in the lower right no structure has evolved. This is due to the fact that i) the state space can be

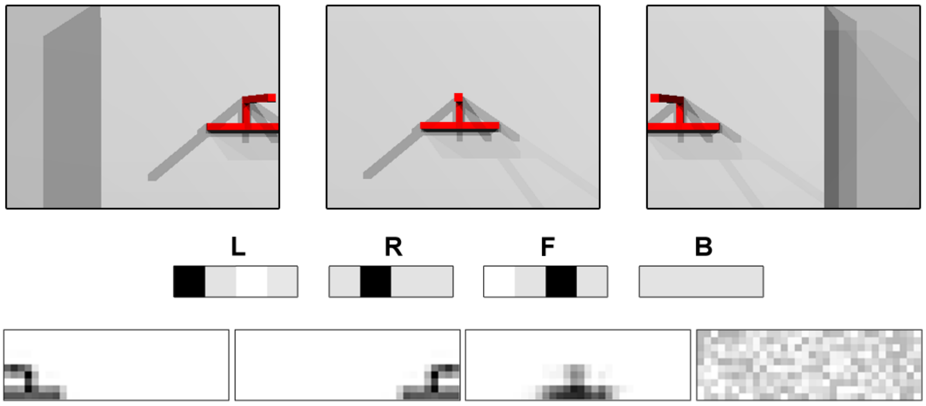


Figure 3.4: Raw visual input (top), receptive fields of action neurons (middle) and hidden feature neurons (bottom) after 100 steps of training. The raw camera image (top) shows three exemplary situations the robot might encounter (corresponding to a robot position left, in front and to the right of the landmark). The hidden neurons (bottom) code for a specific state, which is then mapped correctly via the action weights Q (middle) to an adequate action, e. g. moving left (L), right (R), forward (F) or backward (B). The action weights for the backward action show no structure, because the backward action is hardly ever executed in this simple scenario. Correspondingly, one hidden unit that is not used by any action unit has no structure. Strong weights are displayed dark.

covered completely with the three states captured in the other RFs and ii) the backward action is hardly ever executed in the simple scenario.

3.4.2 Experiment 2 – Complex Navigation Scenario

In the second simulation the possible initialization region of the robot was gradually increased, and due to the vastly growing state space, a much harder problem was preserved. Nevertheless, after training (2000 trials, $\approx$ 2 days[2]) the humanoid robot was able to reach the goal position in 95 % of the cases (690 out of 725 trials). Testing was performed on newly initialized simulation trials, stressing the generalization potential of this approach. The network is capable of identifying the relevant visual features, as shown by the evolved receptive fields (RFs) of the feature and action layer in Fig. 3.5 and to generate task-specific sensorimotor laws needed for navigation. However, these features do not clearly reflect the shape of the landmark anymore. Due to the large variations of the landmark's position, scale and perceived shape, the network is not capable of representing all combinations. Therefore, now not only a single state is linked to a specific action, but a mixture of different ones (Fig. 3.5 top). This "population" coding might be useful for resolving ambiguities. Note, the predominant visual feature for a specific action can still be recognized in the receptive fields (Fig. 3.5 bottom, RFs framed in red).

[2]The combination of the Webots simulator with the Aldebaran Nao API runs in real-time only.

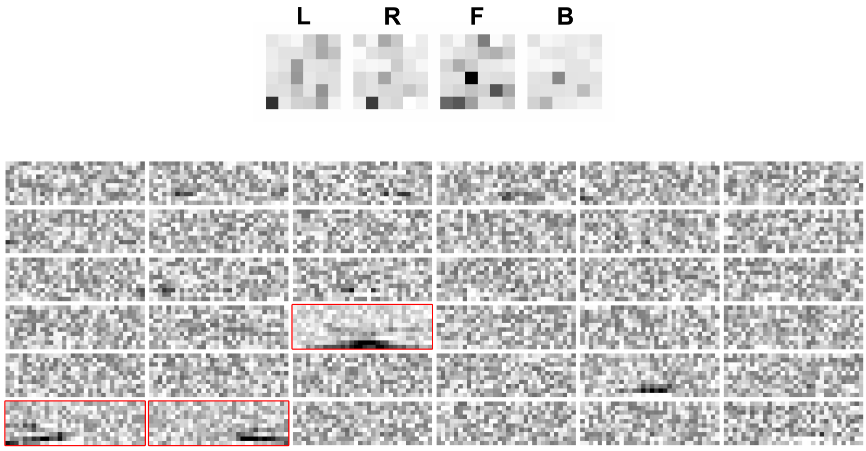


Figure 3.5: Receptive fields of hidden feature neurons (bottom) and action neurons (top) after 2000 steps of training. The RFs of the action units correspond to left (L), right (R), forward (F) and backward (B) movement. The RFs of the feature neurons that have the strongest contribution on the action units for L,R & F are framed in red. Strong weights are displayed dark.

In Fig. 3.6 sample trajectories of the robot are shown. Green trajectories were successful trials, whereas the red ones represent failures. Note that an identical initialization point can result in a completely different trajectory. This is mainly due to noise in movements imposed by the Webots simulator to reflect real-world robot behavior. Furthermore, this noise may lead to a rotation of the robot, which is currently not compensated, because no rotation movements are implemented. This is actually the reason for most of the unsuccessful trials (red trajectories, Fig. 3.6).

After training, receptive fields are linked to specific actions, jointly composing sensorimotor laws. These laws capture the relationship of actions and perceptions that is necessary for goal-directed navigation behavior. If the robot is confronted with a (previously unseen) input, the winning feature unit is the one where the receptive field is most similar to the current input. Hence, the properties (e.g. shape) of this input will trigger the movement embedded in the sensorimotor law. Hence, the perceived shape reflects the robot position in relation to the goal position.

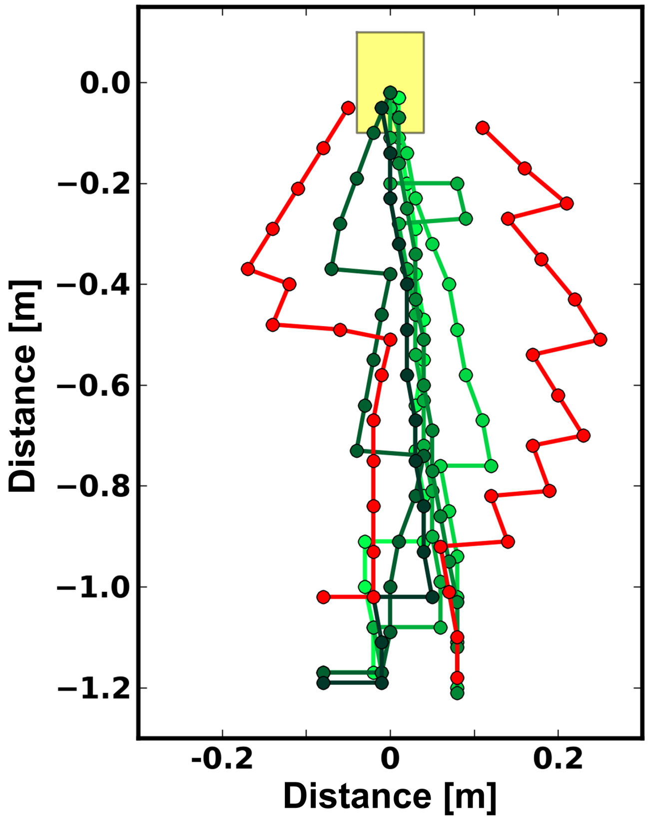


Figure 3.6: Sample trajectories of the robot. Different shades of green represent successful trials and red failures. The yellow docking region measures 8×20 cm.

3.5 Summary

We presented a Webots simulation of a Nao robot that learns to navigate towards a virtual target. Instead of relying on a pre-given or acquired world model, our approach allows a robot to navigate to a target position in an unknown environment, once it has learned what visual features it has to attend to and how those relate to its own actions. Learning of these sensorimotor laws is accomplished using a two-layer network, integrating feature and motor learning, all in a single-step procedure. A 3-D geometrical shape served as a landmark, which led to perspective distortions depending on the robot's position and locomotion. This relationship is learned by the robot, enabling it to successfully reach its target in 95 % of the trials. The results of this chapter have been published (Kleesiek *et al.*, 2011).

3.5.1 Connection to the Other Experiments of this Thesis

In this first experiment, the focus lay on the methodology of this special type of artificial neural network. The two-layer architecture combines a sensory layer with a reinforcement learning layer. This allows the robot to learn in a 'top-down' *action-driven* way which visual features it has to attend to if it wants to master the given navigation task successfully. The RL prediction error not only leads to a state-action mapping; it is also essential for the receptive field development of the sensory layer. In conjunction, the RFs and the state-action mapping represent the *sensorimotor laws* that enable the agent to perform goal-directed behavior.

As in all the conducted studies, the choice of the architecture and the design of the experiment have been influenced by *sensorimotor principles* rooted in the research field of embodied cognition (Ch. 2). Perception is a *sensorimotor experience* and the interaction of the robot with its world results in *information self-structuring*. In this whole process, the *world serves as an outside memory* and the agent just has to learn how its own actions relate to its own sensations.

The knowledge gained during the implementation of the two-layer architecture of this experiment paved the way for the development of the novel bio-inspired model presented in the next chapter.

We not only see, but we look for,
not only hear, but we listen to.
Robert S. Woodworth

Chapter 4
Sigma-Pi Reinforcement Learning for Reaching

4.1 Introduction

Given an object at some location in space, how can we successfully reach for it? We tackle this demanding problem and present a simulation of an artificial agent that learns reaching for a target. The RL agent interacts with its environment, initially guided by trial and error, trying to find an optimal (motor) strategy which is adequate for the current situation. Often, the agent is confronted with invariant situations, e.g. the same relative position between hand and object, where an identical motor response is appropriate (at least in a servo-driven robot). It is therefore desirable for the agent to identify those invariances and combine them to a unique state. A natural way to achieve this is to use *Sigma-Pi* neurons (Softky and Koch, 1995). By multiplying two inputs of a Sigma-Pi network, e.g. hand and object positions, it is possible to detect co-activations of the input units. An additional summing over these units results in a single output node that responds to the manifold of co-activated input units, i.e. it captures the same relative distance between hand and object.

Here we propose a novel architecture which is capable of learning both, invariant hand-object relations and the movement of the hand towards the target, in a single-step procedure. In this architecture the prediction error δ of the top layer is not only used to modulate learning of action weights that encode both, value function and action strategy (Q-values), it is also used to adapt the weights of the Sigma-Pi neurons of the lower layer. In this way, the agent learns in an action-driven approach the relevant sensory input manifold as well as an associated action, i.e. the sensorimotor laws linking the situation to an appropriate movement. Furthermore, the suggested model performs an implicit coordinate transformation combining inputs from the same or different sensory modalities. Previously, Weber and Wermter (2007) were able to show that Sigma-Pi units can perform a frame of reference transformation using unsupervised learning.

Next to the presentation and evaluation of this novel artificial neural architecture we also compare it with two related learning methods devised to solve the given reaching task. One of the models of comparison is an RL architecture with a cubic Q-table that links all possible hand and object positions directly to an action. The other model resembles the classical two-step approach usually used to tackle such a task in the literature. First, an unsupervised method is employed that self-organizes the possible hand-object invariances. Then, in a separate step, once this unsupervised learning of the state space has been completed, an RL algorithm follows.

4.1.1 Biological Inspiration

For successful reaching, coordinate transformations are necessary. Depending on the orientation of the body, the arm and eyes, and their relation to each other, the object falls on different locations on the retina. To cope with this vexing problem, at least one stable frame of reference (FOR) is needed that is invariant to changes in position of the other involved body parts. For reaching, the upper body serves as such a stable FOR. The limb is physically connected to it and, hence, all movements are constrained given this anatomical relation. To facilitate the coordination of movements (Cohen and Andersen, 2002), hand and object-coordinates have to be converted in a common FOR (Snyder, 2000; Crawford *et al.*, 2004).

Several brain areas have been identified that are involved in the process of coordinate transformations. The ventral intraparietal area (VIP) encodes visual and somatosensory information (Duhamel *et al.*, 1998; Sereno and Huang, 2006). This structure is highly interconnected with the lateral intraparietal area (LIP), the parietal reach region (PRR), premotor areas (PMA) and the anterior intraparietal area (AIP). AIP not only responds during the act of grasping but also during the passive viewing of graspable objects (Valyear *et al.*, 2007). The head movement is influenced by PMA (Duhamel *et al.*, 1997) and information about the preparation for reaching is integrated in the PRR (Batista *et al.*, 1999). LIP, on the contrary, preferentially encodes information about an upcoming eye movement towards a target (Andersen *et al.*, 1990; Snyder *et al.*, 1997).

It has been shown that different frames of reference can be represented within a single area, apparently adjusted to the specific function of the particular region. For instance, it is suggested that parts of VIP and LIP encode the position of a visual target in both, head-centered as well as eye-centered, coordinate systems (Duhamel *et al.*, 1998; Mullette-Gillman *et al.*, 2005). Further, it has been found that neurons of the PRR not only represent the position of a visual target in eye-centered coordinates (Batista *et al.*, 1999), but also encode the distance of this target to the current position of the hand (Chang *et al.*, 2009). It has been

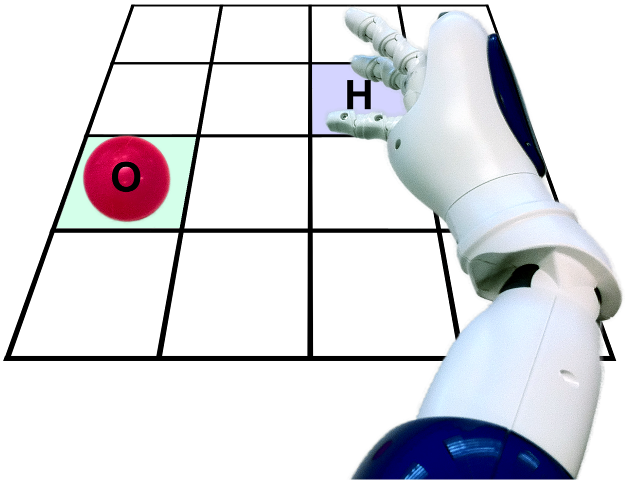


Figure 4.1: Reaching Scenario. The 2-D world is discretized. Hand (H) and object (O) are located at specific positions within this grid. The task for the simulated agent is to reach for the object with a single movement. If hand and object are both shifted by an identical displacement vector, their relative position with respect to each other does not change. Hence, another goal of the agent is to identify those invariant positions.

speculated whether this information serves as an error signal for the reaching movement (ibid.). Neurons with a comparable function have also been found in the posterior parietal cortex, an area adjacent to the PRR (Buneo *et al.*, 2002).

Using *gain modulation*, the coordinate transformation can be computed implicitly (Blohm and Crawford, 2009; Chang *et al.*, 2009). In short, gain modulation is a way of combining several sources, e.g. sensor and motor information, in a nonlinear way (Salinas and Abbott, 2001). Pouget *et al.* (2002) showed that this combination can indeed be realized using a multiplication, encouraging us to use Sigma-Pi neurons for our architecture.

4.2 Scenario

An overview of the simulated scenario can be seen in Fig. 4.1. The 2-D world is modeled as a grid of size $n \times m$. In this discretized situation, it is assumed that the position of hand and object are known. However, it is not important from which modality, e.g. visual or proprioceptive, this information comes. Also the frame of reference, e.g. eye-centered or hand-centered coordinates, in which this information is encoded in is irrelevant for the presented algorithms.

The agent can perform actions in this 2-D grid world. Its movements are encoded as relative changes to the current position of the hand. This is inspired by the control methods of the application programming interface (API) of the Nao humanoid robot. The function `changePosition()` allows an end effector, e.g. the right hand, to be moved in Cartesian space relative to its current position. For this purpose, a movement vector is required as input to the function. The agent is able to perform actions corresponding to all possible movement vectors of a world with a given size, i.e. it can move from its current position to any destination within the grid.

During training the target object is placed at a random position and the agent is allowed to perform movements until it successfully reaches the target. Of course, this initially very frequently results in movements that cause the end effector to leave the specified area. Consequently, to avoid this behavior a negative reward penalizes those movements (see below) and thereafter the object and hand are reinitialized at random positions. The task of the agent is to learn to reach for the object with a single movement. In this context, it is confronted with invariant situations where hand and object have the same relative distance to each other. Thus, another goal for the agent is to identify those invariant situations during learning.

4.3 Theory

In this section theories of three different architectures that are capable of solving the given reaching task are presented. First, we introduce our novel unified architecture and learning algorithm. Next, a reinforcement learning scheme based on a cubic Q-table is elucidated. Those first two architectures are both suited to learn the given task in a one-step procedure. In contrast, the last paradigm that will be explained is a two-step approach. Initially, the possible hand-object relations are self-organized using an unsupervised learning method (Weber and Wermter, 2007). Afterwards RL is employed to link the obtained invariances to suitable actions.

4.3.1 Unified Neural Architecture and Algorithm

The structure of this novel unified algorithm is similar to the one employed for the navigational task of Ch. 3.3. Again, the model is a two-layer feedforward network with full connectivity between adjacent layers (Fig. 4.2). In contrast to the previous architecture the network now has two input layers H and O, each consisting of a binary vector encoding the position of the hand and the object, respectively. In principle, these vectors do not have to be equal in size. However, for the experiments presented here their size was identical.

The two input vectors are combined to a hidden layer via the cubic weights of the Sigma-Pi neurons. The size of the hidden layer, i.e. the number of invariant situations (states), depends on the number of possible unique movements the agent can perform, which in turn depend on the size of the world. For a world of size $n \times m$ this can be determined by

$$N_{states,actions} = (2n-1)(2m-1)\ . \tag{4.1}$$

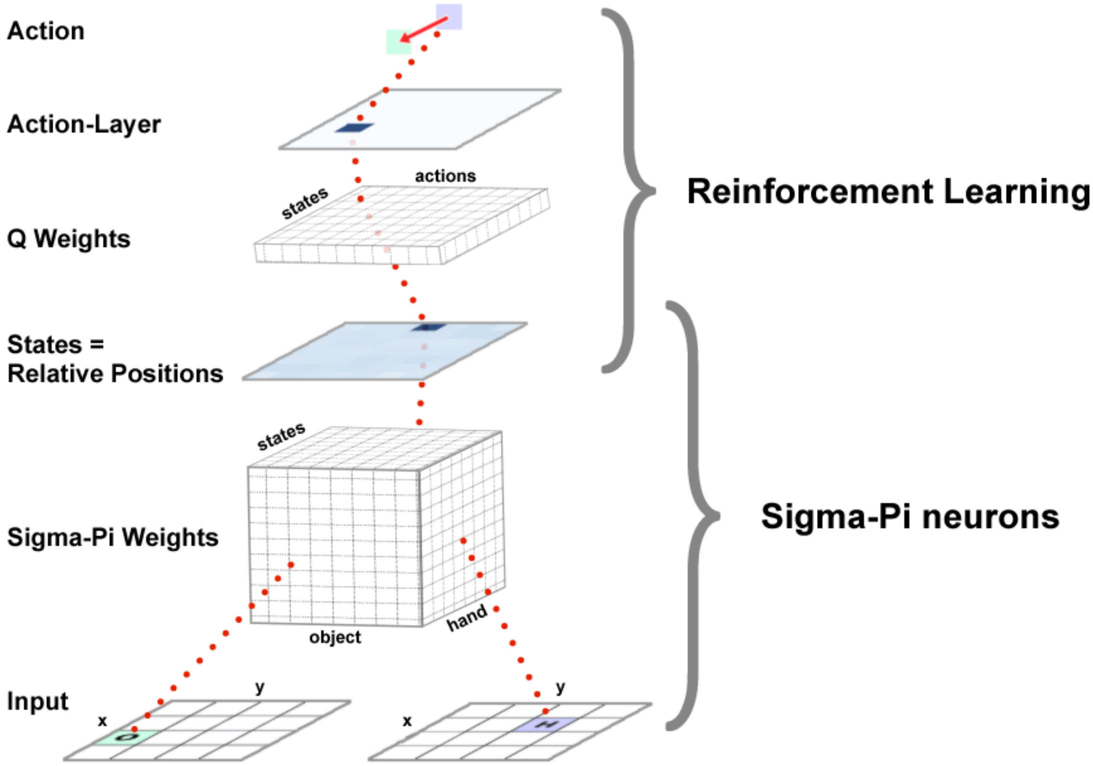


Figure 4.2: Neural architecture consisting of Sigma-Pi neurons and RL layer. The two-layer feedforward network architecture has two binary input layers, one representing the position of the hand (H), the other the location of the object (O). This information is combined using Sigma-Pi neurons to generate an activity in the hidden state layer, signaling the relative position of hand and object. Using RL this state is linked to an action. During training, the RL prediction error δ is not only used to update the Q weights. Additionally, it also modulates the weights of the Sigma-Pi neurons. Please see text for details. Only one example connection between any two layers is shown.

Influenced by a softmax activation, the information of the hidden layer is encoded in a state vector s that in turn is mapped via the action weights Q to an output layer of size $N_{states,actions}$, representing the currently selected action a.

The learning algorithm inherits the top-level structure of the Q-Learning algorithm (Sutton and Barto, 1998) and can be summarized as follows (for the derivation of the gradient descent learning rule, please refer to Ch. 3.3.2). At the beginning of each trial, hand and object are placed at random positions. Based on the sensor information H and O the state activation s_j of neuron j can be obtained via softmax:

$$h_j = \sum_{n,m} W_{jnm} H_n O_m \; , \tag{4.2}$$

$$s_j = \frac{e^{\beta^s h_j}}{\sum_k e^{\beta^s h_k}} \tag{4.3}$$

We used $\beta^s = 10$ for training to facilitate the learning of similar states. Next, an action a_i for neuron i is chosen stochastically (via softmax):

$$h_i = \sum_j Q_{ij} s_j \ , \tag{4.4}$$

$$\texttt{Prob}(a_i{=}1) = \frac{e^{\beta^a h_i}}{\sum_k e^{\beta^a h_k}} \ . \tag{4.5}$$

During training we used $\beta^a = 2$ to make the agent explore. Based on the state activation and on the chosen action the value v is computed:

$$v = \sum_{k,l} Q_{kl} a_k s_l \ . \tag{4.6}$$

A reward $r = 1$ is assigned if the target has been reached. If an invalid action has been performed, i.e. the hand has been moved outside the specified region, the agent is punished with a slightly negative reward value $r = -0.1$. In all other situations no reward is given ($r = 0$). The current value v and the reward r are used to determine the prediction error δ.

$$\delta = r - v \ . \tag{4.7}$$

Using a δ-modulated Hebbian rule with state s and action a as pre- and post-synaptic values, respectively, the action layer weights Q can be updated:

$$\Delta Q_{ij} \propto \ \delta \, a_i s_j \ . \tag{4.8}$$

In addition to this regular RL-update we use the δ signal to modulate the weights of the Sigma-Pi layer. For this purpose, we first compute the outer product I of the two sensory vectors:

$$I = HO^T \ . \tag{4.9}$$

This can be used subsequently to determine the weight update:

$$\Delta W_{jnm} \propto \delta s_j I_{nm} (Q_{ij} \ - \ \sum_k Q_{ik} s_k) \ . \tag{4.10}$$

In any case (positive, no reward or punishment) the weights of both layers are updated according to Eqs. 4.8 and 4.10. If the agent has been punished or it has successfully reached for the object, the simulation is reinitialized after this update. Otherwise the agent is allowed to continue with its movements until one of the conditions (failure or success) is met. For testing, $\beta^a = \beta^s = 100$ was chosen to exploit the learned skills.

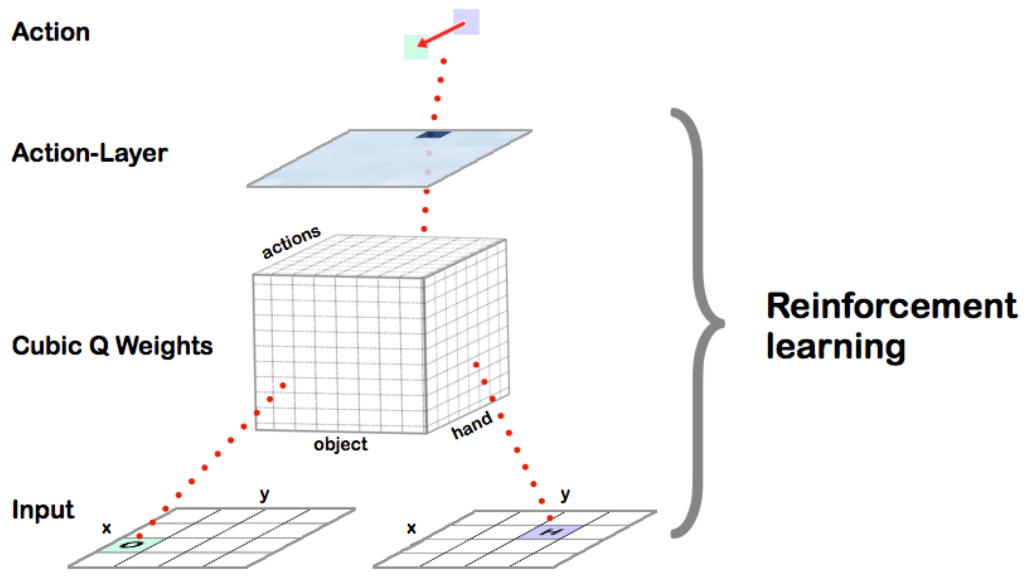


Figure 4.3: Single-layer neural architecture consisting of a cubic RL layer. This single-layer RL architecture has two binary input layers, one representing the position of the hand (H), the other the position of the object (O). This information is combined and directly linked to an action using a cubic Q-table. During training, the RL prediction error δ is used to adjust the weights of this Q-table. Please see text for details. Only one example connection between any two layers is shown.

Adaptive Learning

Not only Eqs. 4.2, 4.7, 4.9 and 4.10 differ from the previous algorithm. Further, an adaptive learning rate η is introduced for each individual neuron of the action layer. This method is loosely inspired by the resilient propagation algorithm of Riedmiller and Braun (1993). Based on the prediction error δ of the current and the last time step the learning rate is either increased by a factor $\xi^+ > 1$ or decreased by a factor $\xi^- < 1$:

$$\eta_{ij}(t) = \begin{cases} \mathbf{max}(\eta_{ij}(t-1)\cdot\xi^-, \eta_{min}) & \text{if } \delta(t-1)\cdot\delta(t) < 0, \\ \mathbf{min}(\eta_{ij}(t-1)\cdot\xi^+, \eta_{max}) & \text{if } \delta(t-1)\cdot\delta(t) > 0, \\ \eta_{ij}(t-1) & \text{else.} \end{cases} \tag{4.11}$$

For our experiments the values of the individual learning rates were allowed to be in the interval $[0.0001, 1.0]$. Depending on the value of the prediction error they were either increased with $\xi^+ = 1.05$ or decreased by a factor $\xi^- = 0.95$.

4.3.2 RL Architecture with Cubic Q Weights

To compare the model presented in Ch. 4.3.1 to a vanilla RL architecture with cubic Q weights (Fig. 4.3), the following algorithm that omits the Sigma-Pi layer was designed. Based on the sensor information H and O the activation h_k for neuron k can be obtained, which in turn directly allows an action a_k to be chosen stochastically (via softmax):

$$h_k = \sum_{ij} Q_{ijk} H_i O_j \ , \tag{4.12}$$

$$\texttt{Prob}(a_k\!=\!1) = \frac{e^{\beta^a h_k}}{\sum_l e^{\beta^a h_l}} \ . \tag{4.13}$$

To facilitate exploration $\beta^a = 2$ is used during training, while for exploitation $\beta^a = 100$ was chosen. Given the action a_k the value v can than be selected directly from the Q weights:

$$v = Q_{ija_k} \ . \tag{4.14}$$

The reward r is assigned as previously and in combination with v used to determine the prediction error

$$\delta = r - v \ . \tag{4.15}$$

Using a δ-modulated Hebbian rule the action layer weights Q can be updated according to:

$$\Delta Q_{ijk} \propto \delta \, H_i O_j a_k \ . \tag{4.16}$$

4.3.3 Unsupervised Learning Followed by RL

Another possibility for learning this challenging reaching task is to subdivide the learning procedure into two neural architectures that are trained consecutively (Fig. 4.4). This is a common strategy in the literature (Legenstein *et al.*, 2010). Initially, the state space is acquired using an unsupervised learning method. This procedure is than followed by canonical reinforcement learning (Sutton and Barto, 1998).

For unsupervised learning we employ an online algorithm with an incremental weight update. While the agent explores the world it seeks for responses that are invariant to variations of input pairs, i.e. hand and object locations, which share the same relative position. In each learning step the algorithm needs two input pairs that are supposed to have an identical output activation (because they share the same relative distance and therefore should result in the same motor response). The first pair is used to obtain an initial neural activation pattern and serves during the weight update as a post-synaptic learning term. A winner unit is determined on which a Gaussian-profile activation function is centered to realize lateral competition and topographic relations between neighboring neurons. The second pair of inputs acts as a pre-synaptic learning term by forming a difference between the data and the weights that are supposed to be updated.

In each iteration of the training algorithm hand and object are placed at random positions. Based on the sensor information H and O the activation h_i of neuron i can be computed:

$$h_i = \sum_{j,k} W_{ijk} H_j O_k \ . \tag{4.17}$$

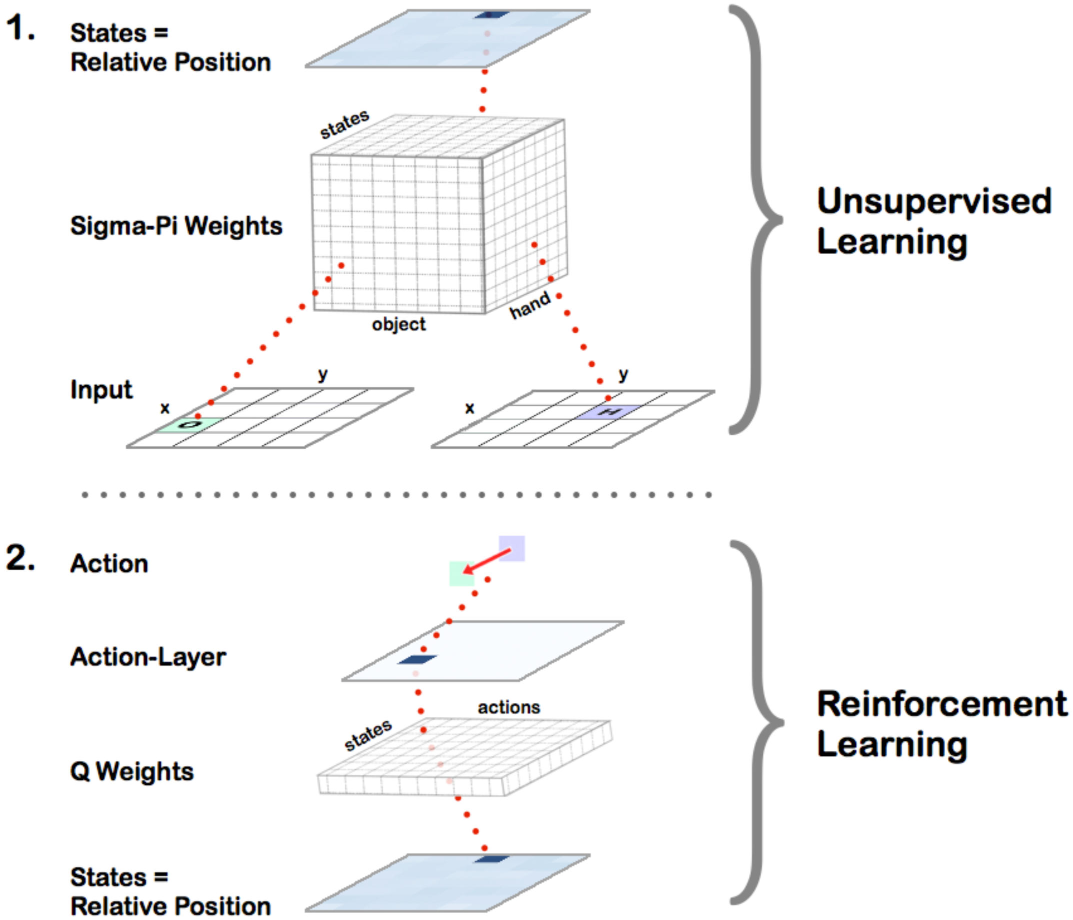


Figure 4.4: Two separate neural architectures. First, unsupervised learning is used to self-organize the state space of possible hand-object relations. Once this learning has been completed, a second RL algorithm links the acquired knowledge of the first procedure to suitable actions. Please see text for details. Only one example connection between any two layers is shown.

Next, we determine the winning unit l:

$$l = argmax(\boldsymbol{h}) \ , \tag{4.18}$$

and its index m. A Gaussian-profile activation function is centered at this position to subsequently weigh the activation of the unit itself and its surround. Next to facilitating topographic relations of neighboring units this ruse also results in lateral competition. The Gaussian function is given as:

$$\begin{aligned} \mathcal{N}(i|m,\sigma) &= \frac{1}{\sum_j e^{-\frac{(j-m)^2}{2\sigma^2}}} \cdot e^{-\frac{(i-m)^2}{2\sigma^2}} \\ &= \frac{1}{Z} \cdot e^{-\frac{(i-m)^2}{2\sigma^2}} \ . \end{aligned} \tag{4.19}$$

The partition function Z normalizes the Gaussian activation function to 1, thereby assigning the highest value to the position of the winning unit determined in Eq. 4.18. This weighted winner scheme is supposed to lead to a reduction of noise, especially at boundaries (Luttrell, 1994). The interaction range of neighboring units is governed by σ. Starting at 2.0 the value of σ is linearly reduced to 0.01 during learning. Next to a linear decay we experimented with other, mostly exponential, forms of decay. None of them proved to be superior to the simple linear regime eventually employed in our experiments.

Multiplying the Gaussian function (Eq. 4.19) centered on the winning unit with the activated neurons $\boldsymbol{h}$ leads to a weighted neural activation:

$$s_i = h_i \cdot \mathcal{N}(i|m, \sigma) \ . \tag{4.20}$$

In the next step a second input pair, $\bar{H}$ and $\bar{O}$, is randomly chosen from the set of all possible hand and object pairs that preserve the same relative distance as the initial input pair. Having acquired the sensor readings of the new positions allows the incremental weight update to be performed:

$$\Delta W_{ijk} \propto s_i(\bar{H}_j \bar{O}_k \ - \ W_{ijk}) \ . \tag{4.21}$$

Note, we provide the algorithm with input pairs sharing the same relative distance. However, this information is strictly used to generate input to the network and has no influence on its output $\boldsymbol{s}$, which self-organizes during this unsupervised learning procedure. This implementation can be interpreted as a form of spatial coherence, a phenomenon which is often exploited for learning in biological organisms.

Once the state space has been learned, regular reinforcement learning follows in a separate procedure (Fig. 4.4). First, an action a_i for neuron i is chosen stochastically (via softmax):

$$h_i = \sum_j Q_{ij} s_j \ , \tag{4.22}$$

$$\texttt{Prob}(a_i = 1) = \frac{e^{\beta^a h_i}}{\sum_k e^{\beta^a h_k}} \ . \tag{4.23}$$

During training we use $\beta^a = 10$ to make the agent explore. Based on the state activation s and on the chosen action a the value v is computed:

$$v = \sum_{k,l} Q_{kl} a_k s_l \ . \tag{4.24}$$

A reward $r = 1$ is assigned if the target has been reached. If an invalid action has been performed, i.e. the hand has been moved outside the specified region, the agent is punished using a slightly negative reward value $r = -0.1$. In all other situations no reward is given ($r = 0$). The current value v and the reward r are

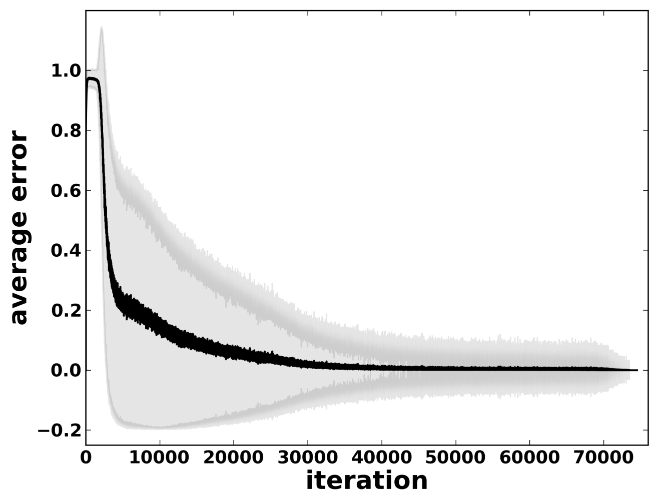


Figure 4.5: Average (1000 runs) error for a 3×3 world. The network is trained with the novel unified architecture (Ch. 4.3.1). The mean squared error decreases monotonically. Only the errors of movements that actually reach the target are considered. Invalid actions, as well as unsuccessful ones, are left out. The gray shaded area delineates ±1 STD.

used to determine the prediction error δ.

$$\delta = r - v \ . \tag{4.25}$$

Using a δ-modulated Hebbian rule with state s and action a as pre- and post-synaptic values, respectively, the action layer weights Q can be updated:

$$\Delta Q_{ij} \propto \delta \, a_i s_j \ . \tag{4.26}$$

4.4 Results

4.4.1 Unified Neural Architecture and Algorithm

To test the general functioning of this novel biologically inspired reaching model (Ch. 4.3.1), scenarios of different sizes have been evaluated. For an initial impression of how the system behaves, a 3×3 world has been trained 1,000 times. The average mean squared error (MSE) of successful iterations, i.e. movements reaching the target, is plotted in Fig. 4.5. It decreases monotonically.

Letting the weights of the system converge, always leads to perfect reaching movements of the agent. Further, all invariances, identical relative positions of hand and object, are identified flawlessly. As an example, the evolution of the Sigma-Pi and Q weights of a 4×4 world are shown at different time points (Fig. 4.6). Initially, no structure can be seen in the receptive fields of the Sigma-Pi weights nor is any state-action mapping visible. As training goes on, invariant sensory situations and corresponding actions evolve. Interestingly, the Q weights of the action-layer develop slightly faster. This stresses the importance of action-driven 'top-down' guidance for sensory receptive field development. There is no 'bottom-up' mechanism like temporal or spatial coherence, e.g. slowness or topography (*cf.* unsupervised learning model, Ch. 4.3.3), for generating the sensory invariances. Instead, actions shape the manifold of hand-object relations.

If hand and object are both shifted by an identical displacement vector, their relative position with respect to each other does not change. This is displayed in Fig. 4.7 for an exemplary invariant hand-object relation in a 4×4 world after training. No matter where the hand-object relation occurs it always leads to the activity of the same neuron in the hidden layer. In turn, this state activation is always linked to the same appropriate motor response.

In general, it has been shown that the novel unified neural architecture learns the given task accurately. But how do changes of the algorithm or the parameters influence the learning? For example, if no mild punishment is given, the model is not able to capture all invariances. Especially the extreme movements, e.g. the object is located in the lower left corner and the hand in the upper right corner, are impaired. This is probably because they are more likely to result in a movement that ends outside of the permitted region.

Omitting the non-local terms (aggregated in brackets in Eq. 4.10) leads to a pure local learning rule. In contrast to the algorithm used to learn the navigational task of Ch. 3.3, it is not possible this time to yield satisfying behavior when training the network with this biologically more realistic approximation. This emphasizes the non-local modulatory effects of the action layer weights Q on the sensory learning.

Yet another result emphasizing the importance of actions for learning of perceptions comes from the augmentation of the algorithm with the adaptive learning rate (Eq. 4.11). Using this modification for the update of the Q weights leads to a 10-fold increase in speed. The described adaptive learning rate basically results in a cautious learning in the beginning that allows the state-action mapping to evolve slowly. A similar adaptive update regime was introduced for the lower sensory-layer. However, no significant improvements in learning speed have been observed and hence, to avoid the computational overhead (time and memory), it was not included in the final version of the algorithm.

4.4.2 RL Architecture with Cubic Q Weights

Again, we start off looking at the MSE during training. For this purpose the single-layer RL model with cubic weights (Ch. 4.3.2) has been trained on a 4×4 world 1,000 times. The average error trace of movements that actually reach the target is shown in Fig. 4.8. Although the world is one order of magnitude larger than in the previous experiment (*cf.* Fig. 4.5), the error decreases more rapidly. Further, the standard deviation, drawn in gray, is also smaller. On average 123,358 ($\pm$4,106) steps were necessary for the learning algorithm to converge. On a contemporary desktop computer this takes about 70 s per training run.

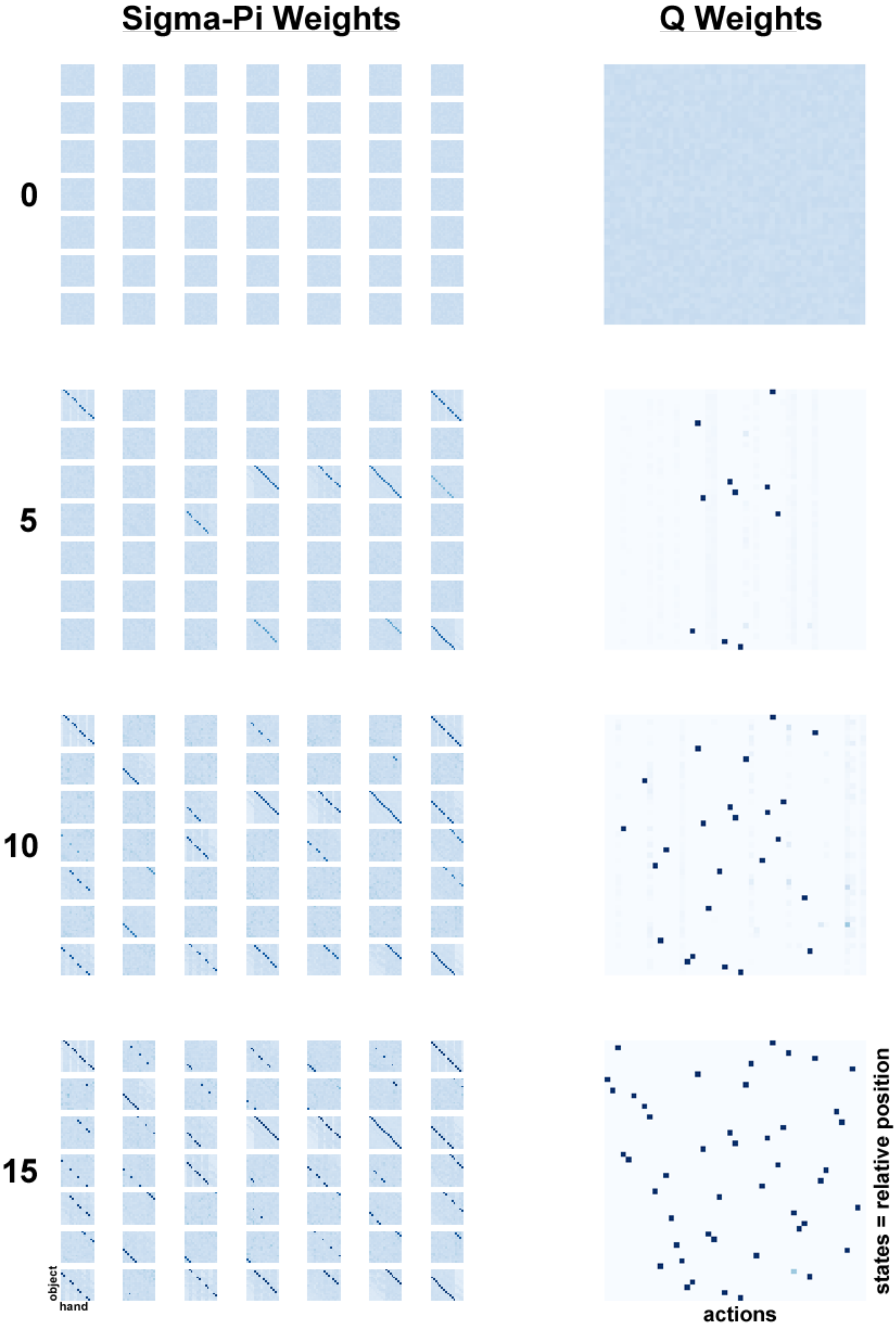


Figure 4.6: Development of Sigma-Pi RFs and Q weights at different time points (0, 5, 10 and 15 $\times 10^6$ iterations). Strong weights are displayed dark.

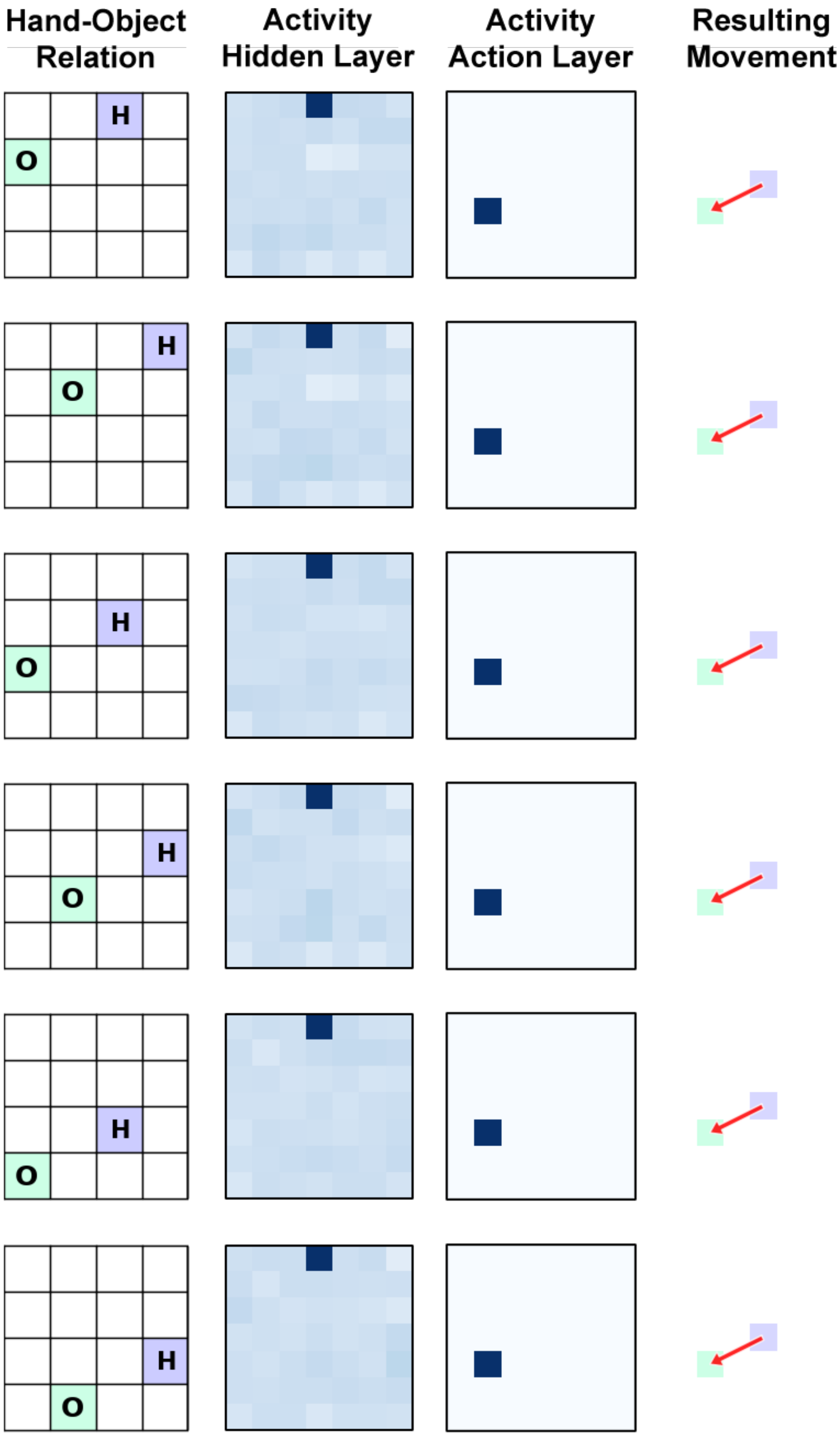


Figure 4.7: Invariant hand-object relations and the resulting movements after learning with the unified architecture. For all shown invariances the identical neurons of hidden and action layers are active. Strong activations are displayed dark.

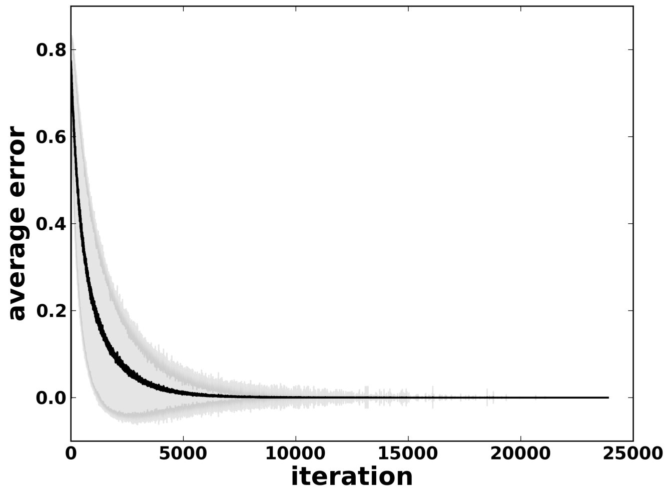


Figure 4.8: Average (1000 runs) error for a 4×4 world. The network is trained using cubic RL (Ch. 4.3.2). The mean squared error decreases monotonically. Only the errors of movements that actually reach the target are considered. Invalid actions, as well as unsuccessful ones, are left out. The gray shaded area delineates ±1 STD.

Also with this cubic RL architecture the virtual agent is able to learn perfect reaching behavior. As an example, the evolution of the cubic Q weights of a 4×4 world is shown at different time points (Fig. 4.9). Initially, no structure can be seen in the receptive fields of the cubic Q weights. However, the structure evolves swiftly. In contrast to the RFs of the unified architecture, the receptive fields obtained during learning with cubic RL are ordered and symmetric. At the center the RF of the state is located that represents the most invariances. This is actually the state where the hand has successfully reached the object and, thus, no further movement is required. Around this RF the other receptive fields are arranged circularly, thereby decreasing with the number of invariant states. Due to the properties of the world, this means at the same time that the distance of the corresponding correct movements increases from the center to the outside. This phenomenon can be explained by properties of the cubic RL architecture. Instead of an action-driven 'top-down' guidance of sensory receptive field development, as it has been seen by the unified architecture (Ch. 4.4.1), the structure of the state space is implicitly incorporated in the learning algorithm. This is realized by including the outer product between the two sensory vectors, encoding hand and object position, during the update of the Q weights (Eq. 4.16).

4.4.3 RL Following Unsupervised Learning

The third algorithm we devised to solve the given reaching task consists of two separate architectures. In a first step the state space, i.e. the hand-object invariances, is self-organized using unsupervised learning. This procedure is followed by classical reinforcement learning (*cf.* Ch. 4.3.3). As for the other two

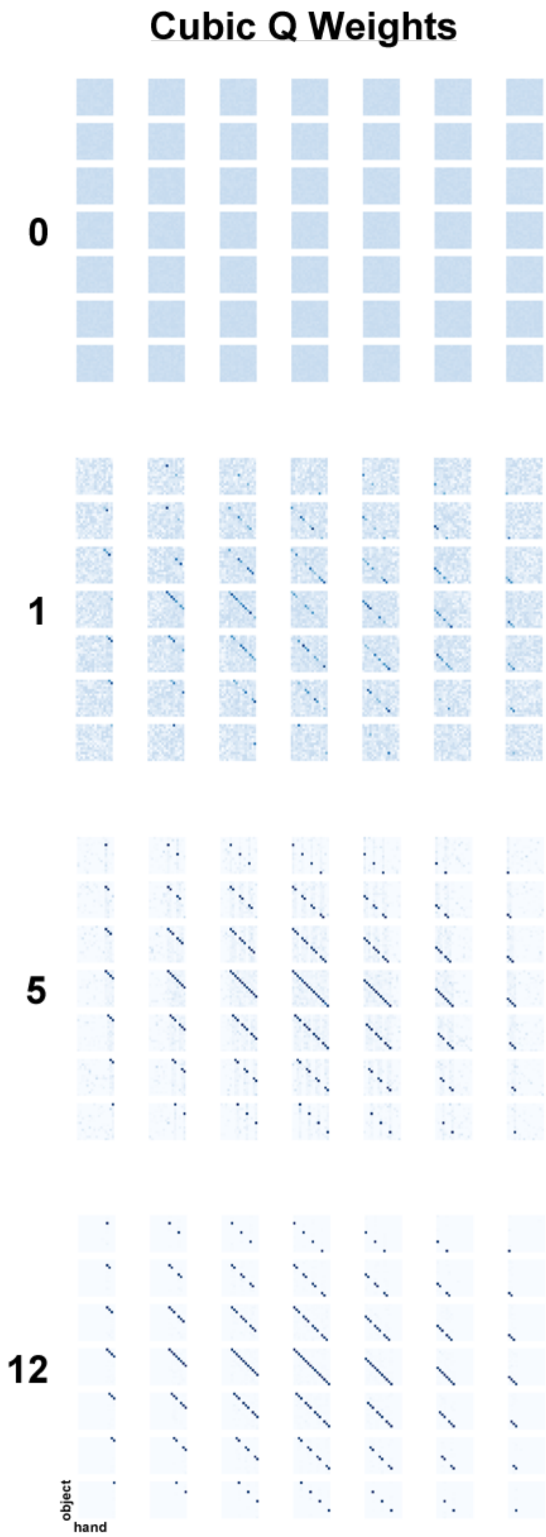


Figure 4.9: Development of the cubic Q weights during RL at different time points (0, 1, 5 and 12×10^4 iterations). Strong weights are displayed dark.

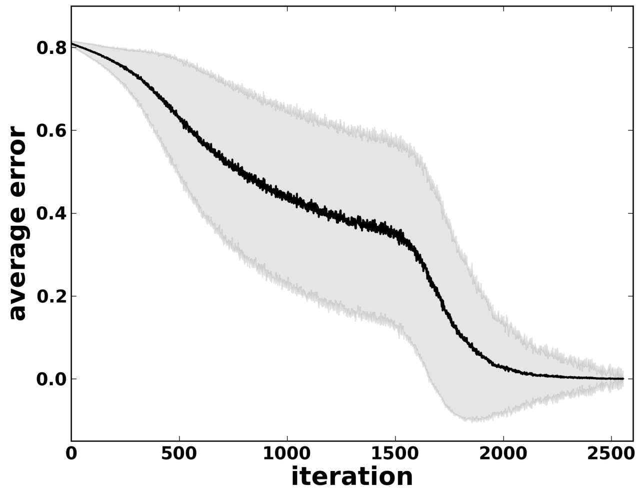


Figure 4.10: Average (1000 runs) error for a 4×4 world. The network is trained using unsupervised learning followed by RL (Ch. 4.3.3). Only the MSE during RL is shown. The error decreases in two phases. This is probably due to the insufficient acquisition of the state space during unsupervised learning. Only the errors of movements that actually reach the target are considered. Invalid actions, as well as unsuccessful ones, are left out. The gray shaded area delineates ±1 STD.

architectures, we initially examine the average MSE during RL. For this purpose the model has been trained on a 4×4 world 1,000 times. The resulting average error trace of successful movements, i.e. movements that actually reached the target, is shown in Fig. 4.10. In contrast to the other learning algorithms, the average error trace shows a two-phasic time course. This is most likely due to the insufficient acquisition of the state space during unsupervised learning. On average it took 100,000 steps of unsupervised learning followed by 15,562 (±973) steps of RL until the system converged. Combined, this two-step procedure then requires about $200\,s$ per training run on a contemporary desktop computer.

As a consequence of the incomplete acquisition of the state space during unsupervised learning, it is not surprising that an agent using this learning scheme is not able to learn perfect reaching movements. The self-organization of the Sigma-Pi weights during unsupervised learning along with the corresponding Gaussians is shown in Fig. 4.11. In the course of learning spatial coherence is exploited for self-organization of the sensory input manifold (Ch. 4.3.3). The neighborhood interactions of the neurons are gradually reduced using a weighted Gaussian winner scheme. For illustration purposes the range influence of the Gaussian-profile activation function is shown schematically for the different time points of RF development. Due to the gradual reduction of the neighborhood interactions the receptive fields developed slower as it was the case for the two previous learning paradigms. In addition, the strength of the weights within a single RF, i.e. all hand-object positions comprising a unique invariance, is not as uniform as it has been for the other architectures.

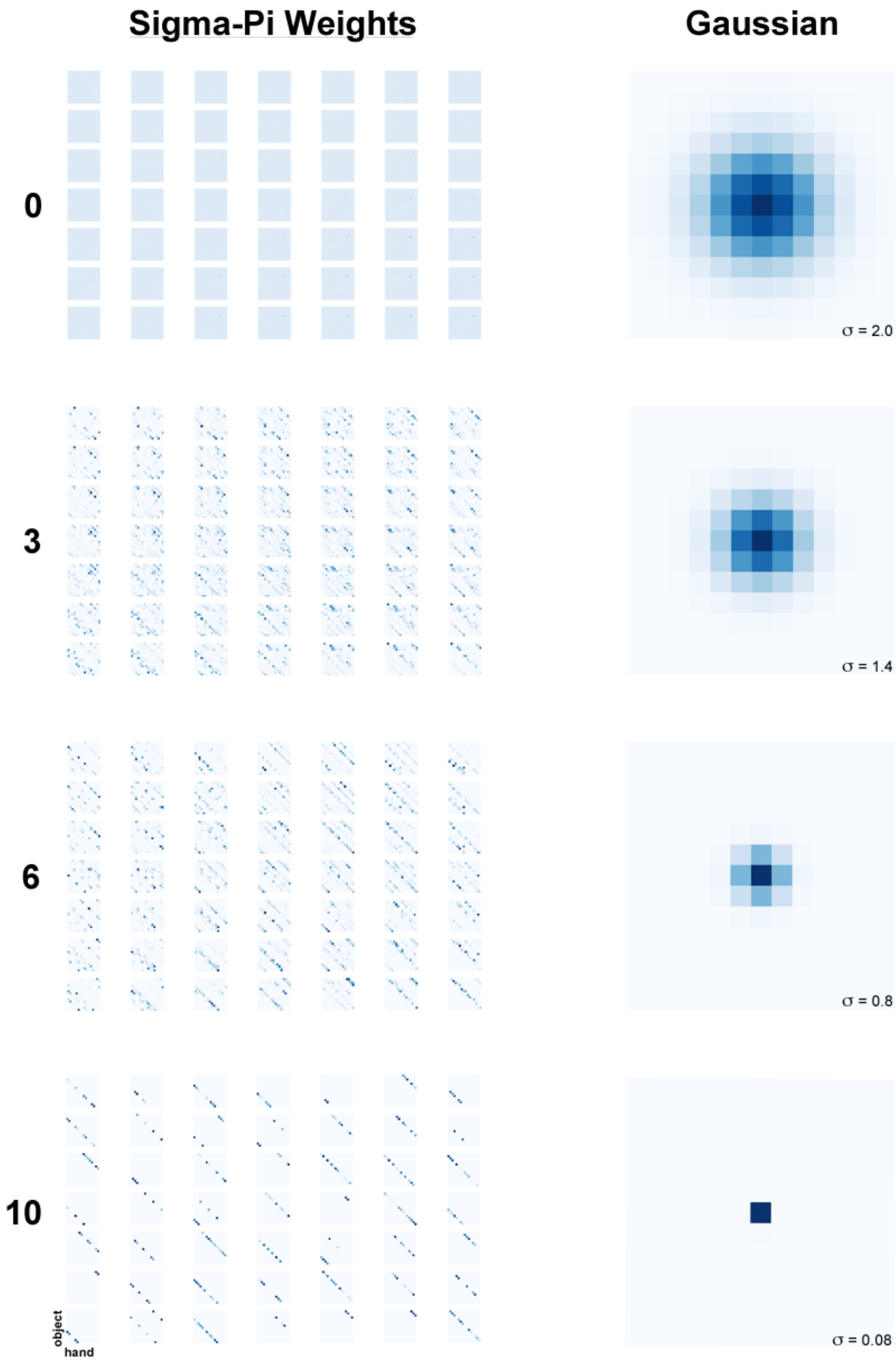


Figure 4.11: Unsupervised RF development and corresponding Gaussians at different times (0, 3, 6 and 10×10^4 iterations). Large values displayed dark.

To further carve out the problems associated with the unsupervised learning algorithm three different invariant hand-object relations of a 4×4 world are displayed after training (Fig. 4.12). As a first example, two different invariances are shown (Fig. 4.12 top and middle) that result in the activation of the same state unit. Consequently, the resulting movement can only be correct for one of them. Another example invariance (Fig. 4.12 bottom) leads to the activation of two different state units. During RL this error was then compensated by linking both states to a single, correct movement. Still, assigning two states to a single invaricance has severe implications because, in turn, another state has to represent two different hand-object relations (as for instance shown in the first example). In general, these two errors reflect the most common problems of this learning algorithm and clearly indicate that those errors are already introduced in the first, unsupervised stage of this two-step learning paradigm.

4.4.4 Comparison of the Three Models

Tab. 4.1 shows how various parameters of the three different learning paradigms scale depending on the size of the world. The number of movements, states (invariances) and weights grows exponentially. Unfortunately, so does the training time for all of the examined architectures.

Our novel unified architecture (Ch. 4.3.1) and cubic RL (Ch. 4.3.2) allow the given reaching task to be learnt flawlessly. In contrast, a two-step learning procedure, unsupervised learning followed by RL (Ch. 4.3.3), resulted in an average behavioral error of 5–6 %. It has been shown above (*cf.* Ch. 4.4.3) that this error is already introduced during the unsupervised self-organization of the state space, i.e. learning of the invariances. Then, the error propagates into the Q-table learned with the RL algorithm. To make a direct comparison possible, we rearranged the Q-tables learned with our novel unified architecture and the canonical two-step learning procedure. It can be seen that every invariance is exactly linked to a single adequate movement for the unified neural architecture (Fig. 4.13 left). However, looking at the Q-table obtained with the two-step procedure (Fig. 4.13 right) reveals a violation of this one-to-one relationship. Several states are connected to a single movement. Further, the weights are not as strong as in the previous case.

Considering our unified architecture, already a world of size 5×5 with its roughly 57,000 weights becomes computationally intractable on contemporary desktop computers. To compare our novel model to a one-layer RL architecture with a cubic Q-table we removed the Sigma-Pi neurons and changed the algorithm accordingly. This architecture scales better, but is still not computationally efficient and therefore cannot handle real-world problems. No results for a world bigger than 7×7 (405,769 weights) have been obtained in reasonable time. Furthermore, we devised a two-step procedure, unsupervised learning followed by RL, to tackle this demanding reaching task. This is a common approach found in the literature (e.g. Legenstein *et al.*, 2010). However, even with this separated learning procedure it was not possible to handle the given task in reasonable time. To illustrate how

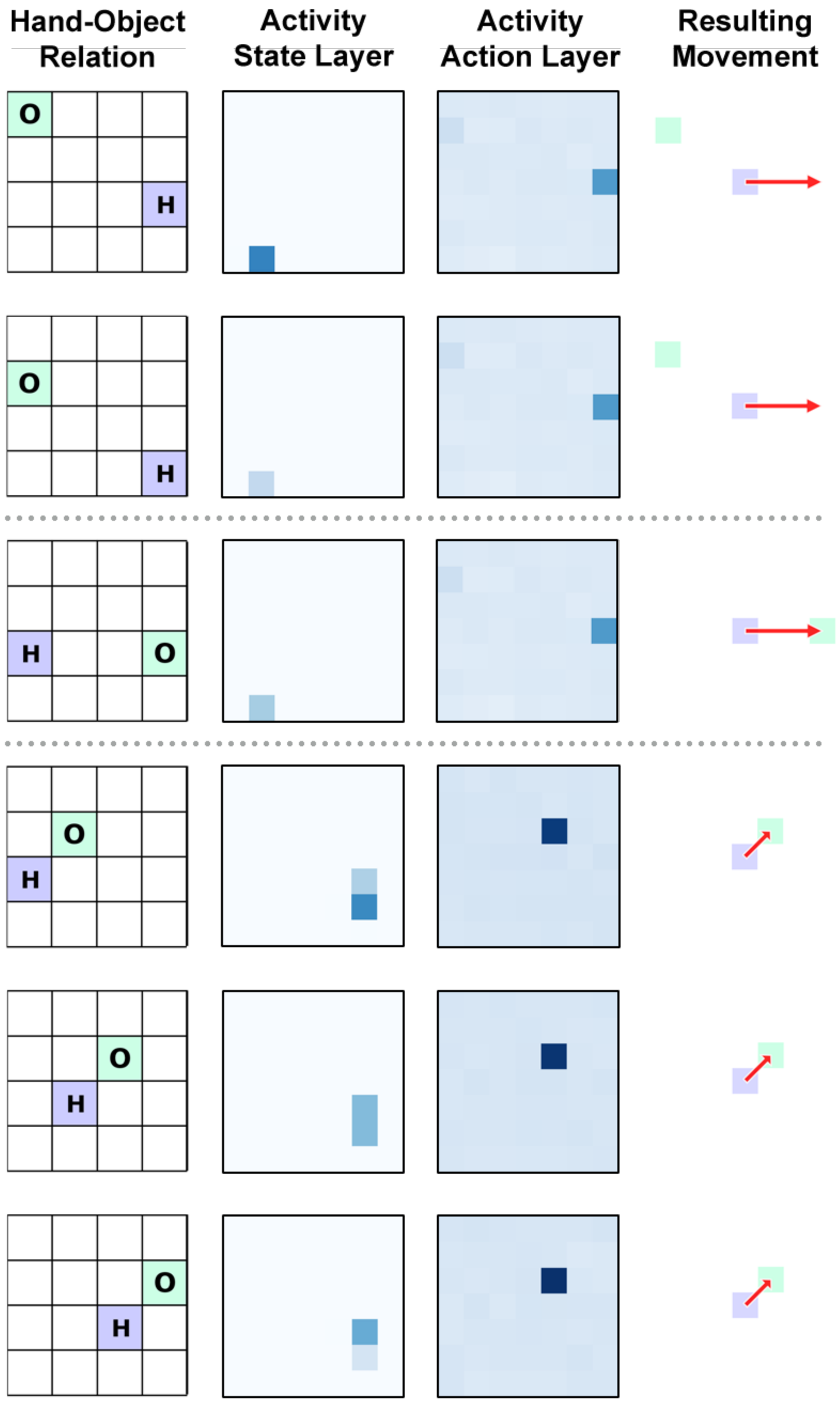


Figure 4.12: Invariant hand-object relations and movements after unsupervised learning followed by RL. Examples of three different invariances are shown. Two different invariances (top and middle) result in the activation of the same state unit and as a consequence the resulting movement is only correct for one of them. Another invariance (bottom) leads to the activation of two different state units. However, during RL both are linked to the correct movement. Strong activations are displayed dark.

Table 4.1: Scaling of world, parameters and training steps.

Size of the world	2×2	3×3	4×4	5×5
Number of actions and states	9	25	49	81
Novel unified architecture				
Weights	225	2,650	14,945	57,186
Iterations	∼3,000	∼300,000	$\sim 13 \times 10^6$	–
Avg. error (%)	0	0	0	–
Cubic RL				
Weights	144	2,025	12,544	50,625
Iterations	∼2,500	∼20,000	∼120,000	∼520,000
Avg. error (%)	0	0	0	0
Unsupervised followed by RL				
Weights unsupervised	144	2,025	12,544	50,625
Weights RL	81	625	2,401	6,561
Weights total	225	2,650	14,945	57,186
Iterations unsupervised	∼2,000	∼20,000	∼100,000	∼200,000
Iterations RL	∼2,000	∼5,500	∼15,000	∼42,000
Iterations total	∼4,000	∼25,500	∼115,000	∼242,000
Avg. error (%)	25	5	6	6

the number of training iterations depend on the size of the world, this relationship is drawn in a semi-log plot (Fig. 4.14). The increase of iterations is most rapid for the unified architecture. The scaling for the other two models is very similar, unfortunately, still exponential.

All in all, the results suggest that yet another mechanism might be responsible for the biological learning of reaching movements and for the identification of invariant situations. It may also be the case that the proposed mechanism is correct and that the learning can be accomplished more efficiently in the brain. Nevertheless, if the proposed unified model really would be the biological mechanism for the learning of reaching movements, it would take about 80 years to learn a 4×4 world when training with 450 examples per day.

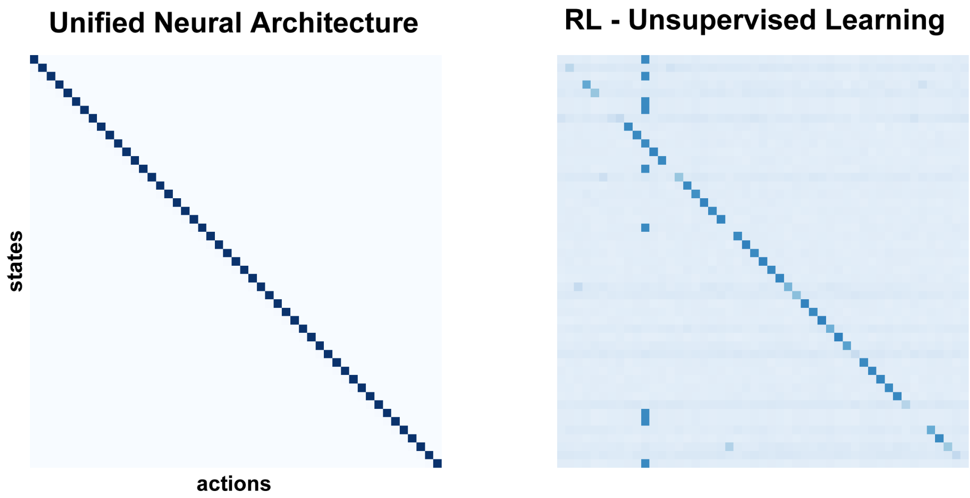


Figure 4.13: Comparison of rearranged Q tables. Every invariance (state) is linked to a single adequate action after learning with the unified architecture (left). However, this is not the case for RL following unsupervised learning (right). Some state-action pairs are not correctly learned. Strong weights are displayed dark.

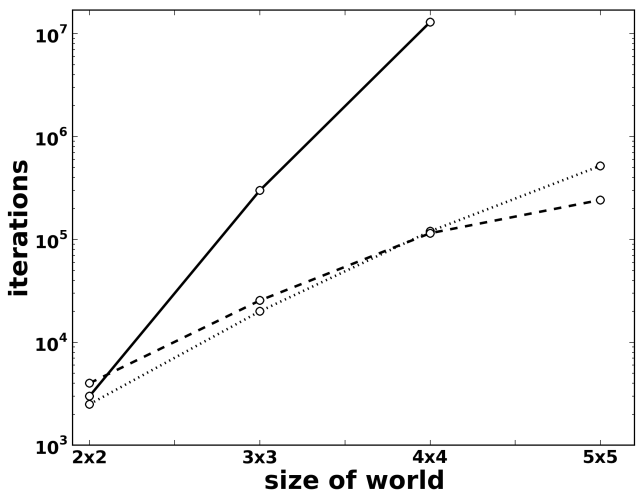


Figure 4.14: Semi-log plot depicting training iterations of the different architectures depending on the size of the world. Comparing the three models reveals that the increase of iterations is much more rapid for the unified architecture. The solid line shows the iterations of the unified architecture, the dashed line unsupervised learning followed by RL and the dotted line cubic RL.

4.5 Summary

We presented a novel unified neural architecture for the learning of invariant hand-object relations and their corresponding movements. The network handles implicit frame of reference transformations, thereby ignoring from which modality the sensory information comes from and in what coordinate system this information is encoded in. We confirmed its functioning with several simulation that all were mastered flawlessly by the proposed learning algorithm. Furthermore, the devised architecture has been compared to two related neural architectures. One of them being a reinforcement learning scheme based on a cubic Q-table, the other one a two-step paradigm, comprising unsupervised learning followed by RL. Unfortunately, all models do not scale very well. The best scaling goes along with the two-step method, i.e. RL following unsupervised learning. However, this procedure leads at the same time to a solution with an average behavioral error of 5–6 %. The other two architectures are able to learn the given task perfectly. Parts of the results of this chapter have been published (Kleesiek *et al.*, 2010).

4.5.1 Connection to the Other Experiments of this Thesis

For the development of the architecture we relied on general *sensorimotor design principles* adopted from the embodied cognition framework (Ch. 2) and the knowledge obtained in Ch. 3. Performing gradient descent on an energy function allows a learning rule to be defined for a network architecture that combines an action (RL) layer with a sensory layer (Ch. 3.3.2). In the network presented in this chapter the sensory layer consists of Sigma-Pi neurons, which can detect a co-activations of input units. In this way, the network is able to detect invariances and to perform a coordinate transformation between different sensory modalities. Again, the agent learns in a ‘top-down’ *action-driven* way which sensory features are relevant for the given task. The RL prediction error modulates the state-action mapping and the receptive field development of the sensory layer, leading to *sensorimotor laws* that enable the agent to perform goal-directed behavior.

Due to limited scaling properties of the two-layer architecture we extended it with an adaptive learning regime. A modified version of this adaptive learning mechanism is also used for the improvement of the learning algorithm for the recurrent neural network with parametric bias that will be presented in the next chapter. The reason for moving to a recurrent architecture was guided by several deliberations, one of them being the scaling issues already mentioned. Further, RNNs are known for their superior generalization potential. The most important reason however, is given by the properties of the parametric bias units, which allow to self-organize time series into clusters. In our experiments, these clusters reflect *sensorimotor laws* that will be used for object categorization (Ch. 5).

It should be axiomatic that perception is not passive, but active. Perceptual activity is exploratory, probing, searching; percepts do not simply fall onto sensors as rain falls onto ground. We do not just see, we look.
RUZENA BAJCSY

Chapter

5 Active Perception Using an RNN with Parametric Bias

5.1 Introduction

Motor actions determine the sensory information that agents receive from their environment. Combining sensory and motor processes dynamically facilitates many tasks (*cf.* Ch. 2), one of those being object classification.

The intention of this experiment is to provide a neuroscientifically and philosophically inspired model for *what do objects feel like?* For this purpose, we stress the active nature of perception within and across modalities. According to the theory of sensorimotor contingencies (*cf.* Ch. 2.4.4), actions are fundamental for perception and help to distinguish the qualities of sensory experiences in different sensory channels (e.g. 'seeing' or 'touching'). O'Regan and Noë (2001) actually suggest that "seeing is a way of acting". Exactly this is mimicked in this computational study.

It has been shown that if the fruit fly drosophila cannot recognize a pattern it starts to move (Dill *et al.*, 1993). It is also known that flies use motion to visually determine the depth of perceived obstacles (Franceschini, 1997). Similarly, pigeons bob their heads up and down to recover depth information (Steinman *et al.*, 2000). Not only living beings, but robots too are embodied, and they have the ability to act and to perceive. In the presented experiments the robot actually needs to act to perceive the objects it holds in its hand. The action-driven sensations are guided by the physical properties of its body, the world and the interplay of both.

A humanoid robot moves toy bricks up and down and rotates them back and forth, while holding them in its hand. The induced multi-modal sensory impressions are used to train an improved version of a recurrent neural network with parametric bias (RNNPB), originally developed by Tani and Ito (2003). As a result, the robot is able to self-organize the contextual information to *sensorimotor laws*, which in turn can be used for object classification. Due to the overwhelming generalization capabilities of the recurrent architecture, the robot is even able to correctly classify unknown objects. Furthermore, we show that the proposed model is very robust against noise.

5.2 Theory

Despite its intriguing properties, the recurrent neural network with parametric bias has hardly been used by anybody other than the original authors. Mostly, the architecture is utilized to model the mirror neuron system (Tani *et al.*, 2004; Cuijpers *et al.*, 2009). Here we apply the variant proposed by Cuijpers *et al.* (2009) using an Elman-type structure (Kolen and Kremer, 2001) at its core. Furthermore, we modify the training algorithm to include adaptive learning rates for training of the weights, as well as the PB values. This results in an improved architecture that is more stable and converges faster (for an evaluation please see Ch. 5.4.1).

5.2.1 Storage of Time Series

The recurrent neural network with parametric bias (an overview of the architecture unfolded in time can be seen in Fig. 5.1) can be used for the storage, retrieval and recognition of sequences. For this purpose, the parametric bias (PB) vector is learned simultaneously and *unsupervised* during normal training of the network. The prediction error with respect to the desired output is determined and backpropagated through time using the BPTT algorithm (ibid.). However, the error is not only used to correct all the synaptic weights present in the Elman-type network. Additionally, the error with respect to the PB nodes δ^{PB} is accumulated over time and used for updating the PB values after an entire forward-backward pass of a single time series, denoted as epoch e. In contrast to the synaptic weights that are shared by all training patterns, a unique PB vector is assigned to each individual training sequence. The update equations for the i-th unit of the parametric bias pb for a time series of length T is given as:

$$\rho_i(e+1) = \rho_i(e) + \gamma_i \sum_{t=1}^{T} \delta_{i,t}^{\mathrm{PB}} \ , \tag{5.1}$$

$$pb_i(e) = \mathrm{sigmoid}(\rho_i(e)) \ , \tag{5.2}$$

where γ is the update rate for the PB values, which in contrast to the original version is not constant during training and not identical for every PB unit. Instead, it is scaled proportionally to the absolute mean value of prediction errors being backpropagated to the i-th node over time T:

$$\gamma_i \propto \frac{1}{T} \left\| \sum_{t=1}^{T} \delta_{i,t}^{\mathrm{PB}} \right\| \ . \tag{5.3}$$

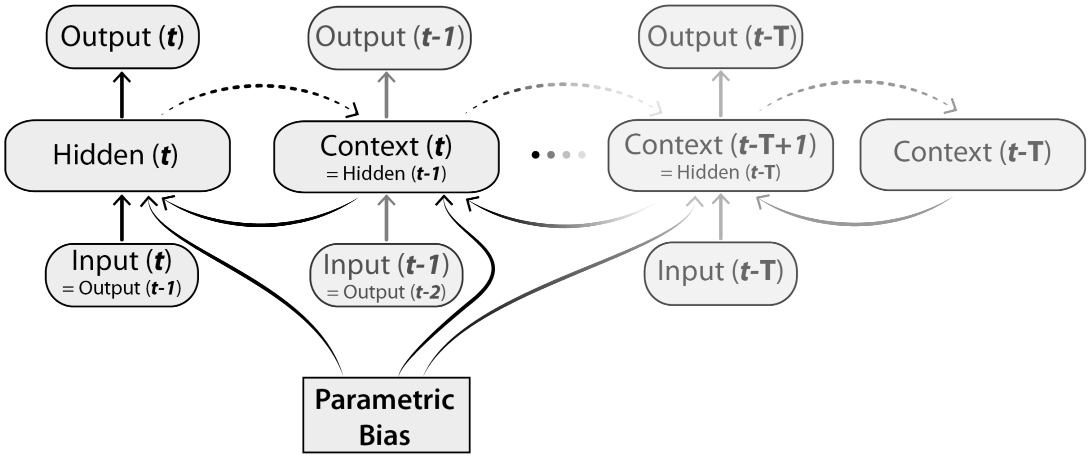


Figure 5.1: Network architecture. The Elman-type Recurrent Neural Network with Parametric Bias (RNNPB) unfolded in time. Dashed arrows indicate a verbatim copy of the activations (weight connections set equal to 1.0). All other adjacent layers are fully connected. t is the current time step, T denotes the length of the time series.

The other adjustable weights of the network are updated via an adaptive mechanism, inspired by the resilient propagation algorithm proposed by Riedmiller and Braun (1993). However, there are decisive differences. First, the learning rate of each neuron is adjusted after every epoch. Second, not the sign of the partial derivative of the corresponding weight is used for changing its value, but instead the partial derivative itself is taken.

To determine if the partial derivative of weight w_{ij} changes its sign we can compute:

$$\epsilon_{ij} = \frac{\partial E_{ij}}{\partial w_{ij}}(t-1) \cdot \frac{\partial E_{ij}}{\partial w_{ij}}(t) \tag{5.4}$$

If $\epsilon_{ij} < 0$, the last update was too big and the local minimum has been missed. Therefore, the learning rate η_{ij} has to be decreased by a factor $\xi^- < 1$. On the other hand, a positive derivative indicates that the learning rate can be increased by a factor $\xi^+ > 1$ to speed up convergence. This update of the learning rate can be formalized as:

$$\eta_{ij}(t) = \begin{cases} \mathbf{max}(\eta_{ij}(t-1) \cdot \xi^-, \eta_{min}) & \text{if } \epsilon_{ij} < 0, \\ \mathbf{min}(\eta_{ij}(t-1) \cdot \xi^+, \eta_{max}) & \text{if } \epsilon_{ij} > 0, \\ \eta_{ij}(t-1) & \text{else.} \end{cases} \tag{5.5}$$

The succeeding weight update Δw_{ij} then obeys the following rule:

$$\Delta w_{ij}(t) = \begin{cases} -\Delta w_{ij}(t-1) & \text{if } \epsilon_{ij} < 0, \\ \eta_{ij}(t) \cdot \frac{\partial E_{ij}}{\partial w_{ij}}(t) & \text{else.} \end{cases} \tag{5.6}$$

In addition to reverting the previous weight change in the case of $\epsilon_{ij} < 0$ the partial derivative is also set to zero ($\frac{\partial E_{ij}}{\partial w_{ij}}(t) = 0$). This prevents changing of the sign of the derivative once again in the succeeding step and thus a potential double punishment.

We use a nonlinear activation function with parameters recommended by LeCun *et al.* (1998) for all neurons in the network, as well as for the PB units (Eq. 5.2):

$$\text{sigmoid}(x) = 1.7159 \cdot \tanh\left(\frac{2}{3} \cdot x\right) . \tag{5.7}$$

5.2.2 Retrieval

The PB vector is usually low dimensional and resembles bifurcation parameters of a nonlinear dynamical system, i.e. it characterizes fixed-point dynamics of the RNN. During training the PB values are self-organized, thereby encoding each time series and arranging it in PB space according to the properties of the training pattern. This means that the values of similar sequences are clustered together, whereas more distinguishable ones are located further apart. Once learned, the PB values can be used for the generation of the time series previously stored. For this purpose, the network is operated in closed-loop mode. The PB values are 'clamped' to a previously learned value and the forward pass of the network is executed from an initial input $I(0)$. In the next time steps, the output at time t serves as an input at time $t+1$. This leads to a reconstruction of the training sequence with a very high accuracy (limited by the convergence threshold used during learning, e.g. as shown in Fig. 5.12 on the left).

5.2.3 Recognition

A previously stored (time) sequence can also be recognized via its corresponding PB value. Therefore, the observed sequence is fed into the network without updating any connection weights. Only the PB values are accumulated according to Eq. 5.1 and 5.2 using a constant learning rate γ this time. Once a stable PB vector is reached (as shown in Fig. 5.13), it can be compared to the one obtained during training.

5.2.4 Generalized Recognition and Generation

The network has substantial generalization potential. Not only previously stored sequences can be reconstructed and recognized. But, (time) sequences apart from the stored patterns can be generated. Since only the PB values but not the synaptic weights are updated in recognition mode, a stable PB value can also be assigned to an unknown sequence.

For instance, training the network with two sine waves of different frequencies allows cyclic functions with intermediate frequencies to be generated simply by operating the network in generation mode and varying the PB values within the interval of the PB values obtained during training. Furthermore, the PB

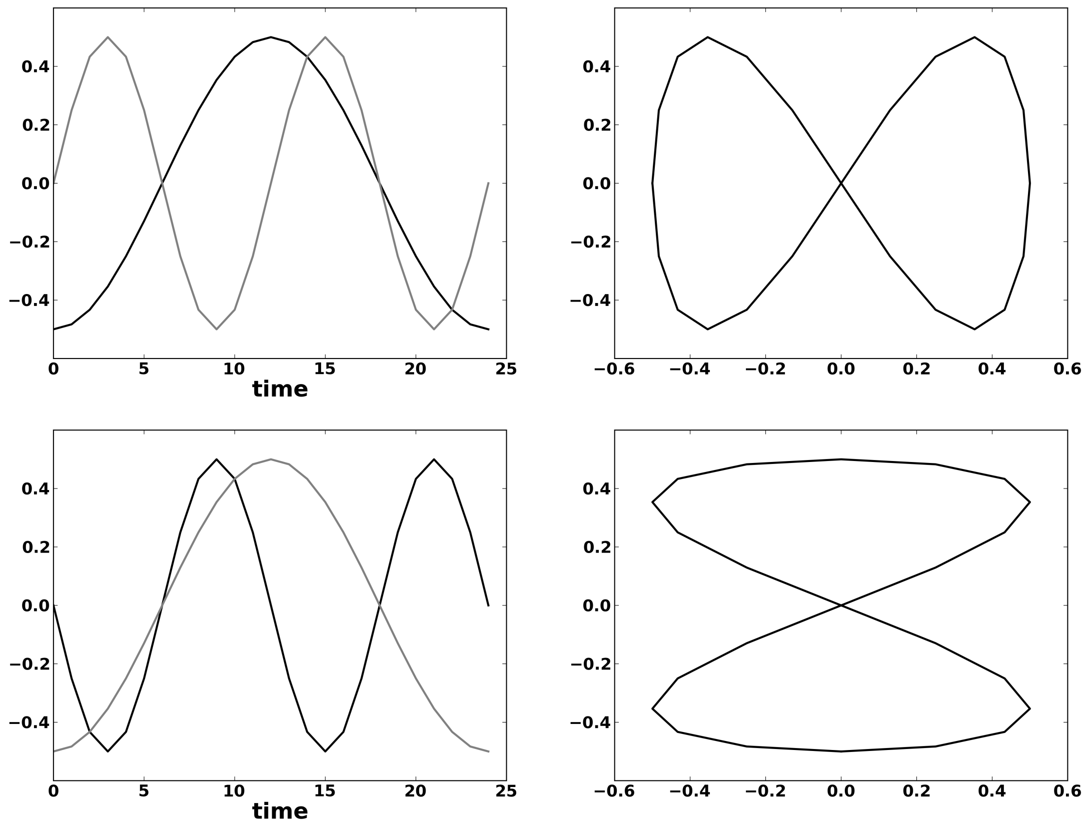


Figure 5.2: Artificial data to demonstrate generalization potential. The 2-D sequences resulting from Eq. 5.8 for $\theta = 90$ (top) and $\theta = 180$ (bottom), respectively. On the left, the two dimensions are plotted separately, whereas on the right x_1 is plotted against x_2, leading to a figure-eight shape.

values obtained during recognition of a previously unseen sine function with an intermediate frequency (w.r.t. the training sequences) will lie within the range of the PB values acquired during learning. Hence, the network is able to capture a reciprocal relationship between a time series and its associated PB value.

These generalized recognition and generation capabilities of the adaptive RNNPB are demonstrated in a more complex example. For this purpose, consider the 2-D sinusoidal sequences described by the following equation:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.5 \begin{pmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{pmatrix} \begin{pmatrix} sin\ \frac{\pi \cdot t}{6} \\ cos\ \frac{\pi \cdot t}{12} \end{pmatrix} \tag{5.8}$$

Plotting x_1 *vs.* x_2 results in a figure-eight shape that is rotated according to the angular value specified by θ. Two 2-D time series of length $t = 25$ were generated using $\theta = 90$ and $\theta = 180$, respectively. These sequences and the resulting figure-eight shapes are shown in Fig. 5.2. The network was trained with the parameters specified in Tab. 5.1. Note, in contrast to the robot experiments presented below, the network only has a single PB unit.

Table 5.1: Network and learning parameters for 'rotating eight' experiment.

Input neurons	2
Hidden neurons	32
Context neurons	32
Output neurons	2
PB units	1
Convergence threshold	3×10^{-4}
Training duration	$\sim$14 h

After training, the network is able to recognize those sequences based on their trained PB values ($\mathrm{PB}_{\theta:90} = -0.01107$ and $\mathrm{PB}_{\theta:180} = -0.65604$), which differ only by a small amount ($\epsilon_{\theta:90} = 0.0005$ and $\epsilon_{\theta:180} = 0.002$) from the ones obtained during storage. The PB values of the two trained sequences are plotted in Fig. 5.3 using white circular markers. Next to the trained sequences, the network is also fed with novel, previously untrained, sequences. These are generated using Eq. 5.8 with varying θ values. The network also generates stable PB values for those unknown sequences (gray and black dots in Fig. 5.3). It can be seen that the PB values are ordered according to the angular value of the underlying time series. For instance, this reciprocal relationship could be used to infer the angular value of a sequence generated with an unknown θ value.

Moreover, using the intermediate PB values (gray markers) located in the interval between the two PB values obtained during training (white markers) allows 2-D sinusoidal functions to be generated. For this purpose, the ordered PB values are 'clamped' to the network one after each other and the forward pass is executed in closed-loop mode (*cf.* Ch. 5.2.2). This allows a 2-D sequence to be generated for each PB value. Plotting x_1 *vs.* x_2 of these newly generated time series in a row results in a figure-eight shape that rotates from 90 to 180 degrees (Fig. 5.4). It shows that the architecture is indeed capable of learning the underlying principle that generated the two training sequences. This generalization potential will be exploited in the robot experiments for the categorization of unknown objects.

5.2.5 Evaluation of the Adaptive Learning Rate

To evaluate the adaptive learning rate proposed in Ch. 5.2.1, artificial 1-D test data of length $T = 11$ in the interval $[-\pi; \pi[$ is generated using the following equations:

$$x = \sin(t)\,, \tag{5.9}$$

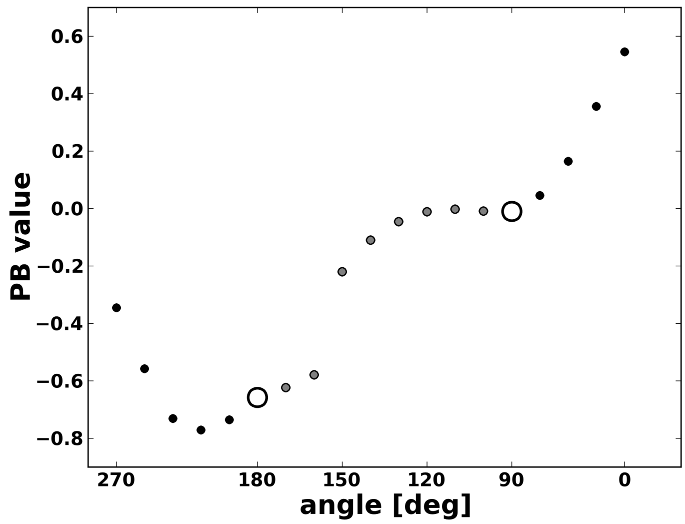


Figure 5.3: Generalized recognition of trained and untrained sequences. The PB values of the two trained 2-D time series using Eq. 5.8 with $\theta = 90$ and $\theta = 180$, respectively, are marked using white circles. In contrast, the PB values obtained by feeding the network with untrained sequences generated with varying θ values are drawn as gray and black dots. These values are arranged in a structured way, emphasizing the self-organization of the PB space. In fact, the PB values can be used to generate time series with intermediate θ values. This reciprocal relationship is shown in Fig. 5.4 for the values marked in gray.

$$x = \frac{\sin(3t) \cdot \sin(t)}{2t^2} - 0.5\,. \tag{5.10}$$

Eq. 5.9 is referred to as `sin` and Eq. 5.10 as `sinc`. Both time series are shown in Fig. 5.5. Except for the following differences, the RNNPB network parameters were identical to the parameters of the robot experiments (see below). The architecture contained only one input and one output node, as well as only 1 PB unit. The convergence criterion was set to 10^{-4}.

5.2.6 Network Parameters for Robot Experiments

To quantify the number of principle components (PCs) actually needed for (almost) lossless reconstruction of the PB space, we determined how many are necessary to explain 99 % of the variance. Increasing the number of PB values, given a bi-modal time series of length $T = 14$, resulted in a constant number of two PCs. Hence, we use a 2-D PB vector for our experiments.

Based on systematic empirical trials, the following parameters have been determined for our experiments. The network contained two input and two output nodes, 24 hidden and 24 context neurons as well as 2 PB units. The convergence criterion for back propagation through time (BPTT) was set to 10^{-6} in the first, and 10^{-5} in the second experiment. For recognition of a sequence, the update rate

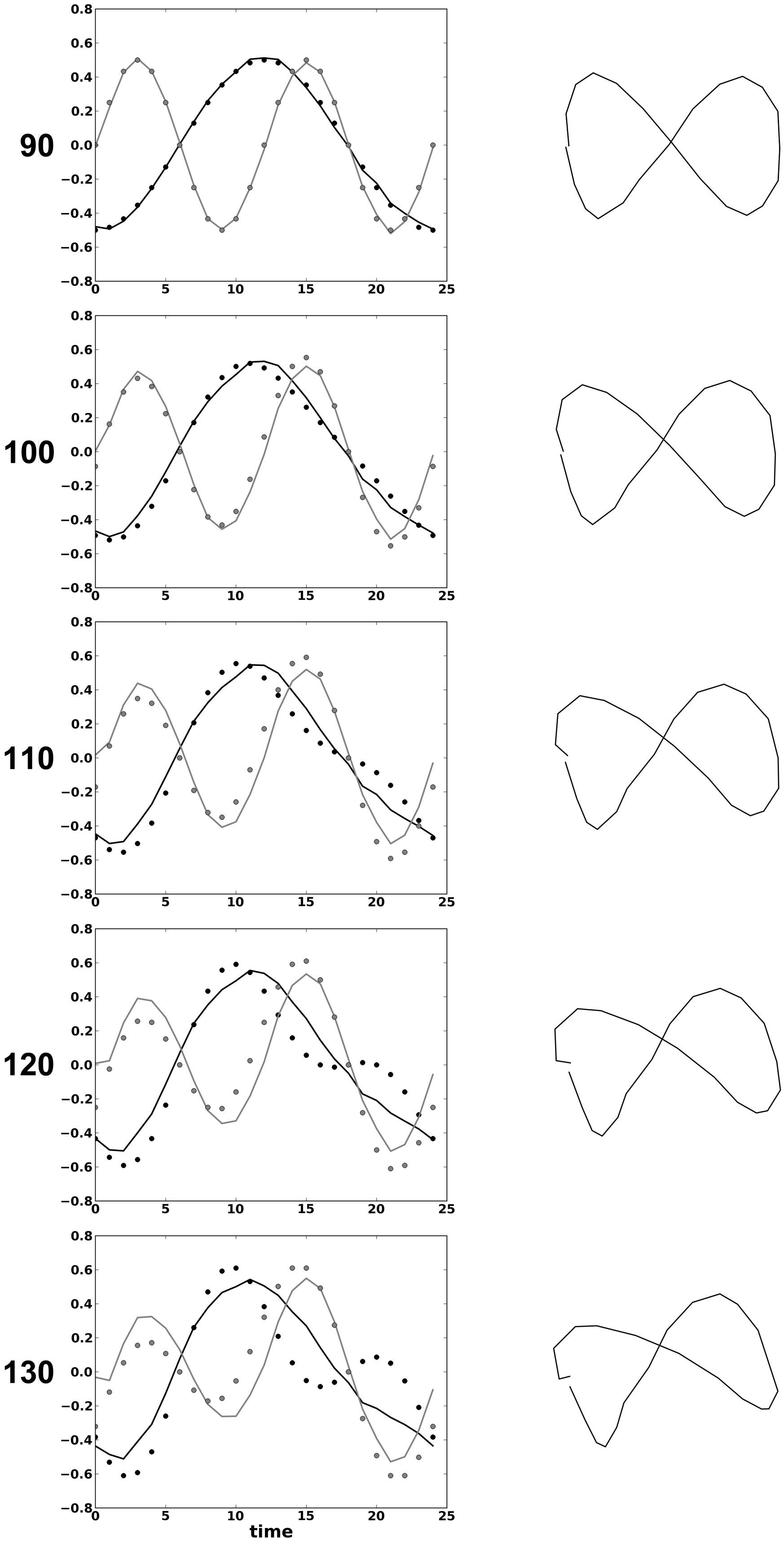


Figure 5.4: The 2-D sequences generated by the network using the PB values shown in Fig. 5.3 as gray dots. The two dimensions are either plotted separately (left) or against each other (right).

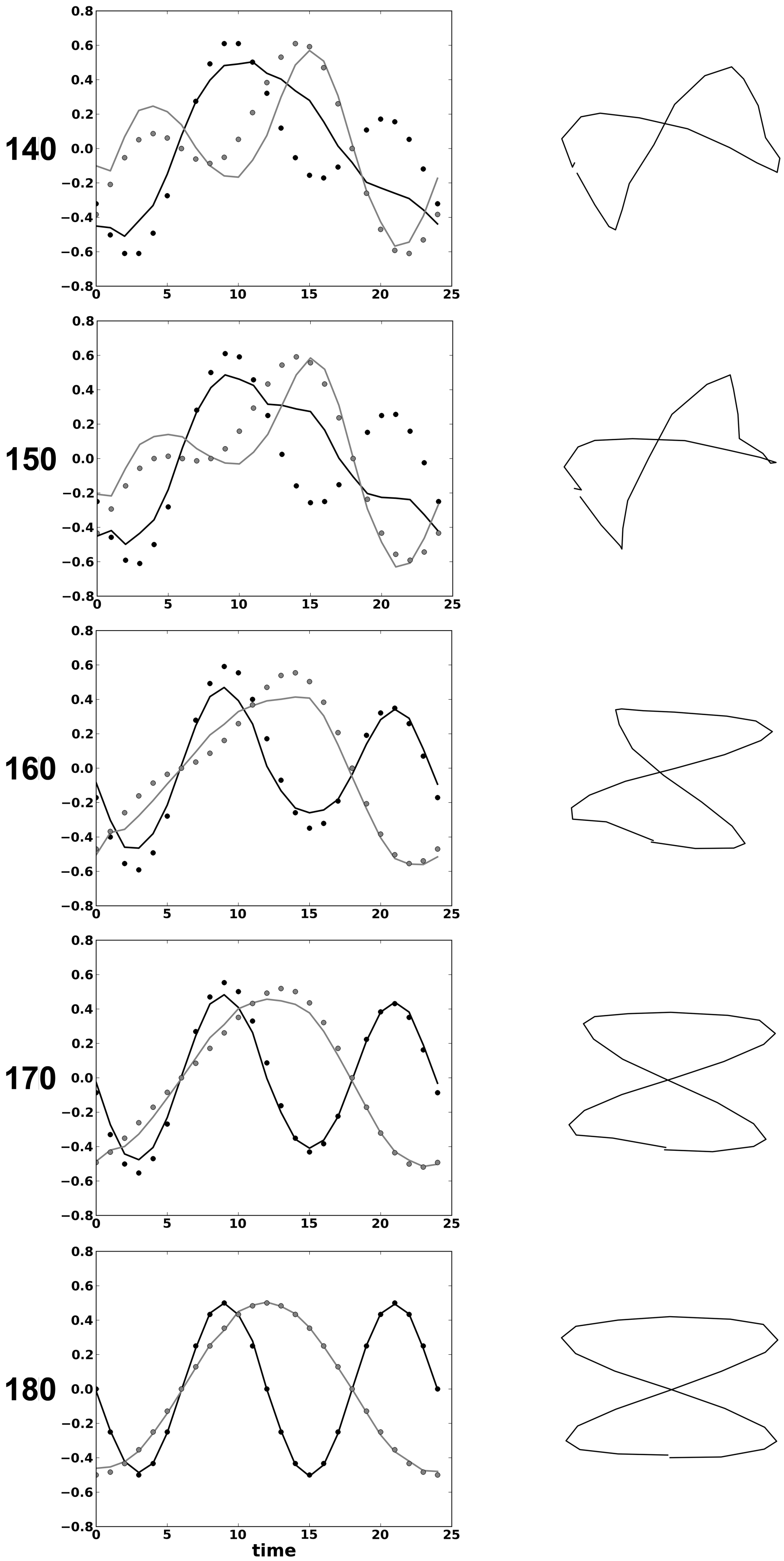


Figure 5.4: Dots represent the function values that would be obtained when using Eq. 5.8 with the angular values for θ shown in bold to the left of each row.

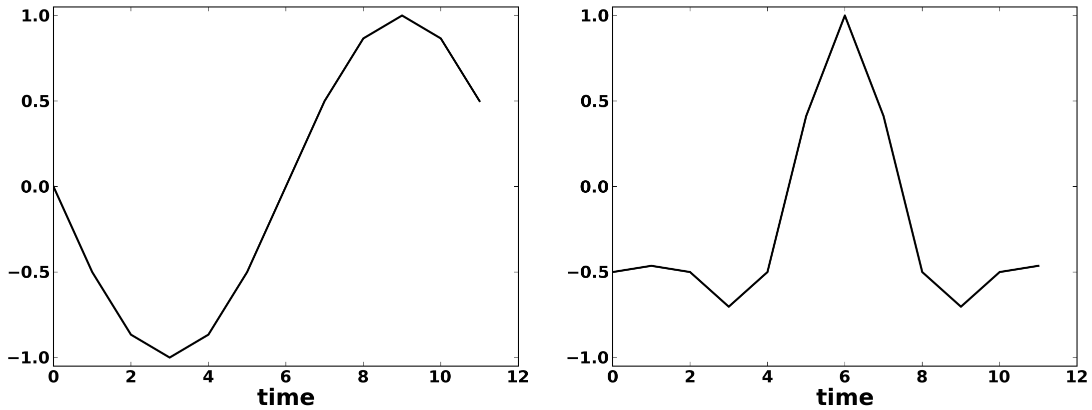


Figure 5.5: Artificial test data for evaluation of the adaptive learning rate. Eq. 5.9 is used to generate the sequence on the left (`sin`) and Eq. 5.10 for the sequence on the right (`sinc`).

γ of the PB values was set to 0.1. The values for all other individual adaptive learning rates (Eq. 5.5) during training of the synaptic weights were allowed to be in the range of $\eta_{min} = 10^{-12}$ and $\eta_{max} = 50$; depending on the gradient they were either increased by $\xi^{+} = 1.01$ or decreased by a factor $\xi^{-} = 0.9$.

5.3 Scenario

The humanoid robot Nao (Aldebaran Robotics) is programmed to conduct the experiments (Fig. 5.6 a). The task for the robot is to identify which object (toy brick) it holds in its hand. In total there are eight object categories that have to be distinguished by the robot: the toy bricks have four different shapes (circular-, star-, rectangular- and triangular-shaped), which each exist in two different weight versions (light and heavy). Hence, for achieving a successful classification multi-modal sensory impressions are required. Additionally, *active* perception is necessary to induce sensory changes essential for discrimination of –depending on the perspective– similar looking shapes (e.g. star- and circular-shaped objects). For this purpose, the robot performs a predefined motor sequence and simultaneously acquires visual and proprioceptive sensor values.

5.3.1 Data Acquisition

The recorded time series comprises 14 sensor values for each modality. In each single trial the robot turns its wrist with the object between its fingers by 45.8 ° back and forth twice, followed by lifting the object up and down three times (thereby altering the pitch of the shoulder joint by 11.5 °) and, finally, turning it again twice.

After an action has been completed, the raw image of the lower camera of the Nao robot is captured, whereas the electric current of the shoulder pitch servo

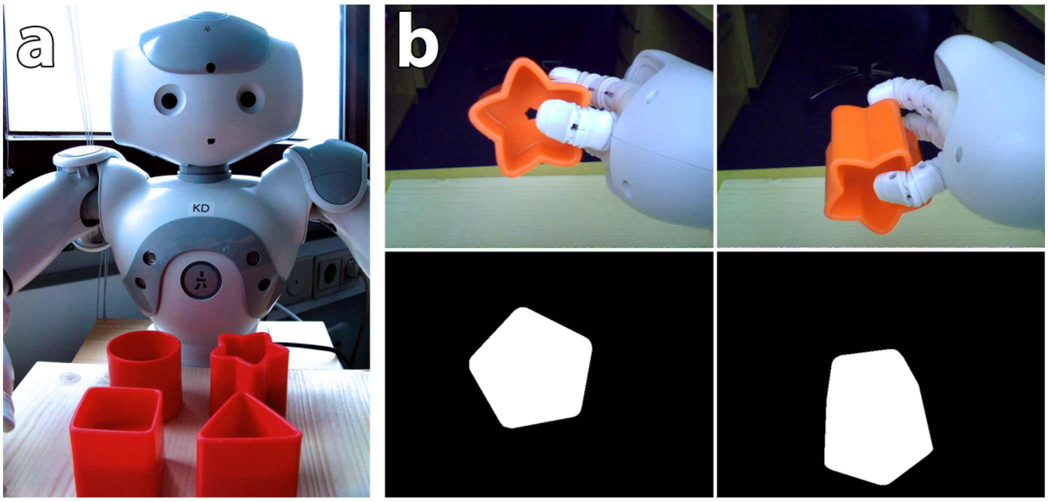


Figure 5.6: Scenario. a) Toy bricks in front of the humanoid robot Nao. The toy bricks exist in four different shapes, have an identical color and are either light-weight ($15\,g$) or heavy ($50\,g$). This results in a total of eight categories that have to be distinguished by the robot. b) Rotation movement with the star-shaped object captured by the robot camera. In the upper row the raw camera image is shown, whereas the bottom row depicts the preprocessed image that is used to compute the visual features.

motor is recorded constantly (sampling frequency $10\,Hz$) over the entire movement interval. For each object category 10 single trial time series are recorded in the described way and processed in real-time. This yields 80 bi-modal time series in total.

5.3.2 Data Processing

For the proprioceptive measurements only the mean values are computed for the time intervals in between movements. The visual processing, on the other hand, involves several steps (Fig. 5.6 b), which are accomplished using OpenCV (Bradski, 2000). First, the raw color image is converted to a binary image using a color threshold. Next, the convex hull is computed and, based on that, the contour belonging to the toy brick is extracted (Suzuki and Be, 1985). For the identified contour the first Hu moment is calculated (Hu, 1962). Finally, the visual measurements are scaled to be within the interval $[-0.5, 0.5]$.

We are aware that more discriminative geometrical features exist, e.g. the orthogonal variant moments proposed by Martín H. *et al.* (2010). However, we deliberately posed the problem this way to make it a challenging task and show the potential of the approach.

5.3.3 Training and Test Data

For testing, the data of single trials are used, i.e. 10 2-D time series per object category (one dimension for each modality). However, for training, a prototype

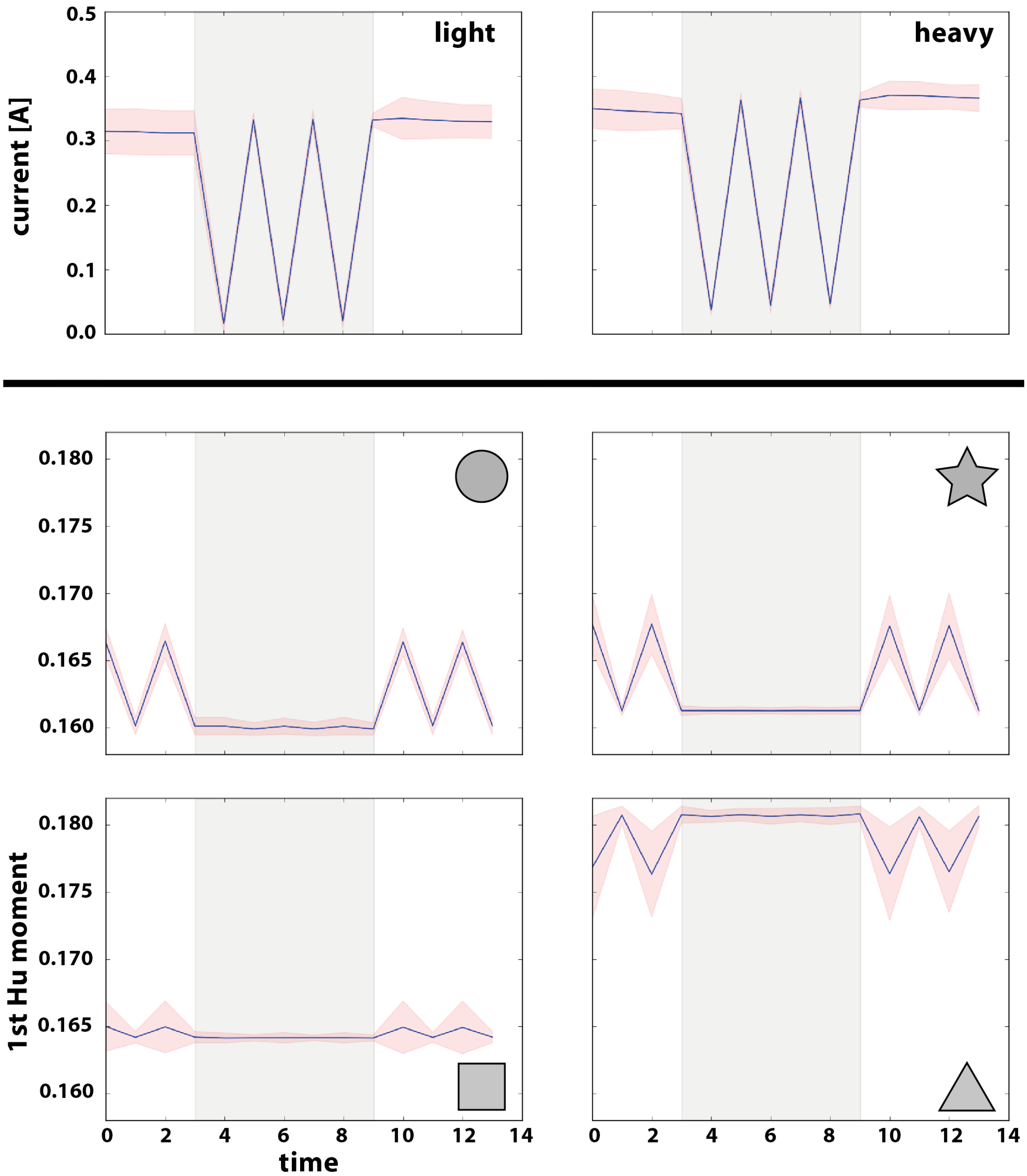


Figure 5.7: Training data. The mean values of the two weight conditions (light and heavy, top) and the four visual conditions (matching symbols, bottom) are shown. These mean time series are used as prototypes for training the RNNPB. Gray shaded areas represent the up and down movement, whereas back and forth movements are unshaded. The red area surrounding the signals delineates two standard deviations from the mean.

is determined for each object category and modality (Fig. 5.7). To obtain this subclass representative, the mean value of pooled single trials, with regard to identical object properties, is computed. This means that, for instance, all circular-shaped objects are combined ($n = 20$) and used to compute the visual prototype for circular-shaped objects. To find the proprioceptive prototype for e.g. all heavy objects, all individual measurements with this property ($n = 40$) are aggregated and used to calculate the mean value at each time step. The subclass prototypes are then combined to form a 2-D multi-modal time series that serves as an input for the recurrent neural network during training.

5.4 Results

5.4.1 Evaluation of the Adaptive Learning Rate

To evaluate the improvements caused by introduction of the adaptive learning rate as described in Ch. 5.2.1, an RNNPB was trained 1000 times with two 1-D sequences (Eqs. 5.9 and 5.10, Fig. 5.5). The results are statistically evaluated using a t-test. To compensate for the sample size bias, the optimal sample size was determined based on the mean value and the standard deviation of the data[1]. This optimal sample size was used to draw 10,000 random subsets of the data, which were subsequently evaluated to obtain an average p-value for the t-test. The results are summarized in Tab. 5.2. The modifications lead, on average, to a 22-fold speedup of the training times (t-test, $p = 0.0$). Also the number of recognition steps has improved significantly (t-test, $p = 0.03$). However, no significant changes of the retrieval accuracy (MSE) can be found.

The PB values that were obtained for the first sequence (`sin`) are plotted against the PB value of the second sequence (`sinc`). This is performed for both algorithms (Fig. 5.8). Values belonging to the adaptive algorithm are depicted in white, whereas the PB values obtained with the classical algorithm are shown in gray. A symmetric relationship can be seen for both conditions, indicating a constant relationship between the PB values of the two sequences despite random initializations of the weights. This stresses the self-organization of the PB space (*cf.* Ch. 5.2.2). Further, the PB values obtained using the adaptive algorithm are on average smaller and display a higher standard deviation (Tab. 5.2). This might be due to the drastically faster training.

Plotting the average MSE against the number of steps needed until convergence further highlights the drastic improvement in speed (Fig. 5.9). The error, shown separately for both sequences, decreases for both algorithms in a similar manner. However, the adaptive version looks ‘compressed’ in comparison to the classical algorithm. In addition, the fluctuations seem to be much less, indicating a more stable behavior of the modified RNNPB.

[1]This was achieved using the DSS Research Web-toolkit.

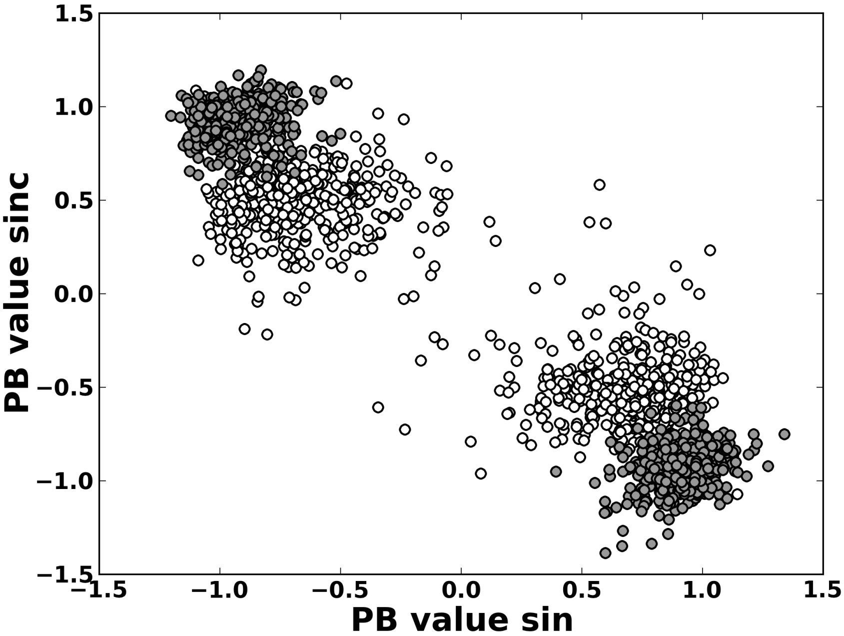


Figure 5.8: Scatter plot of PB values for the two test sequences. The PB values obtained for the first sequence (`sin`) are plotted against the PB value for the second sequence (`sinc`). Gray dots represent values obtained with the classical and white dots values obtained with the adaptive algorithm.

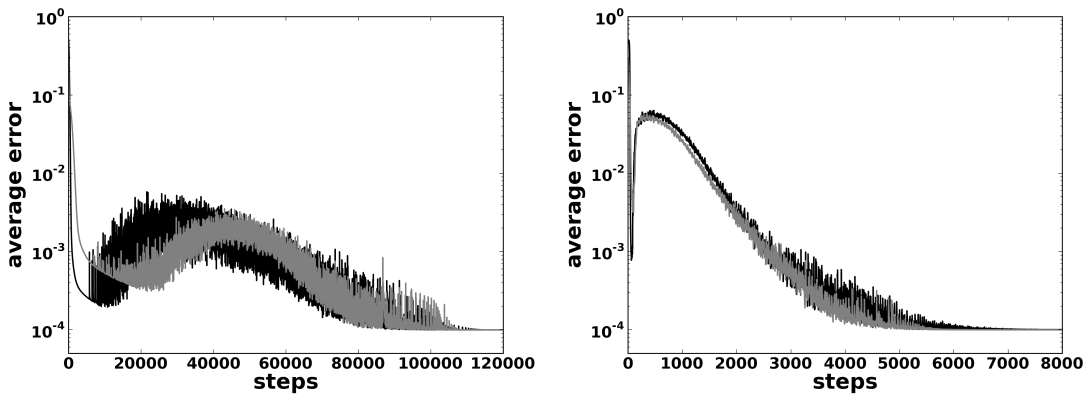


Figure 5.9: Error plot comparing classical (left) to adaptive (right) RNNPB. The average MSE of the `sin` sequence is shown in black, whereas the average MSE of the `sinc` sequence is drawn in gray.

Table 5.2: Statistical evaluation of the adaptive learning rate. Mean values and standard deviations are shown, significant changes (t-test, $p < 0.05$) are marked bold.

	Modified RNNPB	Classical RNNPB	Factor
Total steps	**5,520 (±1,713)**	122,709 (±20,027)	22.2
Total time	**34 s (±10)**	751 s (±124)	22
MSE *sin*	4.3×10^{-4} ($\pm 1.2 \times 10^{-3}$)	5.5×10^{-4} ($\pm 3 \times 10^{-4}$)	–
MSE *sinc*	4.7×10^{-4} ($\pm 8.7 \times 10^{-4}$)	1.9×10^{-4} ($\pm 1.9 \times 10^{-4}$)	–
Recognition steps	**192 (±85)**	284 (±101)	1.48
PB *sin*	±0.59 (±0.22)	±0.92 (±0.11)	–
PB *sinc*	±0.60 (±0.22)	±0.93 (±0.11)	–

Looking at the number of average steps per epoch of the two algorithms reveals that actually the total number of epochs does not seem to be much less in the modified version of the algorithm. Nevertheless, the average steps per epoch are about one order of magnitude smaller (semi-log plot Fig. 5.10). Furthermore, the fluctuations of the trajectories obtained for the adaptive algorithm are again smoother. Overall, the results suggest that the adaptive modifications resulted in a more stable architecture that converges much faster.

5.4.2 Experiment 1 – Classification Using All Object Categories for Training

In the first experiment the improved recurrent neural network with parametric bias was trained with the bi-modal prototype time series of all eight object categories (see Fig. 5.7 and Ch. 5.3.3). During training, the PB values for the respective categories emerged in an unsupervised way. This means, the two-dimensional PB space self-organizes based on the inherent properties of the sensory data that was presented to the network. Hence, objects with similar dynamic sensory properties are clustered together. This can be seen in Fig. 5.11. For instance, the learned PB vectors representing star- and circular-shaped objects, either light-weight (light gray) or heavy (dark gray), are located in close proximity, whereas the PB values coding for the triangular-shaped objects are positioned more distantly. This is due to the deviating visual sensory impression they generate (Fig. 5.7). The experiment has been repeated several times with different random initializations of the network weights. However, the obtained PB values of the different classes always demonstrate a comparable geometric relation with respect to each other.

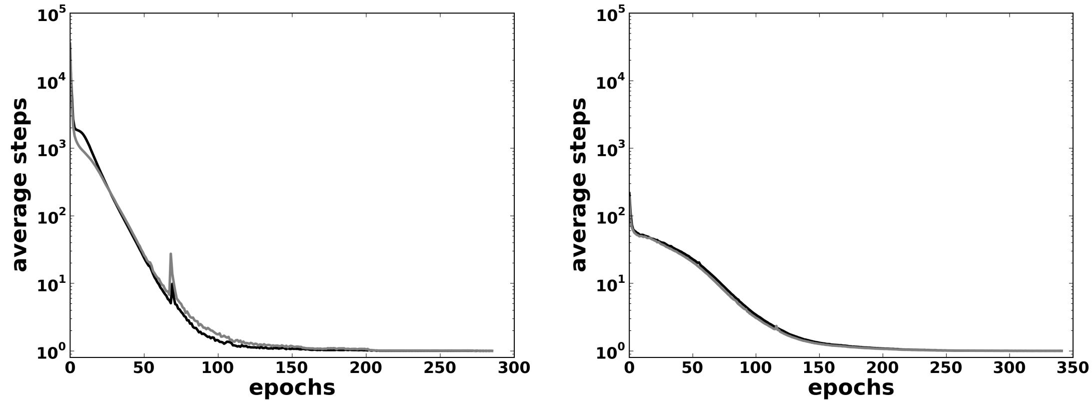


Figure 5.10: Semi-log plot showing average steps per epoch comparing classical (left) to adaptive (right) RNNPB. The average number of steps per epoch of the `sin` sequence are shown in black and for the `sinc` sequence in gray.

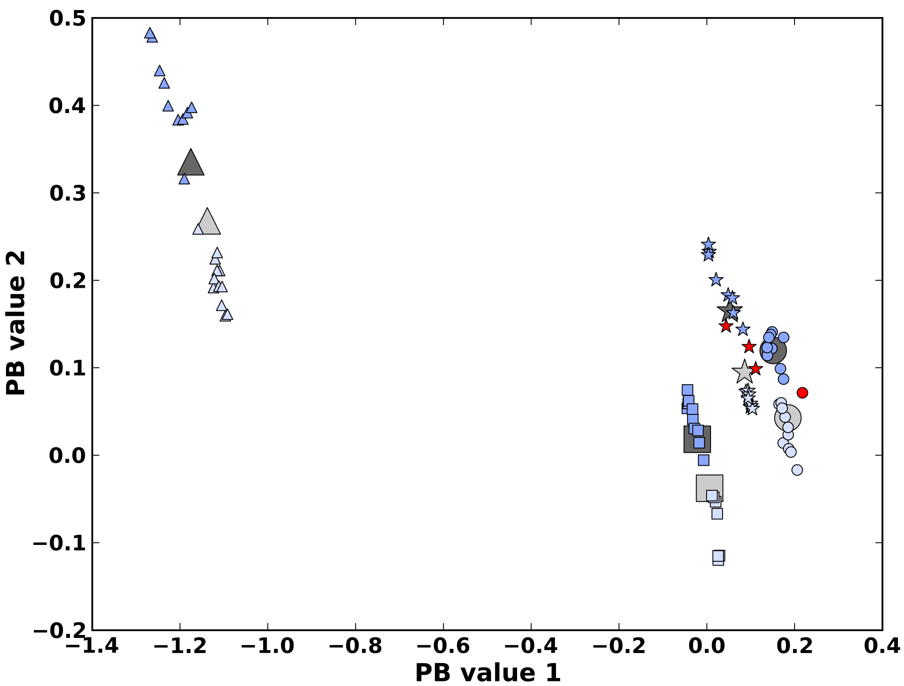


Figure 5.11: Experiment 1 – Classification using all object categories for training. PB values of the class prototypes used for training are depicted in light and dark gray and with a symbol matching the corresponding shape. Smaller symbols depict PB values obtained during testing with bi-modal single trial data. If the objects have been correctly classified, they are shown in light or dark blue, otherwise in red. Light colors are used for light-weight, dark colors for heavy-weight objects.

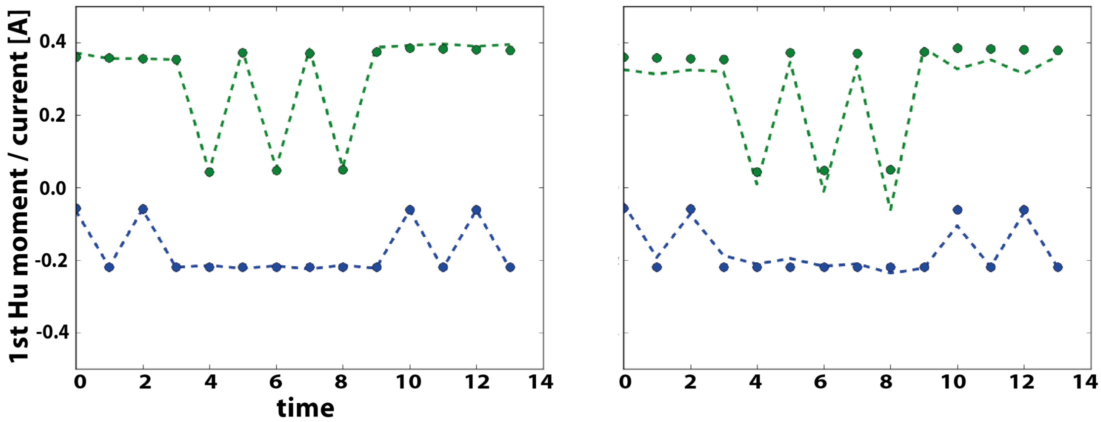


Figure 5.12: Retrieval and generation capabilities. Proprioceptive (green) and visual (blue) dots represent the sampling points of the heavy star-shaped prototype time series (Fig. 5.7). Dashed lines are the time series generated by the network operated in closed-loop with 'clamped' PB values as the only input. The PB values have been acquired unsupervised either during full training (left) or partial training (right). During partial training (right) the network has only been trained with the prototype sequences for the light-weight circle and the heavy triangle. Still, the network is able to generate a fairly accurate sensory prediction for the (untrained) heavy star-shaped object.

To demonstrate the retrieval properties (Ch. 5.2.2) of the fully trained architecture, the PB values acquired during training were 'clamped' to the network. Operating the network in closed-loop mode showed that the input sequences used for training can be retrieved with a very high accuracy. As an example this is shown in Fig. 5.12 (left) for the heavy star-shaped object.

The steps needed until stable PB values are reached, which in turn can be used for recognition, are illustrated in Fig. 5.13. The bi-modal sensory sequences for all light-weight and heavy objects were fed consecutively into the network. On average it took less than 100 steps (about 200 ms) until the PB values converged. The convergence criterion was set to 20 consecutive iterations where the cumulative change of both PB values was $< 10^{-5}$. To assure that the PB values reached a stable state, this number has been successfully increased to 100,000 consecutive steps in preliminary experiments (not shown). Note, that the network and PB values was not reinitialized when the next sensory sequence was presented to the network. Thus, the robot can continuously interact with the toy bricks and is able to immediately recognize an object based on its sensorimotor sequence.

For testing, the network was operated in generalized recognition mode (Ch. 5.2.4). Single trial bi-modal sensory sequences were presented to the network that in turn provided an 'identifying' PB value. The class membership, i.e. which object the robot holds in its hand and how heavy this object is, was then determined based on the minimal Euclidean distance to the PB values of the class prototypes (gray symbols). In Fig. 5.11 the PB values of all 80 single trial test patterns are depicted.

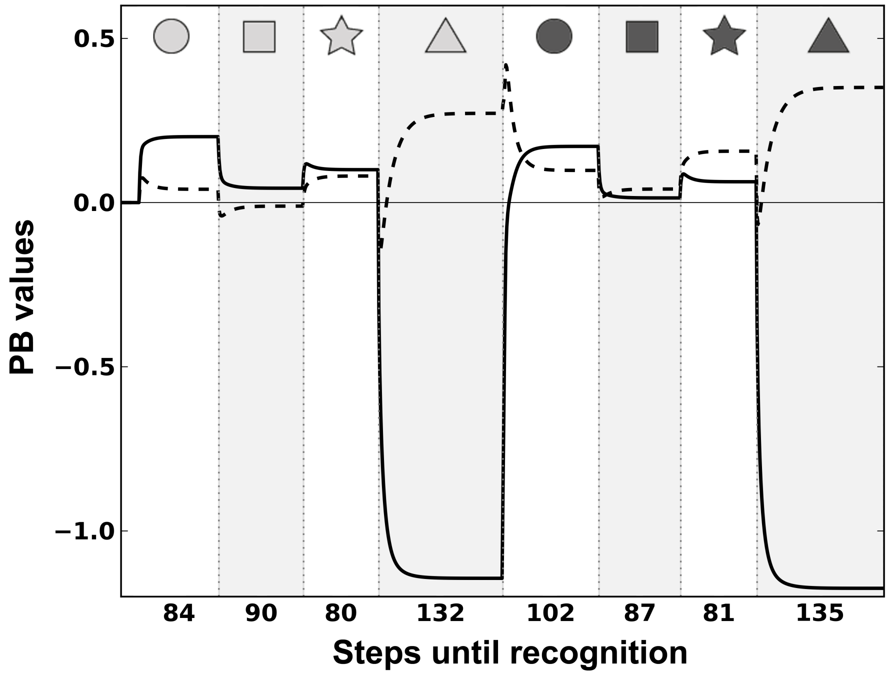


Figure 5.13: Steps until stable PB values are reached. Bi-modal sensory sequences for all light-weight and heavy objects (represented by matching symbols in light and dark gray, respectively) are consecutively fed into the network. The time courses of PB value 1 (solid line) and PB value 2 (dashed line) during the recognition process are plotted.

Only 4 out of 80 objects are misclassified (shown in red), yielding an error rate of 5 %. Interestingly, only star- and circular-shaped objects are confused by the network, which indeed generate very similar sensory impressions (Fig. 5.7). To assess the meaning of the error rate and estimate how challenging the posed problem is, we evaluated the data with two other commonly used techniques in machine learning. First, we trained a multi-layer perceptron (28 input, 14 hidden and one output unit) with the prototype sequences. Testing with the single trial data resulted in an error rate of 46.8 %, reflecting weaker generalization capabilities of the non-recurrent architecture. Next, we trained and evaluated our data with a support vector classifier (SVC) using default parameters (Chang and Lin, 2011). In contrast, this method is able to classify the data perfectly.

5.4.3 Experiment 2 – Classification Using Only the Light Circular-Shaped and the Heavy Triangular-Shaped Object for Training

In experiment 2, only the bi-modal prototypes for the light circular- and heavy triangular-shaped objects were used to train the RNNPB. Although, the absolute PB values obtained during training differ from the ones being determined in the

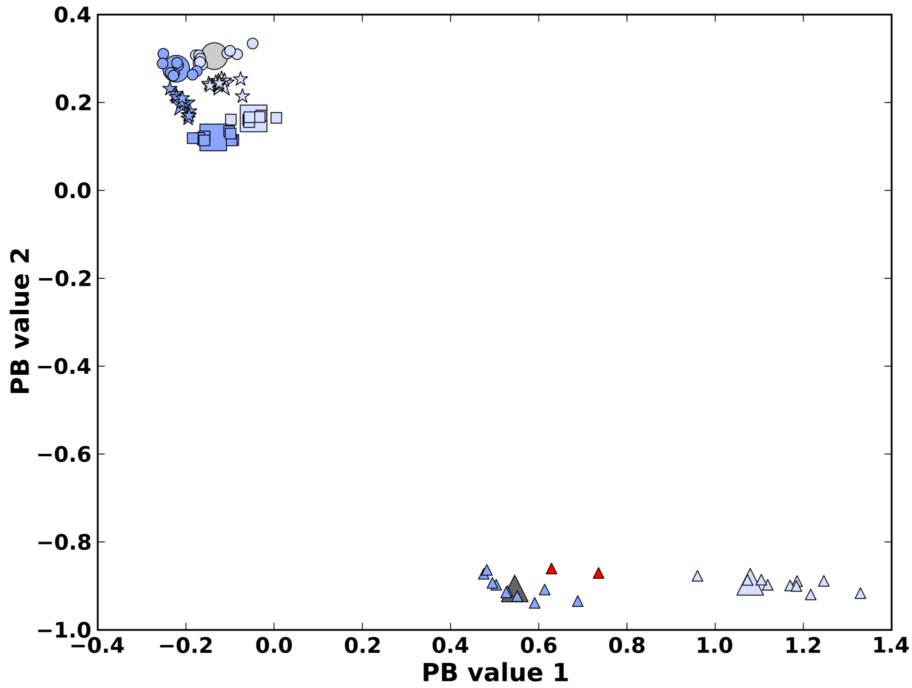


Figure 5.14: Experiment 2 – Classification using only the light circular-shaped and the heavy triangular-shaped object for training. PB values of the class prototypes used for training are depicted in light and dark gray and with a symbol matching the corresponding shape. The a posteriori computed cluster centers of the untrained object categories are depicted using larger symbols in either light or dark blue. Smaller symbols are used for PB values of sensory data of single trials. If the objects have been classified correctly they are shown in light or dark blue, otherwise in red. Light colors are used for light-weight, dark colors for heavy-weight objects.

previous experiment, their relative Euclidean distance in PB space is nearly the same (1.39 *vs.* 1.35), stressing the data-driven self-organization of the parametric bias space.

For testing, initially only the bi-modal sensory time series matching the two training conditions were fed into the network, thereby determining their PB values. Using the Euclidean distance subsequently to obtain the class membership resulted in a flawless identification of the two categories.

Further evaluation of the single trial test data was performed in two stages. In a primary step the remaining test data was presented to the network and the respective PB values were computed (generalized recognition, Ch. 5.2.4). Despite not having been trained with prototypes for the remaining six object categories, the network is able to cluster PB values stemming from similar sensory situations, i.e. identical object categories. In a succeeding step we computed the centroid for each class (mean PB value) and classified again based on the Euclidean distance. This time only two single trial time series were misclassified by the network (error rate 2.5 %). The results are shown in Fig. 5.14.

The generalization potential (Ch. 5.2.4) of the architecture is presented in

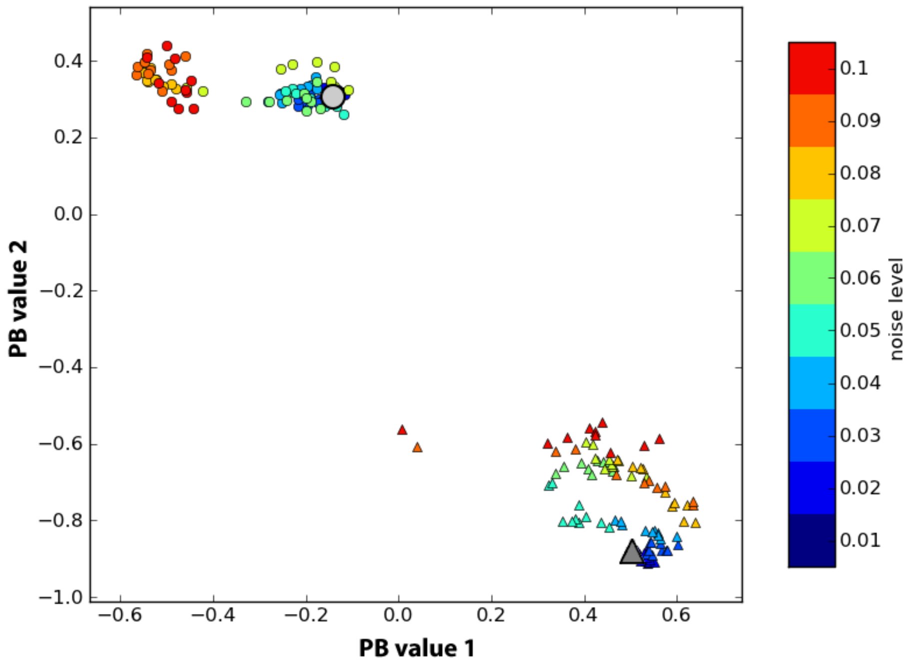


Figure 5.15: Uni-modal noise tolerance. Uniformly distributed noise of increasing levels (color coded) is only added to the visual prototype time series for the light-weight circle and the heavy triangle. The PB values are determined and marked with a matching symbol.The light gray circle and dark gray triangle show the PB values obtained during training without noise.

Fig. 5.12 (right) for the heavy star-shaped object. For this purpose, the mean PB values (centroid of the respective class) were clamped to the network, which was operated in closed-loop mode. The network had only been trained with the light circular- and the heavy triangular-shaped object. Still, it was possible to generate sensory predictions for unseen objects, e.g. the heavy star-shaped toy brick, that match the real sensory impressions fairly well.

5.4.4 Experiment 3 – Noise Tolerance Within and Across Modalities

Based on the network weights that had been obtained in experiment 2 (training the RNNPB only with the bi-modal prototypes for the light circular- and heavy triangular-shaped objects), we evaluated the noise tolerance of the recurrent neural architecture. For this purpose, uniformly distributed noise of increasing levels was either added to the visual prototype time series only (Fig. 5.15) or to the time series of both modalities (Fig. 5.16).

As it can be seen for both conditions, even high levels of noise allow for a reliable linear discrimination of the two classes. Furthermore, the PB values of increasing noise levels show commonalities and are clustered together, again

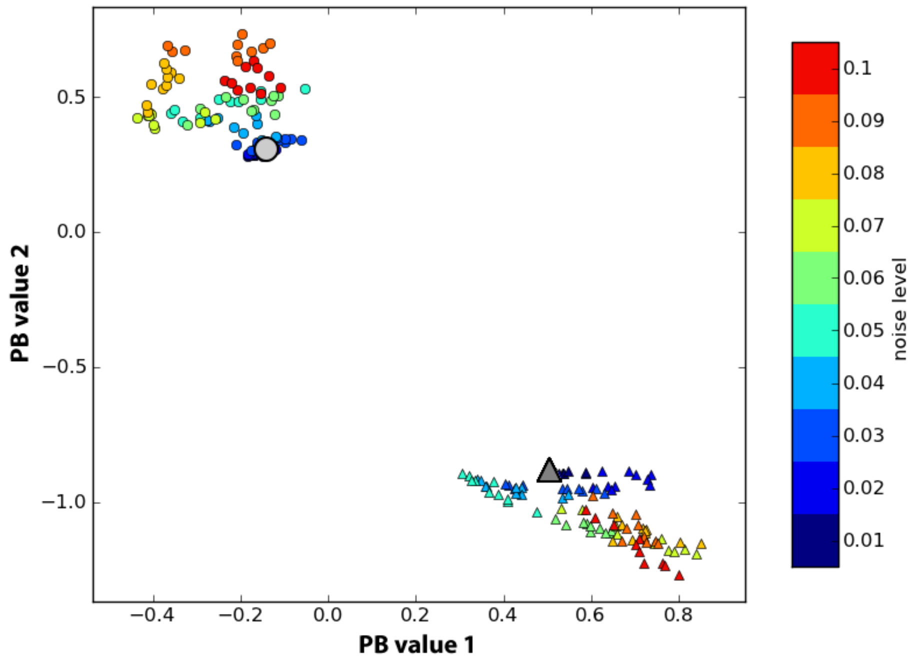


Figure 5.16: Bi-modal noise tolerance. Uniformly distributed noise of increasing levels (color coded) is added to both (visual and proprioceptive) prototype time series for the light-weight circle and the heavy triangle. The PB values are determined and marked with a matching symbol. The light gray circle and dark gray triangle show the PB values obtained during training without noise.

providing evidence for a data-driven self-organization of the PB space. Thus, determining the Euclidean distance of the PB values obtained from the noisy signals to the class representatives enables not only the class membership to be determined, it also allows the noise level to be estimated with respect to the prototypical sensory impression.

5.5 Summary

We present active object categorization experiments with a real humanoid robot. For this purpose, the training algorithm of a recurrent neural network with parametric bias has been extended with adaptive learning rates. This modification leads to a 22-fold increase in training speed. After confirming the improved operation of the new training algorithm we conducted three experiments aiming at object categorization. While holding different objects in its hand, the robot executes a motor sequence that induces multi-modal sensory changes. During learning, these high-dimensional perceptions are 'engraved' in the network. Simultaneously, low-dimensional PB values emerge unsupervised, coding for a sensorimotor sequence that characterizes the interplay of the robot with a specific object. These sequences can be stored and retrieved with a high accuracy.

Further, the geometrical relation of the PB vectors of different objects can be used to infer relations between the original high dimensional time series, e. g. the sensation of a star-shaped object 'feels' more like a circular-shaped object than a triangular-shaped one. Even sensations belonging to unknown objects can be discriminated from known (learned) ones and kept apart from each other reliably. Additionally, we have shown that the network tolerates both uni- and bi-modal noise very well. The results of this chapter have been accepted for publication (Kleesiek *et al.*, 2012).

5.5.1 Connection to the Other Experiments of this Thesis

This chapter represents the final study in a sequence of related studies that build upon each other. In the first experiment, the robot learns to navigate towards a target region (Ch. 3). This is followed by a reaching study (Ch. 4) and the dynamic object recognition task of this chapter. They all have in common that they are inspired by sensorimotor principles originating from the research field of embodied cognition (Ch. 2). According to those themes *perceiving is a way of acting*, stressing the importance of organism-world interactions. Further, the notion that *the world serves as an outside memory* and contributes to the cognitive process as a whole has influenced the design of the experiments. Especially important for this final experiment is the concept of *information self-structuring*, according to which actions of an agent help to structure (sensory) information that is readily available in its environment. This is reflected by the PB values of the recurrent architecture that emerge unsupervised during the storage of the sensory sequences.

Next to those common *sensorimotor design principles* the studies also exhibit a constant technical improvement. In the first study (Ch. 3), the focus lay on the methodology of an artificial neural architecture that combines RL with the learning of sensory features. This paved the way for the development of the novel bio-inspired architecture presented in Ch. 4. Due to an excellent generalization potential of recurrent neural networks and the intriguing properties of the parametric bias units, we decided on this type of architecture for the third study. Again, the knowledge gained previously helped the design of the experiment and improvement of the learning algorithm.

Erst indem wir unsere Sinnesorgane nach eigenem Willen in verschiedene Beziehungen zu den Objekten bringen, lernen wir sicher urteilen über die Ursachen unserer Sinnesempfindungen.
HERMANN VON HELMHOLTZ

Chapter 6
Discussion

6.1 Neural Architectures Based on Sensorimotor Principles

Three computational neural architectures have been presented in Ch. 3 – 5. Despite their different natures and the fact that they have been used to solve different tasks, they do indeed have commonalities. All the architectures, as well as the conducted experiments, are intended to substantiate the intertwined relationship of action and perception. They are all inspired by sensorimotor principles put forward by von Helmholtz, Dewey, Merleau-Ponty, Gibson, Clark, O'Regan and others. These concepts and their importance for the research field of embodied cognition have been presented in detail in Ch. 2.

In all our experiments, the agent actually needs to act to perceive. Based on the hypothesis *perceiving (seeing) is a way of acting* , the agent explores its environment. By doing so, it is able to learn *sensorimotor laws* which subsequently can be exploited for *goal-directed behavior*. Importantly, the architectures do not simply store an internal representation of the world. In contrast, *the world itself serves as an outside memory.* Relying on this external memory enables the neural architecture to learn how actions and perceptions relate to each other.

In the first part of the discussion the three different neural architectures and their corresponding experiments will be discussed individually. In the second part, their joint relevance for the action-driven perception approach will be elucidated. Further, the general significance of computational architectures for the embodied cognition paradigm will be critically examined.

6.2 Reinforcement Learning Architectures

The first two experiments that have been conducted share related two-layer architectures. In both cases, the top layers inherit their structure from classical reinforcement learning algorithms (Sutton and Barto, 1998). However, the lower sensory layers differ in their composition. In the docking experiment of Ch. 3 this layer consists of canonical artificial neurons. To obtain their activations the sum of the products of all ingoing connections and the corresponding weights is computed. In contrast, the lower layer of the reaching experiment presented in Ch. 4 is composed of Sigma-Pi nodes (Softky and Koch, 1995; Weber and Wermter, 2007). The activation of post-synaptic neurons in this model depends on the weighted sum of a multiplicative term, representing the co-activation of input units. For training, both models use a similar mechanism. The RL prediction error δ of the top layer is not only used to modulate learning of action weights, but at the same time to adapt the weights of the sensory neurons of the lower layer, all in a single-step procedure (for details, please see Ch. 3.3 and 4.3, respectively).

In traditional cognitive science and GOFAI, perceptual learning is usually defined as the problem of extracting useful *features* from passively received (visual) stimuli. Subsequently, these features are manipulated to generate an output (e.g. classification result or movement direction). Even in robotics, pre-given or handcrafted features (*cf.* anthropomorphic bias Ch. 2.4.9) are often assumed that are not customized to the motor repertoire of the machine (Pezzulo *et al.*, 2011). In contrast, the agent in our experiments shapes its receptive fields by *enacting* its world. Depending on its *embodiment*, its *situatedness* and the interplay of both it is able to identify the relevant perceptual stimuli and encode their relation to its own movements within the neural architecture, i.e. it is able to identify the *sensorimotor laws*. Furthermore, it can be compensated for wear and tear of the robot by constantly adjusting the RFs to the current situation (lifelong learning). Yet another advantage of our algorithms is their modality independence. For the action-driven learning of the RFs it does not matter if visual, auditory, proprioceptive or any other (sensory) information is given.

The information flow in our architectures during exploitation of the learned skills is reminiscent of classical ideomotor theory (Ch. 2.2.3, James, 1950 [1890]). During learning, actions are linked to their perceptual effects. If the agent is later confronted with a known stimulation, this evokes the corresponding action. This happens directly, showing similarities to the 'directives' of the pragmatic neuroscience framework (Ch. 2.4.5). Admittedly, there is no dynamic interaction between highly distributed neural populations in our model as suggested by Engel (2010).

According to Gibson (1966), organisms seek to discover invariant features that stay constant during movement-induced transformations. He speculates that the acquisition of this information might act in a reinforcing manner (*cf.* Ch. 2.2.7). In fact, the importance of RL for perception has previously been suggested by Woodworth (1947). He claimed that the reinforcing occurs at two levels. First,

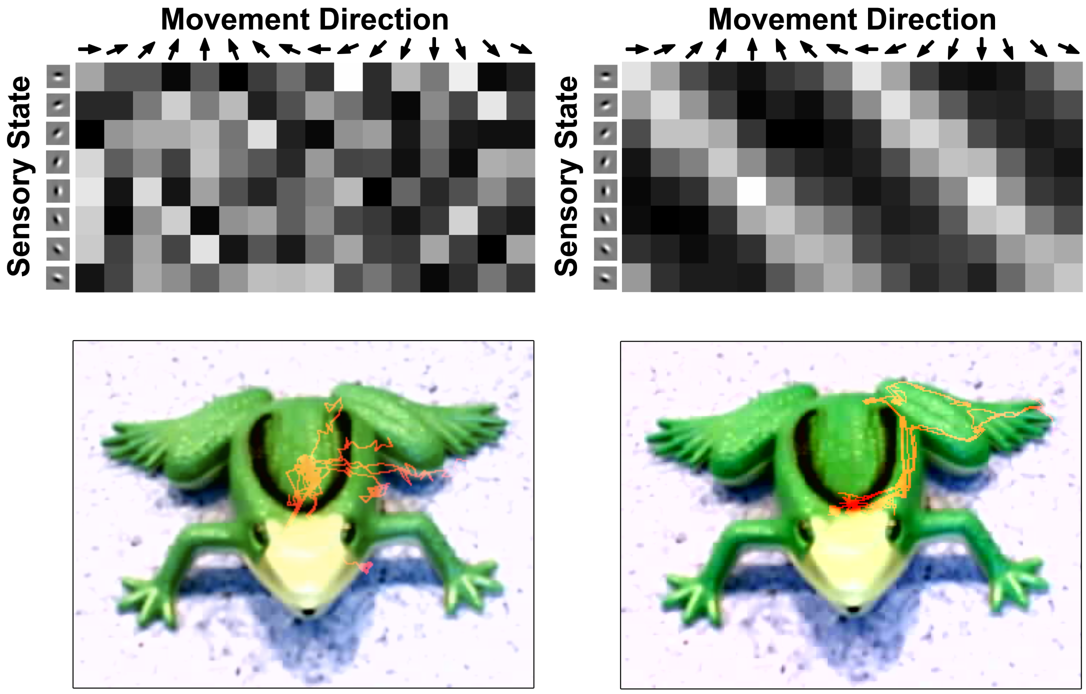


Figure 6.1: Reinforcement learning of sensory invariances. Implementation of the algorithm proposed by Choe *et al.* (2008). The sensory states, i.e. the orientation of the edges, are mapped on actions that keep the internal neural activity constant. At initialization (left) no structure can be seen in the Q-table (top) and the gaze trajectories do not follow edges in a lawful manner (bottom, red lines). After 100,000 steps of training (right) a structure has evolved in the Q-table weights that links orientation and movement direction in an expected way. Furthermore, the resulting gaze trajectories (bottom, red lines) trace edges reliably.

successful gathering of information rewards the explanatory adjustments of the sensory organs “hunting” for clarity. Next, the neural activity generated by the sensory information flow is also reinforced.

Gibson’s theory, however, does not include how invariant information is discovered and applied by an agent. Next to the proposed architectures of this thesis, Choe *et al.* (2008) proposed another computational possibility for detecting internal state invariances without the necessity of an external observer. In their scenario, the gaze of a virtual agent could be moved around natural images with the goal of keeping the internal neural activation constant, i.e. of following edges in the image. Employing RL allowed the learning of actions that correspond to edges with a specific orientation (Fig. 6.1). At the same time receptive fields (RF) evolved with competitive learning that showed a classical Gabor patch-like structure. Furthermore, they showed an enhanced shape recognition based on motor representations in their study.

6.2.1 Reward-Driven Learning of Sensorimotor Laws for Navigation

In the first experiment (Ch. 3) a Webots simulation for navigation and docking towards a virtual target is presented. To solve this hard-delayed RL problem (Gross *et al.*, 1998) we applied a two-layer network, integrating feature and motor learning in a single-step procedure (Weber and Triesch, 2009). As a landmark we used a 3-D geometrical shape (Fig. 3.2), which leads to perspective distortions depending on the robot's position and locomotion (Fig. 3.4). The network learns to exploit this for navigation.

In general, a robot should be aware of the effects of its own actions on objects in the environment and consequently be able to use this knowledge for goal-directed behavior. This is achieved by the presented architecture. The network discovers relevant sensory features and stores this information in the weights of the hidden layer (see receptive fields in Fig. 3.4 and 3.5). Simultaneously, sensorimotor laws are learned which link the current state (comprising physical properties of the object and the sensor, as well as the position of the robot) to a goal-directed action.

Another restriction imposed on an autonomous agent is that noise (e.g. other red objects within the visual field) should not affect performance. This criterion is also fulfilled by our model (ibid.). The sensorimotor laws acquired during learning are grounded in sensorimotor interactions, i.e. perception is, in fact, a sensorimotor experience (*cf. extended mind* Ch. 2.4.3, Clark, 2008). The prediction error δ modulates the learning in a top-down, action-driven way and allows relevant and irrelevant sensory features to be distinguished, despite the fact that the information stemming from the sensory apparatus of the agent can be ambiguous, incomplete and noisy. To improve robustness, a memory layer could be added to the network (Saeb *et al.*, 2009), which would help to maintain focus on relevant features in phases where they are masked by noise or incomplete.

A drawback results from the physical constraints of the robots' camera. Due to monocular vision and a lack of zoom control it is not possible to perform navigation at a distance larger than 1.5 m. The resolution of the camera simply does not permit the geometrical properties of the target marker to be discriminated anymore. Further, without depth information the robot cannot discriminate between different situations stemming from a single line of sight. The perceptual information is blended into a single state. This phenomenon can actually be seen in Fig. 3.5. Including motor knowledge, i.e. parameters for zoom control, would most likely resolve this problem and allow for the development of an increased number of informative receptive fields.

It is conceivable that sensory information from other modalities, e.g. sonar sensors or auditory cues, could be used for robot navigation with the presented model. This could be done in addition, i.e. multi-modal, or separately. As a matter of fact, the architecture should also be capable of learning without any landmark at all, instead exploiting the geometrical shape of the surrounding area of the docking position. However, this is not feasible due to the narrow field-of-view of the built-in camera. While getting closer, the target region is eventually too close to be captured *in toto*.

The presented network learns goal-relevant features within a single united framework. However, if one is willing to give up on training the network with one energy function, a self-organizing map (Kohonen, 2001) could be used to first learn the visual features. Then, this resulting map could represent the state space for RL in a succeeding step. There are further methods of unsupervised learning, e.g. exploiting a sparseness or slowness principle (Wiskott and Sejnowski, 2002), that could also be used to learn the perceptual features. All these models have in common that they do not discriminate between action-relevant and irrelevant features, nor do they solve the delayed RL problem. Therefore, they would be separate parts that would need to be linked *ad hoc* to the RL architecture.

Once the sensorimotor laws have been learned and the visual features needed for navigation have been captured in the weights, the robot is able to navigate to any position where a landmark with the same physical properties is located. This is a clear advantage compared to algorithms that rely on a world model for navigation (*cf.* Ch. 3.1). In our scenario, the landmark serves as an *outside memory* (O'Regan, 1992), a perceptual channel between the world and the neural system that is constantly 'online' (*cf. The Extended Mind* Ch. 2.4.3). Thus, according to Clark's *parity principle* (2008) the landmark could be seen as a *constituent* for cognition, just like Otto's notebook.

Related work

From a technical point of view it is straightforward to first learn the state space, i.e. extract features, with an unsupervised method and then use RL on top of this to find the mapping between states and actions. This two stage learning is a common approach in the literature. For instance, Legenstein *et al.* (2010) trained a simple neural network based on rewards on top of features, which before had been extracted with a hierarchical slow feature analysis network. In contrast, the attention-gated reinforcement learning (AGREL) model of Roelfsema and Ooyen (2005) represents a link between supervised and reinforcement learning. The learning rules lead to the same average weight changes as supervised back-propagation learning. However, learning is slower due to insufficient feedback when the network guesses incorrectly and, hence, the temporal credit assignment problem is not addressed with this model.

Faudzi and Shibata (2010) criticize the way that many researchers treat RL and neural networks as separate modules, the former for learning actions and the latter as non-linear function approximators for the state space. To overcome this shortcoming they propose a method that they call Actor-Q learning. In contrast to our method, their method is designed to have two neural networks, called Q-net and Actor-net, respectively, in parallel. The Q-net is responsible for choosing an action, whereas the Actor-net controls the movements. During training each subnet generates and employs its own error signal. For a given task, i.e. camera motion for the identification of two patterns, they were not able to achieve successful recognition within a reasonable number of trials when training the network simultaneously. However, when first learning camera motion (*cf.* motor babbling, Ch. 2.4.1) and then later on including training for pattern recognition, they succeeded with their task.

Future work

Future work should address several issues. Most importantly, the action repertoire of the agent should be augmented to include turning movements. In fact, most of the 5 % unsuccessful trials (red trajectories, Fig. 3.6) are caused by motion-induced rotation which cannot currently be compensated by the robot. Further, an implementation on a real robot, possibly including additional modalities, would be desirable. However, it has to be considered that this might not be computationally tractable given contemporary hardware.

6.2.2 Sigma-Pi Reinforcement Learning for Reaching

In Ch. 4 a novel two-layer architecture is proposed that is capable of learning invariant hand-object relations and a corresponding movement of the hand towards the target. The network handles implicit frame of reference transformations, unaffected by which modality the sensory information comes from or what coordinate system this information is encoded in. The biological inspiration is drawn from neurons in the parietal reach region, which keep track of the distance between a target and the current position of the hand, presumably to serve as an error signal for reaching movements (Chang *et al.*, 2009). The choice to utilize Sigma-Pi neurons for the lower layer has been influenced by a mechanism called gain modulation, a way to combine several sources in a non-linear way, e.g. a multiplication (Pouget *et al.*, 2002). For details regarding the biological inspiration, please refer to Ch. 4.1.1.

The functioning of the method is demonstrated in a simulation of a 2-D grid world. Given the position of the hand of a virtual agent, as well as the position of the target object, all possible invariant hand-object relations and corresponding movements can be learned perfectly (Fig. 4.6 and 4.7). Despite mastering the given task flawlessly, the method has a major limitation. The number of weights in the network grows exponentially (Tab. 4.1) and just for a world of size 5×5 approximately 57,000 synapses have to be adjusted by the training algorithm. As a matter of fact, this *curse of dimensionality* (Bellman, 1957) does not permit the network to be trained on a world larger than 4×4 on contemporary desktop computers. Alternatively, we solved the reaching task with two related learning methods and compared the results to our unified architecture. The first model of comparison is a reinforcement learning scheme based on a cubic Q-table, the other one a two-step paradigm where unsupervised learning of Sigma-Pi neurons (Weber and Wermter, 2007) is followed by a canonical RL algorithm. Unfortunately, these two alternative models do not scale very well either. The best scaling goes along with the two-step procedure. However, at the same time this procedure leads to an average error rate of 5–6 %. A potential explanation for this weaker outcome could be the missing influence of actions during unsupervised learning, which is based on the self-organization of sensory information only.

The analysis of our novel unified method reveals some interesting properties. Indeed, the model captures law-like relations between motor actions and sensor readings. A top-down action-driven mechanism structures the sensory information that is readily available in the environment. Thus, the world serves again as an outside memory. After learning, when confronted with a known stimulation, this sensation is directly translated into an action, i.e. reaching for the target object, reminding of ideomotor theory and the directive minds (see above, Ch. 6.2).

Interestingly, if illegal actions, i.e. movements that leave the specified area, are not penalized, the reaching task is not learned perfectly. Similarly, if the approximation to the learning rule of the lower sensory layer, which omits non-local terms from the action layer, is used, the results are impaired as well. On the contrary, including an adaptive learning scheme for the action layer (and not for the sensory layer) leads to a 10-fold increase in learning speed. In fact, all these observations might have a common cause: actions have a tremendous effect on the learning of perceptual stimuli. This action-driven perception is also affirmed by the development of the Q weights, which is slightly faster than the development of the Sigma-Pi weights. All these findings stress the influence of actions for the learning of perceptions and is, for instance, in agreement with the result of Faudzi and Shibata (2010, *cf.* Ch. 6.2.1).

Related work

Kuperstein (1988) demonstrated that adaptive hand-eye coordination results in successful reaching. He trained a neural model so that a multi-joint arm could reach for a cylinder arbitrarily placed in space. For this purpose, self-generated motor commands were used to explore many different arm positions while holding a cylinder in the hand. During this process the associated topographic sensory information was stored in maps. This knowledge later on allowed the system to learn to reach for the object solely based on sensory and motor feedback. Similar work has been presented by Andry *et al.* (2004). A robot learns associations between vision and arm movements and stores this information in visuo-motor maps, which are eventually exploited for reaching via a neural network control mechanism. Recently, Rolf *et al.* (2011) modeled early goal-directed movements of infants and showed that inverse models can actually be bootstrapped within a few hundred movements. Using a simulated arm with up to 50 degrees of freedom their statistical online algorithm could successfully reach in a 2-D plane. However, employing the algorithm in a 3-D world had a severe impact on the scaling (personal communication), emphasizing the challenging nature of the problem.

6.3 Recurrent Architecture

Recurrent neural networks can be conceived as dynamical systems (*cf.* Ch. 2.4.7). A special type of such an architecture is called RNNPB, a canonical RNN extended with so-called parametric bias units (Tani and Ito, 2003). We modified a variant proposed by Cuijpers *et al.* (2009) with an adaptive weight update regime (Ch. 5.2) and used this novel algorithm to solve a robotic object categorization task. The PB units of the recurrent architecture emerge unsupervised ('self-tuned') during training of the network and can be regarded as fixed-points of this dynamical system that 'resonate' to object-specific perceptual stimuli. This *self-tuning* and *resonating* is reminiscent of Gibson's ecological theory of perception (*cf.* Ch. 2.2.7). Gibson (1966) and others (Ch. 2) suggested that perception is indeed an active process, requiring that an organism moves around in its environment to learn about dynamic sensory structures. Depending on the morphology of the organism, different kinds of actions might be adequate to unveil the necessary information and to *self-structure* them (*cf.* Andy Clark's six elements of EC research, Ch. 2.4.3). This is exactly what is happening in our experiments. A humanoid robot moves toy bricks up and down and rotates them back and forth, while holding them in its hand. Visual and proprioceptive information of this process is structured using an RNNPB. This helps to distinguish different sensory experiences and mimics *seeing is a way of acting* (O'Regan and Noë, 2001).

6.3.1 Active Perception using an RNN with Parametric Bias

By introducing modifications to the learning algorithm of the RNNPB we were able to achieve a significant 22-fold increase in speed (Tab. 5.2) for the storage of the two 1-D signals shown in Fig. 5.5. It was also confirmed that the storage and retrieval of those time series was stable and that learning converged in a well-behaved manner (Fig. 5.9 and 5.10). Admittedly, the storage of other sequences with e.g. a different dimensionality, length or dynamic, may well result in a different increase in performance.

After confirming flawless operation of the training algorithm we conducted three experiments aiming at object categorization, a fundamental cognitive ability (Shapiro, 2011). While holding different objects (Ch. 5.3) in its hand, the robot executes a motor sequence that induces multi-modal sensory changes. During learning these high-dimensional perceptions are 'engraved' in the network. Simultaneously, low-dimensional PB values emerge unsupervised, coding for a sensorimotor sequence characterizing the interplay of the robot with an object. We show that 2-D time series of length $T = 14$ can be reliably represented by a 2-D PB vector and that this vector allows learned sensory sequences to be recalled with a high accuracy (Fig. 5.12 left). Furthermore, the geometrical relation of PB vectors of different objects can be used to infer relations between the original high dimensional time series, e. g. the sensation of a star-shaped object 'feels' more like a circular-shaped object than a triangular-shaped one. Due to the experimental noise of single trials, identical objects cause varying sensory impressions. Still, the RNNPB can be used to recognize those (Fig. 5.11). Additionally, sensations belonging to unknown objects can be discriminated from known (learned) ones. Moreover, sensations arising from different unknown objects can be kept apart from each other reliably (Fig. 5.14).

Humans are able to immediately divide the perceived world into different physical objects, seemingly without effort, even when they are confronted with previously unseen objects. Indeed, it makes perfect sense that the discrimination between different sensory qualia is possible without training (Ch. 5.4.3). However, actively generating (retrieving) sensorimotor experiences does require training and generalization capabilities. Similar findings have been reported recently for humans (Held *et al.*, 2011). Previously blind subjects, regaining sight after a surgical procedure, were able to visually discriminate different objects right away. Cross-modal mappings between seen and felt, however, had to be learned.

Comparing the classification results of the fully trained RNNPB with the SVC reveals a superior performance of the support vector classifier. Nevertheless, it has to be kept in mind that the maximum margin classifier cannot be used to generate or retrieve time series. Interestingly, the error rate is lower if the recurrent network is only trained with two object categories (Ch. 5.4.3). A potential explanation, besides random fluctuations, could be that during training a common set of weights has to be found for all object categories. This process presumably interferes, due to the challenging input data, with the self-organization of the PB space.

A drawback of the presented model is that it currently operates on a fixed motor sequence. It would be desirable if the robot performed *motor babbling* (*cf.* Ch. 2.4.1) leading not only to a self-organization of the sensory space, but to a self-organization of the sensorimotor space. A simple solution to this problem would be to train the network additionally with the motor sequence most appropriate for an object, i.e. reflecting its affordance (*cf.* Ch. 2.2.7, Gibson, 1977). This would lead to an even better classification result because the motor sequences themselves would help to distinguish the objects from each other and, thus, the emerging PB values would be arranged further apart in PB space (conversely, this means currently it does not make sense to train the network with the identical motor sequences in addition). However, that does not address the fact that the robot should identify the object affordances, the movements characterizing an object, by itself. For a possible extension mechanism specifically addressing this issue, please see the section on future work below.

De-noising

There are several potential applications of the presented model. As shown in Fig. 5.15 and 5.16, the network tolerates noise very well. This fact can be exploited for sensor de-noising. Despite receiving a noisy sensory signal, the robot still will be able to determine the PB values of the class representative based on the Euclidean distance. In turn, these values can be used to operate the RNNPB in retrieval mode (Ch. 5.2.2), generating the noise-free sensory signal previously stored, which then can be processed further. In fact, Körding and Wolpert (2004) suggested that the central nervous system combines, in nearly optimal fashion, visual, proprioceptive and other sensory information to overcome sensory and motor noise. Next to their Bayesian framework an RNNPB might also be a possible way to model this 'de-noising' happening in the brain.

Sensorimotor Imagery

It is also conceivable that the network is used to model sensory (sensorimotor) imagery. Due to the powerful generalization capabilities of the network, not only the trained sensory perceptions can be recalled, but interpolated 'feelings' can be generated (Fig. 5.12 right). For instance, setting (top-down modulated) a PB value allows how an object feels and what actions are associated with it to be imagined. It has been reported that visual and motor imagery share the same neural circuits as actual perceptions and actions do (Jeannerod, 1995). Further, they are conditional on the same constraints, e.g. timing and metric spatial information. This blends in with the concept of simulators, which can be seen as dynamical systems that generate experience depended context-specific simulations (Barsalou, 1999; 2008; 2009). A simulation is "the re-enactment of perceptual, motor, and introspective states acquired during experience with the world, body, and mind" (Barsalou, 1999, pp. 618–619). Based on associative mechanisms among modalities the concept of simulations might actually lead to the development of categorical representations (ibid.). Indeed, one is tempted to claim that the PB values of the presented experiments reflect such a development.

It is also known that the neural circuits underlying motor imagery and motor preparation are tightly coupled (Cisek and Kalaska, 2005). This is important for the planning of future actions and the prediction of their perceptual consequences. Of course, this immediately reminds one of the notion of forward models and the related efference copy principle (Ch. 2.4.1). Based on PB values a sensory prediction can be generated. This prediction can be compared to the reafference signal, i.e. the real sensory feedback, and in this way allows ambiguities to be resolved that might be present in sensory information. In turn, this knowledge can then be exploited to control behavior or to discriminate intrinsically from extrinsically induced sensations. According to Clark and Grush (1999), forward models (emulators) are a simple way to predict the next sensory and motor state of a system. Decoupling of the forward model from reality leads back to the concept of simulations (emulations). For instance, this principle has been exploited for obstacle avoidance in a computational study of Hoffmann (2007, *cf.* Ch. 2.4.9).

Garbarini and Adenzato (2004) establish a link between simulation and mirror neurons (Gallese *et al.*, 1996). "In fact, a third term must be added to the relation between *action* and *perception*, i.e. that of *simulation*. While observing an object, the neural system is activated as-if the observer were interacting with it" (Garbarini and Adenzato, 2004). This view, however, is discussed controversially (Hickok, 2009). Despite correlations that have been found in monkeys between the firing of mirror neurons and the observation of actions, evidence that this may indeed have influence on how an organism perceives actions is still very limited (Iacoboni, 2009). Nevertheless, the RNNPB architecture has been used to model the mirror neuron system previously (Tani *et al.*, 2004; Cuijpers *et al.*, 2009).

Related work

In related research, Ogata *et al.* (2005) also extract multi-modal dynamic features of objects, while a humanoid robot interacts with them. However, there are distinct differences. Despite using fewer objects in total, the problem posed in our experiments is considerably harder. Our toy bricks have approximately the same circumference and identical color. Furthermore, they exist in two weight classes with an identical in-class weight that can only be discriminated via multi-modal sensory information. We provide classification results, compare the results to other methods (MLP and SVC) and evaluate the noise tolerance of the architecture. In addition, only prototype time series are used for training (in contrast to using all single-trial time series), resulting in a reduced training time. Further, it is demonstrated that, if the network has already learned sensorimotor laws of certain objects, it is able to generalize and provide fairly accurate sensory predictions for unseen ones (Fig. 5.12 right).

The stimulus response compatibility effect establishes a clear link between visual and motor processes (Ellis *et al.*, 2007). In a series of psychological experiments it is demonstrated that the visual categorization of humans is influenced by the micro-affordances, e.g. precision grip *vs.* power grasp, associated with the objects. This phenomenon has been implemented on a simulated iCub robot (Macura *et al.*, 2009) using a Jordan-type RNN (Kolen and Kremer, 2001) at its core. In this study, four different objects have to be discriminated by the artificial agent, specifically two big objects and two small objects have to be kept apart (ball and cube in each case). During training the robot is trained to grasp the objects with either a precision or power grip, depending on the size of the object. Based on these different grips and the associated embodied knowledge acquired during training the objects can be categorized successfully. Not surprisingly, the error of trials for congruent conditions (categorization grip is in agreement with the grip matching the size of the object) was lower than for incongruent ones (mismatch between used and appropriate grip). This was interpreted to be in agreement with psychological effects observed in human studies, where subjects showed shorter reaction times in congruent than in incongruent trials.

Future work

As has been noted above, it would be desirable if the robot were able to identify the actions by itself that correspond to the most discriminative features of the objects. At the same time, the affordances of the objects should also be acquired and utilized for the categorization task, giving the experiments a truly embodied flavor. To accomplish this goal, the RNNPB could be combined with methods used in sensor (view) planning. These methods deal with the problem where a robot has to 'look' for an unambiguous identification of an object (*cf.* Ch. 2.5).

The resulting sensorimotor sequence could then be stored in the RNNPB, leading to representative PB values for the respective category. Due to the fantastic generalization potential that has been demonstrated, the PB values of unknown objects would probably cluster in proximity to the trained ones, reflecting similar affordances and informative view sequences of unknown objects, which in turn could not only be used for recognition but also for sensory and motor imagery as well as for 'offline' movement planning (see above). Next to Bayesian methods, RL is also conceivable for realizing this challenging task. For a review of additional methods used for view planning, please see Roy *et al.* (2004).

Another study that could readily be conducted with the established framework is the modeling of the experimental observation that actions alter shape categories (Smith, 2005). Ross *et al.* (2007) investigated in a VR experiment the influence of arbitrary actions during category learning on the recognition of objects. During the recognition phase the subjects either performed consistent or inconsistent movements (w.r.t. to the actions executed during training). It was shown that consistent movements facilitated the categorization task, letting the authors conclude that arbitrary action information is incorporated into object representations. Using the RNNPB to store different motor sequence in conjunction with the sensory impressions of identical objects will result in distinct PB values. It would be interesting to investigate if the number of recognition steps (and the Euclidean distance) show similar effects to the ones obtained from the human experiments when feeding the network with consistent and inconsistent sensorimotor sequences, respectively.

6.4 Action-Driven Perception – A Critical Assessment

6.4.1 Ingredients for Embodied Cognitive Systems

Are the presented architectures (Ch. 3 – 5) really (prime) examples for embodied cognitive systems? Hardly anyone (not even a traditional cognitive scientist) doubts the (causal) influence of action on perception and that the physical properties of the world contribute to perception. Including a (learned) description of the body properties in the mind's symbol manipulation algorithm, makes it possible for a traditional cognitive science approach to be compliant with an embodied approach. It could also be argued that the contact area of the body with the world generates symbols which in turn can be processed within the brain (Shapiro, 2011). Thus, computational modeling seems to meet the requirements of first-generation cognitive science in the sense of being a solipsistic solution to a given problem. But how can it be shown that cognition of a computational model with a clearly specified I/O interface extends beyond the boundaries of its artificial (neural) architecture?

Andy Clark (2008) proposed six ingredients that capture the general spirit of an embodied approach (*cf.* Ch. 2.4.3). "[A] given research project may not include all six of these elements, but its adoption of several should suffice to distinguish it from research within more standard cognitive science, as well as to ally it with other embodied approaches." (Shapiro, 2011, p. 61). Interestingly, one of Clark's elements, the *Dynamic-Computational Complementarity* establishes a bridge to canonical cognitive science, because in his view certain explanatory concepts, like computation and representation, are vital for the understanding of cognition[1]. Next, to this element our experiments are also in accordance with *Open Channel Perception*, *Information Self-Structuring* and *Perception as Sensorimotor Experience*. Thus, based on this classification they can be conceived as embodied research projects.

And yet, other authors claim "to implement a truly embodied cognitive system, multiple modalities are essential. In addition to sensory and motor modalities, internal modalities, including affect, motivation, and reward, are essential from the embodied perspective" (Pezzulo *et al.*, 2011). In a qualified sense, these characteristics are also in compliance with our experiments. The RL architectures could easily be extended to include additional modalities and they already include a reward system. On the other hand, the recurrent architecture comprises, next to visual sensory impressions, also proprioceptive information and the proposed extension, which is supposed to include sensor planning, aims at incorporating further elements like attention and possibly reward.

Another counterargument suggesting that our architectures do not qualify as embodied cognitive systems could be based on the point that some of our experiments were only performed in simulation. However, even if simulations do not cover the full complexity of real environments and their transferability is not always granted it is still believed that they are very valuable for cognitive science research (Ziemke, 2003a).

6.4.2 Robotics – A Valuable Tool for EC Research?

It has been argued that robots cannot be *embodied*. "The reason that Artificial Intelligence originally adopted the view of a 'disembodied brain' is that a robot 'is' disembodied: it is just a container for a mind. Our bodies are not mere containers of minds: our bodies have been shaped by evolution to be the natural object and subject of the mind" (Scaruffi, 2006). Until now, self-evolving robots are still dreams of the future. Therefore, to cope with this argument, we used a state-of-the-art humanoid robot for our experiments. Further, we demonstrated that due to the nature of our algorithms the processing, i.e. the cognitive abilities, extend beyond the neural architectures and are general enough to adopt to the physical properties of the body. Nevertheless, this 'evolution' argument and another reasoning, namely "that *machines act according to plans* (their human designers'), whereas *living organisms are acting plans*" (Ziemke, 2003b) are hard

[1] Also Shapiro (2011) thinks that "retaining [the traditional cognitive science] conceptual apparatus, when possible, seems reasonable" (p. 205).

to confute. Nevertheless, as it has been shown in our experiments, if a robot self-acquires *sensorimotor laws*, i.e. relating its own actions to sensory impressions, this *anthropomorphic bias* can be minimized, if not totally abandoned.

Cognitive robotics and theories for embodied cognitive systems should cross-fertilize each other (Ch. 2.4.9). Indeed, it has been questioned “how designing efficacious robots could be a convincing argument for psychologists and neuroscientists for or against a certain theory” (Pezzulo *et al.*, 2011). Currently, it seems that the focus in robotics research is on the design of better autonomous machines rather than on trying to understand the neural mechanisms of the brain. Biological-inspired theories help to improve robots (Wermter *et al.*, 2005), but reciprocal impact for refining neuroscientific hypotheses is rather scarce. This might be due to the fact that “most current computational models are mere *proofs of concept* and lack the adequate level of detail to start deriving precise predictions, or to simply be considered as useful tools by psychologists and neuroscientists” (Pezzulo *et al.*, 2011). This might be true, however there are several ways where results from robotic experiments can be beneficial for EC theories. Most importantly, converting theories into algorithms helps to think through the proposed conjectures and, hence, it might point out shortcomings in one’s reasoning. Further, robot demonstrators open up new avenues of research and might help the research field of embodied cognition to become a well-defined paradigm.

Considering our experiments, some of these advantages become apparent. Bearing the core concepts of EC research in mind during the design of our algorithms, we were able to obtain systems that ground perception in *sensorimotor principles*. This paved the way for several lines of future research (see above). Furthermore, the results of our bio-inspired reaching experiment (Ch. 4) have led to novel insights. In principle, the agent learned to accomplish the given task perfectly. However, the proposed algorithm is not computationally tractable. Thus, due to the unrealistic number of training examples we are able to refine both the proposed architecture and our theory.

6.4.3 Using Computational Architectures to Prove a Theory

We have demonstrated that our architectures realize two of the core concepts proposed by the sensorimotor contingencies theory (*cf.* Ch. 2.4.4, O’Regan and Noë, 2001): *seeing is a way of acting* and *the world serves as an outside memory*. These concepts have indeed been helpful for the design of the computational models and the robotics experiments, but does this suffice to prove that the SMCs theory itself is correct? Even mimicking a phenomenon described as a prime example of this theory, e.g. change blindness, does not allow this conclusion to be drawn. In general, devising a novel mechanistic theory and then implementing it on a machine will most certainly yield the desired phenomenological results. The underlying reason for this has been summarized by Beer (in press):

> “[W]e must recognize that computationalism, connectionism and dy-

> namicism are not really scientific theories at all, because they themselves do not make falsifiable predictions."

Making the situation even worse, the *cognitive subject* who is doing the describing, implementing and interpretation of the experiment, influences the whole process additionally (Froese, 2010). A scientific experiment cannot be an observer-independent reality because "the environment as we perceive it is our invention" (von Foerster, 1988).

6.4.4 Dynamical Systems

Van Gelders (1995) dynamical hypothesis and his example of the Watts governor for steam machines has been presented in Ch. 2.4.7. This system can be formulated and solved with differential equations. Being a coupled system suggests that all parts of it are really constituents[2] of the process and raises the question of whether representations do exist at all in this model. Nevertheless, the same problem can be formulated using representations leading to a solution rooted in cybernetics (Ch. 2.2.5) that works equally well, and more importantly, also explains the phenomenon, albeit differently[3]. Which solution should be taken for granted now? How does the 'brain' of the steam machine work and "can we really benefit from conceiving of cognitive processes as dynamical systems" (Shapiro, 2011, p. 123)?

Another dynamical concept, the free-energy principle, has been proposed in order to account for action, perception and learning (Friston, 2010). According to this theory, the brain seeks to minimize its free energy via optimization of value (expected reward, expected utility) or its complement surprise (prediction error, expected cost). This can be formulated with three sets of differential equations, one set that describes how actions and neuronal representations of expected states change with time, another set that reflects the agent's model of how sensory data are generated and, finally, a set that formalizes the prediction error of the brain. Again, the coupled systems are fully described. For instance, if the agent's equations of motion for the mountain car problem are known, the principle works very well (Friston *et al.*, 2009). But can it be assumed that equations for the complete environment and brain are given?

It is not only for these reasons that recurrent neural architectures, which can either be learned or evolved, are an interesting dynamical alternative. Beer's (2003) categorical perception experiment that uses a CTRNN has been presented in Ch. 2.4.7. In our experiments we also perform object categorization. For this purpose, we employ an RNNPB to self-structure the information present in multi-modal data. Both studies demonstrate in an impressive way how actions are indeed fundamental for perception. If the agent did not move (itself or the object) it would be insensate. Moreover, in our experiment passive movements are not

[2] *cf. coupling-constitution fallacy* Ch. 2.4.3

[3] Yet another example is Ashby's homeostat, whose behavior appears to be intelligent to an external observer but can potentially be realized by a simple mathematical system (Ashby, 1947).

sufficient to differentiate between heavy and light objects because proprioceptive information necessary for successful categorization is missing.

6.4.5 Representations

In contrast to Beer's categorical perception where the agent's behavior is used for object classification, the steady state (stable PB value representing a fixed-point) is used to determine the object in our system. This might attribute a rather *representational* approach to our experiment. According to van Gelder (1995), representations in computational systems are discrete and deterministic states, they *stand in* for something, which is used subsequently for computation. This is not the case for his experiments, Beer (2003) claims. Both the object and the agent always move and, thus, do not reach an equilibrium point. Does this make our approach to object recognition less embodied, less cognitive? In fact, observing the neural activation patterns in Beer's experiments *over time* does allow which object the agent is faced with to be determined, it represents the state of the system. However, Beer (ibid.) would reply:

> "Rather than assigning representational content to neuronal states, the mathematical tools of dynamical systems theory are used to characterize the structure of the space of possible behavioral trajectories and the internal and external forces that shape the particular trajectory that unfolds. Indeed, a dynamical approach to situated action raises important questions about the very necessity of notions of representation and computation in cognitive theorizing."

Admittedly, in both experiments (Beers and ours) the meaning of the categorization, i.e. which object is 'felt' by the agent, is assigned by an external observer. Without this, a self-structuring of the action-perception cycle of agent-world interactions still happens. Nevertheless, to interpret the processes, representational vehicles are most certainly necessary. Also for higher cognitive functions, e.g. transfer abilities in comparison to purely reactive behavior or sensory and motor imagery of an object (see above), some sort of memory system is necessary. Following the same notion, Shapiro (2011) states that the continuous interaction between agent and object "without representation is useless". Furthermore, Ward and Ward (2009) claim to have found *weak-substantive representations* (*cf.* Ch. 2.4.3, Clark, 1997b) after performing an extensive behavioral analysis of an agent's discrimination abilities, which was obtained by replicating Beer's (2003) experiments. In contrast, the PB values of our experiments rather reflect strong representations (Clark, 1997b), which refer to internal states that can also be used offline for planning or mental simulations (see above).

6.5 Conclusion

We have presented three different artificial neural architectures and discussed their relevance for an action-driven perception paradigm within the research field of embodied cognition. Actions help to resolve ambiguities and enhance the perceptual experience and, hence, considering their importance for perception, contribute to the design of better robotic systems. Inspired by principles rooted in various disciplines (*cf.* Ch. 2), we demonstrated that the concept of action-driven perception can be used to operate control structures (*cf.* experiments Ch. 3 and 4) and to determine how objects 'feel' (*cf.* experiment Ch. 5).

In all our experiments the agent actually needs to act to perceive. Based on the hypothesis *perceiving (seeing) is a way of acting* it interacts with its environment, thereby learning *sensorimotor laws* that can be exploited for *goal-directed behavior*. The agent relies on information that is readily available in the environment and, thus, uses *the world as an outside memory*. Further, the concept of *information self-structuring*, according to which actions help to structure (sensory) information, can be found in all our studies.

But how can other researchers benefit from the underlying *sensorimotor design principles*? The scope of this thesis was not to prove that actions are fundamental for perception. Evidence for this hypothesis has been given in Ch. 2. Instead, we used the sensorimotor principles listed in Ch. 1.2 as an inspiration for the design of our architectures and experiments. We have shown, if one follows these guidelines that indeed, artificial neural architectures can be devised or refined that allow different robotic tasks to be solved successfully. However, this does not mean that the proposed architectures and the underlying design principles are in general better suited for all sorts of robotic tasks. The claim that the design principles are superior to other approaches cannot be proven because computational theories (e.g. modeling studies) themselves are not falsifiable. Just because a solution performs better on a given test set does not mean that the theory itself is correct (w.r.t. neuroscientific plausibility). In general, the intertwined relationship of action and perception should be kept in mind by every researcher who designs (robotic) experiments. However, if the goal is to make 'better robots', the method which works best should be adopted, even if it is not biologically inspired. Especially hybrid architectures, e.g. combining neural networks with a statistical method, can be very helpful and may even scale better than the proposed methods.

The experiments comprise an entire sequence of several related tasks. In the first experiment, the robot learns to navigate towards a target region (Ch. 3). This is followed by a reaching study (Ch. 4) and a dynamic object recognition task where a humanoid robot moved objects up and down and rotated them back and forth, while holding them in its hand (Ch. 5).

Besides being a sequence of related tasks, the conducted experiments also represent the evolutionary process of finding suitable artificial neural architectures and appropriate experiments for *action-driven* learning based on *sensorimotor principles*. In the first study (Ch. 3), the focus was on the methodology of this special type of artificial neural architecture. This paved the way for the development of the novel bio-inspired architecture presented in Ch. 4. Due to an excellent generalization potential of recurrent neural networks and the known self-organization of the parametric bias units, we decided on the RNNPB architecture for the final experiment (Ch. 5). Again, the knowledge gained previously helped the design of the experiment and the improvement of the learning algorithm.

The main scientific contributions of this thesis can be summarized as follows. First, we introduced an innovative navigation paradigm that is independent of a world model, because *the world itself serves as an outside memory.* Second, we proposed a novel bio-inspired neural architecture that combines reinforcement learning and Sigma-Pi neurons. Third, we extended an RNNPB with an adaptive learning regime, leading to a drastically reduced training time. In addition, several future experiments have been suggested that can be conducted based on the theoretical and methodological framework established within this thesis.

Appendix A
References

Adams, F. and K. Aizawa (2001). “The Bounds of Cognition”. *Philosophical Psychology* 14.1, pp. 43–64.

Adams, F. and K. Aizawa (2010). “Defending the bounds of cognition”. *The extended mind* 1998, pp. 67–80.

Adams, F. and K. Aizawa (2009). “Why the Mind is Still in the Head”. *The Cambridge Handbook of Situated Cognition*, pp. 78–95.

Adams, F. and K. Aizawa (2008). *The Bounds of Cognition*. Wiley-Blackwell.

Aloimonos, J., I. Weiss, and A. Bandyopadhyay (1988). “Active vision”. *International Journal of Computer Vision* 1 (4), pp. 333–356.

Amari, S (1977). “Dynamics of pattern formation in lateral-inhibition type neural fields”. *Biol Cybern* 27.2, pp. 77–87.

Andersen, R. A., R. M. Bracewell, S Barash, J. W. Gnadt, and L Fogassi (1990). “Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque”. *J Neurosci* 10.4, pp. 1176–96.

Anderson, M (2003). “Embodied Cognition: A field guide”. *Artificial Intelligence* 149.1, pp. 91–130.

Andry, P., P. Gaussier, J. Nadel, and B. Hirsbrunner (2004). “Learning Invariant Sensorimotor Behaviors: A Developmental Approach to Imitation Mechanisms”. *Adaptive Behavior* 12.2, pp. 117–140.

Asada, M., S. Noda, S. Tawaratsumida, and K. Hosoda (1996). “Purposive Behavior Acquisition for a Real Robot by Vision-Based Reinforcement Learning”. *Machine Learning* 23 (2), pp. 279–303.

Ashby, W. R. (1947). “The nervous system as physical machine; with special reference to the origin of adaptive behaviour”. *Mind* 56.221, pp. 44–59.

Aytekin, M., C. F. Moss, and J. Z. Simon (2008). “A sensorimotor approach to sound localization”. *Neural Comput* 20.3, pp. 603–35.

Bach-y-Rita, P. (1972). *Brain mechanisms in sensory substitution.* New York: Academic Press.

Bach-y-Rita, P. (2004). “Tactile sensory substitution studies”. *Ann N Y Acad Sci* 1013, pp. 83–91.

Bach-y-Rita, P. and S. W. Kercel (2003). “Sensory substitution and the human-machine interface”. *Trends Cogn Sci* 7.12, pp. 541–6.

Bajcsy, R. (1988). “Active perception”. *Proceedings of the IEEE* 76.8, pp. 966 –1005.

Ballard, D. H. (Feb. 1991). “Animate vision”. *Artificial Intelligence* 48.1, pp. 57–86.

Barsalou, L. W. (1999). “Perceptual symbol systems”. *Behav Brain Sci* 22.4, 577–609; discussion 610–60.

Barsalou, L. W. (2008). “Grounded cognition”. *Annu Rev Psychol* 59, pp. 617–45.

Barsalou, L. W. (2009). “Simulation, situated conceptualization, and prediction”. *Philos Trans R Soc Lond B Biol Sci* 364.1521, pp. 1281–9.

Batista, A. P., C. A. Buneo, L. H. Snyder, and R. A. Andersen (1999). “Reach plans in eye-centered coordinates”. *Science* 285.5425, pp. 257–60.

Beckhaus, S. and **J. Kleesiek** (2011). “Intuitive Interaction: Tapping into body skills to find rich and intuitive interaction methods for Virtual Reality”. *CHI ’11 Workshop 27 – Embodied Interaction: Theory and Practice in HCI.* Vancouver, BC, Canada.

Beer, R. D. (2003). “The Dynamics of Active Categorical Perception in an Evolved Model Agent”. *Adaptive Behavior* 11.4, pp. 209–243.

Beer, R. D. (in press). “Dynamical systems and embedded cognition”. *The Cambridge Handbook of Artificial Intelligence.* Ed. by K. Frankish and W. Ramsey. Cambridge University Press.

Bellman, R. (1957). *Dynamic programming.* Princeton: Princeton University Press.

Berti, A and F Frassinetti (2000). “When far becomes near: remapping of space by tool use”. *J Cogn Neurosci* 12.3, pp. 415–20.

Bjorkman, M. and D. Kragic (2004). "Combination of foveal and peripheral vision for object recognition and pose estimation". *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. Vol. 5, pp. 5135–5140.

Blohm, G. and J. D. Crawford (2009). "Fields of gain in the brain". *Neuron* 64.5, pp. 598–600.

Boden, M. A. (2006). *Mind as machine: a history of cognitive science*. Oxford: Clarendon Press.

Bompas, A. and J. K. O'Regan (2006). "More evidence for sensorimotor adaptation in color perception". *J Vis* 6.2, pp. 145–53.

Bovet, S. and R. Pfeifer (2005). "Emergence of coherent behaviors from homogenous sensorimotor coupling". *Advanced Robotics, 2005. ICAR '05. Proceedings., 12th International Conference on*, pp. 324 –330.

Bradski, G. (2000). "The OpenCV Library". *Dr. Dobb's Journal of Software Tools*.

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

Brooks, R. A. (1991a). "New approaches to robotics". *Science* 253.5025, pp. 1227–32.

Brooks, R. A. (1991b). "Intelligence without representation". *Artificial Intelligence* 47 (1-3), pp. 139–159.

Buneo, C. A., M. R. Jarvis, A. P. Batista, and R. A. Andersen (2002). "Direct visuomotor transformations for reaching". *Nature* 416.6881, pp. 632–6.

Burt, P. (1988). "Smart sensing within a pyramid vision machine". *Proceedings of the IEEE* 76.8, pp. 1006 –1015.

Bushnell, E. W. and J. P. Boudreau (1993). "Motor development and the mind: the potential role of motor abilities as a determinant of aspects of perceptual development". *Child Dev* 64.4, pp. 1005–21.

Butko, N. and J. Movellan (2010). "Learning to look". *IEEE 9th International Conference on Development and Learning (ICDL)*, pp. 70 –75.

Byun, J. and T Nagata (1996). "Active visual sensing of the 3-D pose of a flexible object". *Robotica* 14, pp. 173–188.

Carpenter, W. B. (1852). "On the influence of suggestion in modifying and directing muscular movement, independently of volition". *Proceedings of the Royal Institution of Great Britain* 1, pp. 147–153.

Chakravarthy, V. S., D. Joseph, and R. S. Bapi (2010). "What do the basal ganglia do? A modeling perspective". *Biol Cybern* 103.3, pp. 237–53.

Chalmers, D. J. (1995). "Facing up to the problem of consciousness". *Journal of Consciousness Studies* 2, pp. 200–219.

Chang, C.-C. and C.-J. Lin (2011). "LIBSVM: A library for support vector machines". *ACM Transactions on Intelligent Systems and Technology* 2 (3), 27:1–27:27.

Chang, S. W. C., C. Papadimitriou, and L. H. Snyder (2009). "Using a compound gain field to compute a reach plan". *Neuron* 64.5, pp. 744–55.

Chemero, A. (2009). *Radical embodied cognitive science.* Cambridge, Mass.: MIT Press.

Chen, S., Y. Li, and N. M. Kwok (2011). "Active vision in robotic systems: A survey of recent developments". *The International Journal of Robotics Research.*

Choe, Y., Y. Choe, H.-f. Yang, and N. Misra (2008). "Motor system's role in grounding, receptive field development, and shape recognition". *Proceedings of IEEE 7th International Conference on Development and Learning.*

Cisek, P. and J. F. Kalaska (2005). "Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action". *Neuron* 45.5, pp. 801–14.

Clark, A. (1995). "Moving Minds: Situating Content in the Service of Real-Time Success". *Philosophical Perspectives* 9, pp. 89–104.

Clark, A. (1997a). *Being there: putting brain, body, and world together again.* Cambridge, Mass.: MIT Press.

Clark, A. (1997b). "The Dynamical Challenge". *Cognitive Science* 21.4, pp. 461–481.

Clark, A. (2008). *Supersizing the mind: embodiment, action, and cognitive extension.* Philosophy of mind. Oxford: Oxford University Press.

Clark, A. and D. Chalmers (1998). "The Extended Mind". *Analysis* 58.1, pp. 7–19.

Clark, A. and R. Grush (1999). "Towards a Cognitive Robotics". *Adaptive Behavior* 7.1, pp. 5–16.

Cohen, Y. E. and R. A. Andersen (2002). "A common reference frame for movement plans in the posterior parietal cortex". *Nat Rev Neurosci* 3.7, pp. 553–62.

Collins, S., A. Ruina, R. Tedrake, and M. Wisse (2005). "Efficient bipedal robots based on passive-dynamic walkers". *Science* 307.5712, pp. 1082–5.

Cooke, E. and E. Myin (in press). “Is Trilled Smell Possible? How the Structure of Olfaction Determines the Phenomenology of Smell”. *Journal of Consciousness Studies*.

Crapse, T. B. and M. A. Sommer (2008). “Corollary discharge across the animal kingdom”. *Nat Rev Neurosci* 9.8, pp. 587–600.

Crawford, J. D., W. P. Medendorp, and J. J. Marotta (2004). “Spatial transformations for eye-hand coordination”. *J Neurophysiol* 92.1, pp. 10–9.

Cuijpers, R. H., F. Stuijt, and I. G. Sprinkhuizen-Kuyper (2009). “Generalisation of action sequences in RNNPB networks with mirror properties”. *Proceedings of the 17th European symposium on Artifical Neural Networks (ESANN)*, pp. 251–256.

David, N., A. Newen, and K. Vogeley (2008). “The "sense of agency" and its underlying cognitive and neural mechanisms”. *Conscious Cogn* 17.2, pp. 523–34.

Davidson, P. R. and D. M. Wolpert (2005). “Widespread access to predictive models in the motor system: a short review”. *J Neural Eng* 2.3, S313–9.

Dennett, D. C. (1991). *Consciousness explained*. 1st ed. Boston: Little, Brown and Co.

Desmurget and Grafton (2000). “Forward modeling allows feedback control for fast reaching movements”. *Trends Cogn Sci* 4.11, pp. 423–431.

Dewey, J. (1896). “The reflex arc concept in psychology”. *Psychological Review* 3, pp. 357–370.

Di Paolo, E. A. (2000). “Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions”. *From animals to animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior*, pp. 440–449.

DiGiovanna, J., B. Mahmoudi, J. Fortes, J. C. Principe, and J. C. Sanchez (2009). “Coadaptive brain-machine interface via reinforcement learning”. *IEEE Trans Biomed Eng* 56.1, pp. 54–64.

Dill, M, R Wolf, and M Heisenberg (1993). “Visual pattern recognition in Drosophila involves retinotopic matching”. *Nature* 365.6448, pp. 751–3.

Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. Cambridge, Mass.: MIT Press.

Doya, K (1999). “What are the computations of the cerebellum, the basal ganglia and the cerebral cortex?” *Neural Netw* 12.7-8, pp. 961–974.

Duhamel, J. R., F Bremmer, S BenHamed, and W Graf (1997). “Spatial invariance of visual receptive fields in parietal cortex neurons”. *Nature* 389.6653, pp. 845–8.

Duhamel, J. R., C. L. Colby, and M. E. Goldberg (1998). "Ventral intraparietal area of the macaque: congruent visual and somatic response properties". *J Neurophysiol* 79.1, pp. 126–36.

Dupuy, J.-P. (2009). *On the origins of cognitive science: the mechanization of the mind*. A Bradford book. Cambridge, Mass.: MIT Press.

Ellis, R., M. Tucker, E. Symes, and L. Vainio (2007). "Does selecting one visual object from several require inhibition of the actions associated with nonselected objects?" *J Exp Psychol Hum Percept Perform* 33.3, pp. 670–91.

Engel, A. K. and W Singer (2001). "Temporal binding and the neural correlates of sensory awareness". *Trends Cogn Sci* 5.1, pp. 16–25.

Engel, A. K., P König, A. K. Kreiter, T. B. Schillen, and W Singer (1992). "Temporal coding in the visual cortex: new vistas on integration in the nervous system". *Trends Neurosci* 15.6, pp. 218–26.

Engel, A. K., P Fries, and W Singer (2001). "Dynamic predictions: oscillations and synchrony in top-down processing". *Nat Rev Neurosci* 2.10, pp. 704–16.

Engel, A. K. (2010). "Directive minds: how dynamics shapes cognition". *Enaction: towards a new paradigm for cognitive science*. Ed. by J Stewart, O Gapenne, and E Di Paolo. MIT Press, pp. 219–243.

Faudzi, A. and K. Shibata (2010). "Acquisition of active perception and recognition through Actor-Q learning using a movable camera". *SICE Annual Conference 2010, Proceedings of*, pp. 1950 –1956.

Feldman, J and S Narayanan (May 2004). "Embodied meaning in a neural theory of language". *Brain and Language* 89.2, pp. 385–392.

Fine, P., E. A. Di Paolo, and E. Izquierdo (2007). "Adapting to your body". *Ninth European Conference on Artificial Life*. Springer, pp. 203–212.

Fitzpatrick, P. and G. Metta (2003). "Grounding Vision through Experimental Manipulation". *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 361.1811, pp. 2165–2185.

Fitzpatrick, P., G. Metta, L. Natale, S. Rao, and G. Sandini (2003). "Learning About Objects Through Action – Initial Steps Towards Artificial Cognition". *Proceedings of the 2003 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3140–3145.

Fitzpatrick, P. M. and G. Metta (2002). "Towards Manipulation-Driven Vision". *in IEEE/RSJ Conference on Intelligent Robots and Systems*, pp. 43–48.

Fodor, J. A. (1980). "Methodological Solipsism as a Research Strategy in Cognitive Psychology". *Behavioral and Brain Sciences* 3, pp. 63–109.

Fodor, J. A. (1981). *Representations: philosophical essays on the foundations of cognitive science*. 1st MIT Press ed. Cambridge, Mass.: MIT Press.

Foerster, H. von (1984). *Observing systems*. 2nd ed. Seaside, Calif.: Intersystems Publications.

Foerster, H. von (1988). "On constructing a reality". *Adolesc Psychiatry* 15, pp. 77–95.

Foerster, H. von (2003). *Understanding understanding: essays on cybernetics and cognition*. New York: Springer.

Franceschini, N. (1997). "Combined optical, neuroanatomical, electrophysiological and behavioral studies on signal processing in the fly compound eye". *Biocybernetics of vision: integrative mechanisms and cognitive processes: proceedings of the International School of Biocybernetics, Casamicciola, Napoli, Italy, 16-22 October 1994*. Ed. by C Taddei-Ferretti. Vol. 2. Singapore: World Scientific.

Fries, P. (2005). "A mechanism for cognitive dynamics: neuronal communication through neuronal coherence". *Trends Cogn Sci* 9.10, pp. 474–80.

Frintrop, S. and P. Jensfelt (2008a). "Active gaze control for attentional visual SLAM". *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3690–3697.

Frintrop, S. and P. Jensfelt (2008b). "Attentional Landmarks and Active Gaze Control for Visual SLAM". *IEEE Transactions on Robotics* 24.5, pp. 1054–1065.

Friston, K. (2010). "The free-energy principle: a unified brain theory?" *Nat Rev Neurosci* 11.2, pp. 127–38.

Friston, K. J., J. Daunizeau, and S. J. Kiebel (2009). "Reinforcement learning or active inference?" *PLoS One* 4.7, e6421.

Froese, T. (2009). "Hume and the enactive approach to mind". *Phenomenology and the Cognitive Sciences* 8 (1). 10.1007/s11097-008-9111-5, pp. 95–133.

Froese, T. (2010). "From Cybernetics to Second-Order Cybernetics: A Comparative Analysis of Their Central Ideas". *Constructivist Foundations* 5.2, pp. 75–85.

Gallese, V, L Fadiga, L Fogassi, and G Rizzolatti (1996). "Action recognition in the premotor cortex". *Brain* 119 (Pt 2), pp. 593–609.

Garbarini, F. and M. Adenzato (2004). "At the root of embodied cognition: cognitive science meets neurophysiology". *Brain Cogn* 56.1, pp. 100–6.

Gelder, T. van (1995). "What Might Cognition Be, If Not Computation?" *The Journal of Philosophy* 92.7, pp. 345–381.

Gelder, T. van (1998). “The dynamical hypothesis in cognitive science”. *Behav Brain Sci* 21.5, 615–28; discussion 629–65.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Gibson, J. J. (1977). “The Theory of Affordances”. *Perceiving, acting, and knowing: toward an ecological psychology.* Ed. by Robert Shaw and John Bransford. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Gross, H.-M., V. Stephan, and M. Krabbes (May 1998). “A neural field approach to topological reinforcement learning in continuous action spaces”. *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on.* Vol. 3, 1992 –1997 vol.3.

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

Guillery, R. W. (2005). “Anatomical pathways that link perception and action”. *Prog Brain Res* 149, pp. 235–56.

Harleß, E. (1861). “Der Apparat des Willens”. *Zeitschrift für Philosophie und philosophische Kritik* 38, pp. 50–73.

Harnad, S. (1990). “The symbol grounding problem”. *Physica D: Nonlinear Phenomena* 42.1-3, pp. 335–346.

Haugeland, J. (1985). *Artificial intelligence: the very idea.* Cambridge, Mass.: MIT Press.

Hebb, D. O (1949). *The Organization of Behavior: A Neuropsychological Theory.* A Wiley book in clinical psychology. New York: Wiley.

Heidegger, M. (1975). *Die Grundprobleme der Phänomenologie.* Frankfurt am Main: V. Klostermann.

Heidegger, M. (1977). *Sein und Zeit.* 14. Auflage. Tübingen: Niemeyer.

Held, R and A Hein (1963). “Movement-produced stimulation in the development of visually guided behavior”. *J Comp Physiol Psychol* 56, pp. 872–6.

Held, R., Y. Ostrovsky, B. Degelder, T. Gandhi, S. Ganesh, U. Mathur, and P. Sinha (2011). “The newly sighted fail to match seen with felt”. *Nat Neurosci* 14.5, pp. 551–3.

Helmholtz, H. von (1867). *Handbuch der physiologischen Optik.* Leipzig: Voss.

Helmholtz, H. von and R. Kahl (1971 [1878]). *Selected writings of Hermann von Helmholtz*. Middletown, Conn.: Wesleyan University Press.

Helmholtz, H. von and A. P. König (1896). *Handbuch der physiologischen Optik*. 2. umbearbeite Auflage. Hamburg und Leipzig: L. Voss.

Herbart, J. F. (1816). *Lehrbuch zur Psychologie*. Königsberg, Germany: Unzer.

Hickok, G. (2009). "Eight problems for the mirror neuron theory of action understanding in monkeys and humans". *J Cogn Neurosci* 21.7, pp. 1229–43.

Hoffmann, H. (2007). "Perception through visuomotor anticipation in a mobile robot". *Neural Netw* 20.1, pp. 22–33.

Holmes, N. P., G. A. Calvert, and C. Spence (2004). "Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools". *Neurosci Lett* 372.1-2, pp. 62–7.

Holst, E. von and H. Mittelstaedt (1950). "Das Reafferenzprinzip". *Naturwissenschaften* 37.20, pp. 464–476.

Holyoak, K. J. (1999). "Psychology". *The MIT encyclopedia of the cognitive sciences*. Ed. by Robert A Wilson and Frank C. Keil. Cambridge, Mass.: MIT Press.

Hommel, B, J Musseler, G Aschersleben, and W Prinz (2001). "The Theory of Event Coding (TEC): A framework for perception and action planning". *Behavioral and Brain Sciences* 24.5, pp. 849–878.

Hommel, B. (2009). "Action control according to TEC (theory of event coding)". *Psychol Res* 73.4, pp. 512–26.

Hu, M.-K. (1962). "Visual pattern recognition by moment invariants". *Information Theory, IRE Transactions on* 8.2, pp. 179 –187.

Humphrey, N. (2006). *Seeing red: a study in consciousness*. Cambridge, Mass.: Harvard University Press.

Hurley, S. L (1998). *Consciousness in action*. Cambridge, Mass.: Harvard University Press.

Husserl, E. (1970). *The crisis of European sciences and transcendental phenomenology: an introduction to phenomenological philosophy*. Evanston: Northwestern University Press.

Iacoboni, M. (2009). "Imitation, empathy, and mirror neurons". *Annu Rev Psychol* 60, pp. 653–70.

Izquierdo-Torres, E. and E. A. Di Paolo (2005). “Is an embodied system ever purely reactive?” *Proc. of the 8th European Conference of Artificial Life*. Springer Verlag, pp. 252–261.

James, W. (1950 [1890]). *The Principles of Psychology*. Vol. 2. New York, NY: Dover.

Jäncke, L., M. Cheetham, and T. Baumgartner (2009). “Virtual reality and the role of the prefrontal cortex in adults and children”. *Front Neurosci* 3.1, pp. 52–9.

Jeannerod, M (1995). “Mental imagery in the motor context”. *Neuropsychologia* 33.11, pp. 1419–32.

Jog, M. S., Y Kubota, C. I. Connolly, V Hillegaart, and A. M. Graybiel (1999). “Building neural representations of habits”. *Science* 286.5445, pp. 1745–9.

Johnson-Frey, S. H. (2003). “What’s so special about human tool use?” *Neuron* 39.2, pp. 201–4.

Jordan, M. I. and D. E. Rumelhart (1992). “Forward models: Supervised learning with a distal teacher”. *Cognitive Science* 16, pp. 307–354.

Kawato, M (1999). “Internal models for motor control and trajectory planning”. *Curr Opin Neurobiol* 9.6, pp. 718–27.

Kleesiek, J., A. K. Engel, S. Wermter, and C. Weber (2010). “Object Affordances in the Context of Sensory Motor Contingencies”. *Front. Comput. Neurosci. Conference Abstract: Bernstein Conference on Computational Neuroscience.*

Kleesiek, J., A. K. Engel, C. Weber, and S. Wermter (2011). “Reward-driven learning of sensorimotor laws and visual features”. *IEEE International Conference on Development and Learning (ICDL)*. Vol. 2, pp. 1–6.

Kleesiek, J., S. Badde, S. Wermter, and A. K. Engel (2012). “What do Objects Feel Like? Active Perception for a Humanoid Robot”. *International Conference on Agents and Artificial Intelligence (ICAART)*. Vol. 1, pp. 64–73.

Koffka, K. (1935). *Principles of Gestalt Psychology*. London: K. Paul, Trench, Trubner.

Kohonen, T. (2001). *Self-Organizing Maps*. 3rd. Vol. 30. Springer Series in Information Sciences. Springer, Heidelberg.

Kolen, J. F. and S. C. Kremer (2001). *A field guide to dynamical recurrent networks*. New York: IEEE Press.

Körding, K. P. and D. M. Wolpert (2004). “Bayesian integration in sensorimotor learning”. *Nature* 427.6971, pp. 244–7.

Korner, A. and H. Kraemer (1972). "Individual differences in spontaneous oral behavior in neonates". *3rd symposium on oral sensation and perception.* Ed. by J. Bosma. US Department of Health Education, and Welfare Publication. Bethesda, MD, pp. 335–346.

Kuperstein, M (1988). "Neural model of adaptive hand-eye coordination for single postures". *Science* 239.4845, pp. 1308–11.

Lakoff, G. and M. Johnson (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Lakoff, G. and M. Johnson (1999). *Philosophy in the flesh: the embodied mind and its challenge to Western thought.* New York: Basic Books.

Laycock, T (1845). "On the Reflex Function of the Brain". *British and Foreign Medical Journal* 19, pp. 298–311.

LeCun, Y, L Bottou, G. Orr, and K. Müller (1998). "Efficient BackProp". *Lecture Notes in Computer Science* 1524, pp. 5–50.

Legenstein, R., N. Wilbert, and L. Wiskott (2010). "Reinforcement learning on slow features of high-dimensional input streams". *PLoS Comput Biol* 6.8.

Levine, J (1983). "Materialism and qualia: the explanatory gap". *Pacific Philosophical Quarterly* 64, pp. 354–61.

Lindblom, J. and T. Ziemke (2006). "The social body in motion: cognitive development in infants and androids". *Connection Science* 18.4, pp. 333–346.

Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele.* Leipzig: Weidmann.

Lungarella, M. and O. Sporns (2006). "Mapping information flow in sensorimotor networks". *PLoS Comput Biol* 2.10, e144.

Lungarella, M., T. Pegors, D. Bulwinkle, and O. Sporns (2005). "Methods for quantifying the informational structure of sensory and motor data". *Neuroinformatics* 3.3, pp. 243–62.

Luttrell, S. P. (1994). "A Bayesian Analysis of Self-Organizing Maps". *Neural Computation* 6.5, pp. 767–794.

Macura, Z., A. Cangelosi, R. Ellis, D. Bugmann, M. H. Fischer, and A. Myachykov (2009). "A Cognitive Robotic Model of Grasping". *Proceedings of the Ninth International Conference on Epigenetic Robotics.*

Maillard, M., O. Gapenne, L. Hafemeister, and P. Gaussier (2005). "Perception as a dynamical sensori-motor attraction basin". *Advances in Artificial Life (8th European Conference, ECAL)*, pp. 37–46.

Maravita, A. and A. Iriki (2004). “Tools for the body (schema)”. *Trends Cogn Sci* 8.2, pp. 79–86.

Maravita, A., C. Spence, and J. Driver (2003). “Multisensory integration and the body schema: close to hand and within reach”. *Curr Biol* 13.13, R531–9.

Martín H., J. A., M. Santos, and J. de Lope (2010). “Orthogonal variant moments features in image analysis”. *Inf. Sci.* 180 (6), pp. 846–860.

Maturana, H. R. and F. J. Varela (1992). *The tree of knowledge: the biological roots of human understanding.* Rev. Boston: Shambhala.

Maye, A and A. K. Engel (2011). “A discrete computational model of sensorimotor contingencies for object perception and control of behavior”. *Proceedings of the 2011 IEEE International Conference on Robotics and Automation, ICRA.*

Meltzoff, A. N., A. N. Meltzoff, and M. K. Moore (1997). “Explaining facial imitation: a theoretical model”. *Early Development and Parenting* 6, pp. 179–192.

Merleau-Ponty, M. (1962). *Phenomenology of perception.* London: Routledge.

Merleau-Ponty, M. (1963). *The structure of behavior.* Boston: Beacon Press.

Merleau-Ponty, M. (1974). “Eye and Mind”. *Phenomenology, language and sociology: selected essays of Maurice Merleau-Ponty.* Ed. by John O’Neill. London: Heinemann Educational, pp. 280–311.

Metta, G. and P. Fitzpatrick (2003). “Better Vision through Manipulation”. *Adaptive Behavior* 11.2, pp. 109–128.

Michel, O. (2004). “Webots: Professional Mobile Robot Simulation”. *Journal of Advanced Robotics Systems* 1.1, pp. 39–42.

Millikan, R. G. (2004). *Varieties of meaning.* Cambridge, Mass.: MIT Press.

Minsky, M. (1988). *The Society of Mind.* Pages Bent. Simon & Schuster.

Mishra, A. K., Y. Aloimonos, L. F. Cheong, and A. Kassim (2011). “Active Segmentation”. *IEEE Trans Pattern Anal Mach Intell.*

Mueller, C. G. (1968). “Review of Helmholtz on Perception: Its Physiology and Development.” *Science* 161.3842, p. 679.

Mullette-Gillman, O. A., Y. E. Cohen, and J. M. Groh (2005). “Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus”. *J Neurophysiol* 94.4, pp. 2331–52.

Myin, E. and J. K. O’Regan (2009). “Situated Perception and Sensation in Vision and Other Modalities: A Sensorimotor Approach”. *The Cambridge handbook* *of situated cognition*. Ed. by Philip Robbins and Murat Aydede. Cambridge: Cambridge University Press.

Natale, L., F. Orabona, F. Berton, G. Metta, and G. Sandini (2005). "From sensorimotor development to object perception". *Humanoid Robots, 2005 5th IEEE-RAS International Conference on*, pp. 226 –231.

Newell, A. and H. Simon (1972). *Human Problem Solving*. Prentice Hall, Englewood Cliffs.

Newell, A. and H. A. Simon (1976). "Computer science as empirical inquiry: symbols and search". *Commun. ACM* 19 (3), pp. 113–126.

Nguyen, M. (2006). "Cooperative Coevolutionary Mixture of Experts. A Neuro Ensemble Approach for Automatic Decomposition of Classification Problems". PhD thesis. Australia: University of Canberra.

Nute, D. (Aug. 2011). "A Logical Hole the Chinese Room Avoids". *Minds and Machines* 21.3, pp. 431–433.

Ogata, T., H. Ohba, J. Tani, K. Komatani, and H. G. Okuno (2005). "Extracting multi-modal dynamics of objects using RNNPB". *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Edmonton*, pp. 160–165.

Olsson, L. A., C. L. Nehaniv, and D. Polani (2006). "From unknown sensors and actuators to actions grounded in sensorimotor perceptions". *Connection Science* 18.2, pp. 121–144.

O'Regan, J. K. (1992). "Solving the "real" mysteries of visual perception: the world as an outside memory". *Can J Psychol* 46.3, pp. 461–88.

O'Regan, J. K (2011). *Why red doesn't sound like a bell: understanding the feel of consciousness*. New York: Oxford University Press.

O'Regan, J. K. and A Noë (2001). "A sensorimotor account of vision and visual consciousness". *Behav Brain Sci* 24.5, 939–73; discussion 973–1031.

Peirce, C. S. (1950 [1897]). "Logic as semiotic: The theory of signs". *Philosophical writings of Peirce*. New York: Dover Publications.

Peirce, C. S., N. Houser, and C. J. W. Kloesel (1998). *The essential Peirce: selected philosophical writings*. Bloomington: Indiana University Press.

Pezzulo, G., L. W. Barsalou, A. Cangelosi, M. H. Fischer, K. McRae, and M. J. Spivey (2011). "The mechanics of embodiment: a dialog on embodiment and computational modeling". *Front Psychol* 2, p. 5.

Pfeifer, R. and C. Scheier (1997). "Sensory–motor coordination: The metaphor and beyond". *Robotics and Autonomous Systems* 20.2-4. Practice and Future of Autonomous Agents, pp. 157 –178.

Pfeifer, R., M. Lungarella, and F. Iida (2007). “Self-organization, embodiment, and biologically inspired robotics”. *Science* 318.5853, pp. 1088–93.

Philipona, D, J. K. O’Regan, and J.-P. Nadal (2003). “Is there something out there? Inferring space from sensorimotor dependencies”. *Neural Comput* 15.9, pp. 2029–49.

Philipona, D, J. K. O’Regan, J. P. Nadal, and O Coenen (2004). “Perception of the Structure of the Physical World Using Unknown Multimodal Sensors and Effectors”. *Advances in Neural Information Processing Systems*. MIT Press, pp. 945–952.

Piaget, J. (1952). *The origins of intelligence in children.* New York: International Universities Press.

Pitti, A., H. Mori, S. Kouzuma, and Y. Kuniyoshi (2009). “Contingency Perception and Agency Measure in Visuo-Motor Spiking Neural Networks”. *IEEE Transactions on Autonomous Mental Development* 1.1, pp. 86–97.

Pouget, A., S. Deneve, and J.-R. Duhamel (2002). “A computational perspective on the neural basis of multisensory spatial representations”. *Nat Rev Neurosci* 3.9, pp. 741–7.

Press, J. (2008). “The scientific use of ‘representation’ and ‘function’: avoiding explanatory vacuity”. *Synthese* 161 (1). 10.1007/s11229-006-9154-5, pp. 119–139.

Proffitt, D. R. (2006). “Embodied Perception and the Economy of Action”. *Perspectives on Psychological Science* 1.2, pp. 110–122.

Rasolzadeh, B., M. Björkman, K. Huebner, and D. Kragic (2009). “An Active Vision System for Detecting, Fixating and Manipulating Objects in Real World”. *The International Journal of Robotics Research*.

Riedmiller, M. and H. Braun (1993). “A direct adaptive method for faster backpropagation learning: the RPROP algorithm”. *Neural Networks, 1993., IEEE International Conference on*, 586 –591 vol.1.

Rochat, P (1998). “Self-perception and action in infancy”. *Exp Brain Res* 123.1-2, pp. 102–9.

Rochat, P. and S. J. Hespos (1997). “Differential rooting response by neonates: evidence for an early sense of self”. *Early Development and Parenting* 6.3-4, pp. 105–112.

Roelfsema, P. R. and A. van Ooyen (2005). “Attention-gated reinforcement learning of internal representations for classification”. *Neural Comput* 17.10, pp. 2176–214.

Rohrer, T. (2007). "The body in space: Dimensions of embodiment". *Body, language, and mind.* Ed. by T Ziemke, Jordan Zlatev, and Roslyn M Frank. Vol. 35.1. Berlin: Mouton de Gruyter.

Rolf, M., J. J. Steil, and M. Gienger (2011). "Online Goal Babbling for rapid bootstrapping of inverse models in high dimensions". *Proceedings of the 1st Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics.*

Rosenbaum, D. A., R. A. Carlson, and R. O. Gilmore (2001). "Acquisition of intellectual and perceptual-motor skills". *Annu Rev Psychol* 52, pp. 453–70.

Ross, B., R. Wang, A. Kramer, D. Simons, and J. Crowell (2007). "Action information from classification learning". *Psychonomic Bulletin & Review* 14 (3). 10.3758/BF03194097, pp. 500–504.

Roy, D. (2005a). "Grounding words in perception and action: computational insights". *Trends Cogn Sci* 9.8, pp. 389–96.

Roy, D. (2005b). "Semiotic schemas: A framework for grounding language in action and perception". *Artificial Intelligence* 167.1-2. Connecting Language to the World, pp. 170 –205.

Roy, S., S Chaudhury, and S Banerjee (Jan. 2000). "Isolated 3-d object recognition through next view planning". *Ieee Transactions On Systems Man and Cybernetics Part A-Systems and Humans* 30.1, pp. 67–76.

Roy, S., S Chaudhury, and S Banerjee (Mar. 2004). "Active recognition through next view planning: a survey". *Pattern Recognition* 37.3, pp. 429–446.

Roy, S. D., S. Chaudhury, and S. Banerjee (2005). "Recognizing large isolated 3-D objects through next view planning using inner camera invariants". *IEEE Trans Syst Man Cybern B Cybern* 35.2, pp. 282–92.

Rupert, R. D. (2009). *Cognitive systems and the extended mind.* Philosophy of mind. Oxford: Oxford University Press.

Saeb, S., C. Weber, and J. Triesch (2009). "Goal-directed learning of features and forward models". *Neural Networks* 22.5-6, pp. 586 –592.

Salinas, E and L. F. Abbott (2001). "Coordinate transformations in the visual system: how to generate gain fields and what to compute with them". *Prog Brain Res* 130, pp. 175–90.

Sanchez, J. C., B. Mahmoudi, J. DiGiovanna, and J. C. Principe (2009). "Exploiting co-adaptation for the design of symbiotic neuroprosthetic assistants". *Neural Netw* 22.3, pp. 305–15.

Sanchez-Vives, M. V. and M. Slater (2005). "From presence to consciousness through virtual reality". *Nat Rev Neurosci* 6.4, pp. 332–9.

Scaruffi, P. (2006). *The Nature of Consciousness*. Omniware.

Schult, J. (2008). *Segler Lexikon*. 12., überarb. u. erw. Auflage. Delius Klasing Verlag.

Schultz, W (1998). "Predictive reward signal of dopamine neurons". *J Neurophysiol* 80.1, pp. 1–27.

Schultz, W, P Dayan, and P. R. Montague (1997). "A neural substrate of prediction and reward". *Science* 275.5306, pp. 1593–9.

Searle, J. (1980). "Minds, Brains and Programs". *Behavioral and Brain Sciences* 3.3, pp. 417–425.

Sereno, M. I. and R.-S. Huang (2006). "A human parietal face area contains aligned head-centered visual and tactile maps". *Nat Neurosci* 9.10, pp. 1337–43.

Shaffer, M. J. (May 2009). "A Logical Hole in the Chinese Room". *Minds and Machines* 19.2, pp. 229–235.

Shapiro, L. A. (2011). *Embodied cognition*. New problems of philosophy. New York: Routledge.

Shin, Y. K., R. W. Proctor, and E. J. Capaldi (2010). "A review of contemporary ideomotor theory". *Psychol Bull* 136.6, pp. 943–74.

Shuler, M. G. and M. F. Bear (2006). "Reward timing in the primary visual cortex". *Science* 311.5767, pp. 1606–9.

Singer, W and C. M. Gray (1995). "Visual feature integration and the temporal correlation hypothesis". *Annu Rev Neurosci* 18, pp. 555–86.

Smith, L. B. (2005). "Action Alters Shape Categories". *Cognitive Science* 29.4, pp. 665–679.

Snyder, L. H. (2000). "Coordinate transformations for eye and arm movements in the brain". *Curr Opin Neurobiol* 10.6, pp. 747–54.

Snyder, L. H., A. P. Batista, and R. A. Andersen (1997). "Coding of intention in the posterior parietal cortex". *Nature* 386.6621, pp. 167–70.

Softky, W. and C. Koch (1995). "Single-cell models". *The handbook of brain theory and neural networks*. Ed. by Michael A Arbib. Cambridge, Mass.: MIT Press, pp. 879–884.

Sommer, M. A. and R. H. Wurtz (2006). "Influence of the thalamus on spatial visual processing in frontal cortex". *Nature* 444.7117, pp. 374–7.

Sperry, R. W. (1950). "Neural basis of the spontaneous optokinetic response produced by visual inversion". *J Comp Physiol Psychol* 43.6, pp. 482–9.

Sperry, R. W. (1952). "Neurology and the Mind-Brain Problem". *American Scientist* 40.2.

Spivey, M. (2007). *The continuity of mind.* Vol. 44. Oxford psychology series. Oxford: Oxford University Press.

Stark, M., P. Lies, M. Zillich, J. Wyatt, and B. Schiele (2008). "Functional Object Class Detection Based on Learned Affordance Cues". *6th International Conference on Computer Vision Systems (ICVS).* Santorini, Greece.

Steinman, S. B., B. A. Steinman, and R. P. Garzia (2000). *Foundations of binocular vision: a clinical perspective.* New York: McGraw-Hill.

Stock, A. and C. Stock (2004). "A short history of ideo-motor action". *Psychol Res* 68.2-3, pp. 176–88.

Sutton, R. S. and A. G. Barto (1998). *Reinforcement learning: an introduction.* Cambridge, Mass.: MIT Press.

Suzuki, S and K Be (1985). "Topological structural analysis of digitized binary images by border following". *Computer Vision, Graphics, and Image Processing* 30.1, pp. 32–46.

Tani, J. and M. Ito (2003). "Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment". *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 33.4, pp. 481 –488.

Tani, J., M. Ito, and Y. Sugita (2004). "Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB". *Neural Netw* 17.8-9, pp. 1273–89.

Thagard, P. (2005). *Mind: Introduction to Cognitive Science.* Cambridge, Massachusetts: MIT Press.

Thagard, P. (2010). *The brain and the meaning of life.* Princeton, N.J.: Princeton University Press.

Thelen, E, G Schöner, C Scheier, and L. B. Smith (2001). "The dynamics of embodiment: a field theory of infant perseverative reaching". *Behav Brain Sci* 24.1, 1–34; discussion 34–86.

Thrun, S., W. Burgard, and D. Fox (2005). *Probabilistic robotics.* Cambridge, Mass.: MIT Press.

Torrance, S. (2005). "In search of the enactive: Introduction to special issue on enactive experience". *Phenomenology and the Cognitive Sciences* 4 (4). 10.1007/s11097-005-9004-9, pp. 357–368.

Turing, A. M. (1950). “Computing Machinery and Intelligence”. *Mind* LIX.236, pp. 433–460.

Tye, M. (2009). “Qualia”. *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta.

Uexküll, J. v. and T. v. Uexküll (1980 [1920]). *Kompositionslehre der Natur: Biologie als undogmatische Naturwissenschaft – ausgewählte Schriften.* Berlin: Propyläen-Verlag.

Ugur, E. and E. Sahin (2010). “Traversability: A Case Study for Learning and Perceiving Affordances in Robots”. *Adaptive Behavior* 18.3-4, pp. 258–284.

Valyear, K. F., C. Cavina-Pratesi, A. J. Stiglick, and J. C. Culham (2007). “Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp?” *Neuroimage* 36 Suppl 2, T94–T108.

Varela, F. J., E. Thompson, and E. Rosch (1991). *The embodied mind: cognitive science and human experience.* Cambridge, Mass.: MIT Press.

Walter, W. G. (1953). *The living brain.* English. Duckworth, London.

Ward, R. and R. Ward (2009). “Representation in dynamical agents”. *Neural Netw* 22.3, pp. 258–66.

Warren, R. M. and R. P. Warren (1968). *Helmholtz on Perception: Its Physiology and Development.* New York: Wiley.

Weber, C. and J. Triesch (2009). “Goal-directed feature learning”. *IJCNN’09: Proceedings of the 2009 International Joint Conference on Neural Networks.* Atlanta, Georgia, USA: IEEE Press, pp. 3355–3362.

Weber, C. and S. Wermter (2007). “A self-organizing map of Sigma-Pi units”. *Neurocomputing* 70.13-15, pp. 2552 –2560.

Weiller, D., L. Läer, A. K. Engel, and P. König (2010). “Unsupervised learning of reflexive and action-based affordances to model adaptive navigational behavior”. *Front Neurorobot* 4, p. 2.

Wermter, S., G. Palm, and M. Elshaw (2005). *Biomimetic neural learning for intelligent robots: intelligent systems, cognitive robotics, and neuroscience.* Vol. 3575. Lecture notes in computer science. Berlin: Springer.

Wilson, M. (2002). “Six views of embodied cognition”. *Psychon Bull Rev* 9.4, pp. 625–36.

Wilson, R. A. and A. Clark (2006). “Situated Cognition: Letting Nature Take its Course”. *The Cambridge handbook of situated cognition.* Cambridge University Press.

Wiskott, L and T. Sejnowski (Apr. 2002). "Slow feature analysis: Unsupervised learning of invariances". *Neural Computation* 14.4, pp. 715–770.

Wolpert, D. and Z. Ghahramani (2004). "Computational motor control". *The cognitive neurosciences*. 3rd. Cambridge, Mass.: MIT Press, pp. 485–494.

Woodworth, R. S. (1947). "Reënforcement of Perception". *The American Journal of Psychology* 60.1, pp. 119–124.

Ziemke, T. (2003a). "On the Role of Robot Simulations in Embodied Cognitive Science". *AISB* 1.4, pp. 389–399.

Ziemke, T. (2003b). "What's that thing called embodiment?" *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*.

B Appendix

Publications Originating from this Thesis

Beckhaus, S. and **J. Kleesiek** (2011). “Intuitive Interaction: Tapping into body skills to find rich and intuitive interaction methods for Virtual Reality”. *CHI ’11 Workshop 27 – Embodied Interaction: Theory and Practice in HCI.* Vancouver, BC, Canada.

Kleesiek, **J.**, A. K. Engel, S. Wermter, and C. Weber (2010). “Object Affordances in the Context of Sensory Motor Contingencies”. *Front. Comput. Neurosci. Conference Abstract: Bernstein Conference on Computational Neuroscience.*

Kleesiek, **J.**, A. K. Engel, C. Weber, and S. Wermter (2011). “Reward-driven learning of sensorimotor laws and visual features”. *IEEE International Conference on Development and Learning (ICDL).* Vol. 2, pp. 1–6.

Kleesiek, **J.**, S. Badde, S. Wermter, and A. K. Engel (2012). “What do Objects Feel Like? Active Perception for a Humanoid Robot”. *International Conference on Agents and Artificial Intelligence (ICAART).* Vol. 1, pp. 64–73.

C Appendix
Excerpts *Handbuch der physiologischen Optik*

A scanned version of the book *Handbuch der physiologischen Optik* can be obtained at http://books.google.com. Please note that at the time of publication (1867) no unified grammatical and spelling norms existed. The quotations are taken 'as is', resembling the original typesetting. Hence, for our contemporary understanding the text may contain orthographic errors.

> Eine zweite allgemeine Eigenthümlichkeit unserer Sinneswahrnehmungen ist die, dass wir auf unsere Sinnesempfindungen nur so weit leicht und genau aufmerksam werden, als wir sie für die Erkenntniss äusserer Objecte verwerthen können, dass wir dagegen von allen denjenigen Theilen der Sinnesempfindungen zu abstrahiren gewöhnt sind, welche keine Bedeutung für die äusseren Objecte haben, so dass meistentheils eine besondere Unterstützung und Einübung für die Beobachtung dieser letzteren, subjectiven Empfindungen nothwendig ist.
> (Helmholtz, 1867, p. 430)

> Unsere Anschauungen und Vorstellungen sind Wirkungen, welche die angeschauten und vorgestellten Objecte auf unser Nervensystem und unser Bewusstsein hervorgebracht haben. Jede Wirkung hängt ihrer Natur nach ganz nothwendig ab sowohl von der Natur des Wirkenden, als von der desjenigen, auf welches gewirkt wird. [...]
> Ich meine daher, dass es gar keinen möglichen Sinn haben kann, von einer anderen Wahrheit unserer Vorstellungen zu sprechen, als von einer praktischen. Unsere Vorstellungen von den Dingen können gar nichts anderes sein, als Symbole, natürlich gegebene Zeichen für die

Dinge, welche wir zur Regelung unserer Bewegungen und Handlungen benutzen lernen. Wenn wir jene Symbole richtig zu lesen gelernt haben, so sind wir im Stande, mit ihrer Hilfe unsere Handlungen so einzurichten, dass dieselben den gewünschten Erfolg haben, d. h. dass die erwarteten neuen Sinnesempfindungen eintreten.
(Helmholtz, 1867, 442 *et seq.*)

Von der größten Wichtigkeit endlich für die Festigkeit unserer Überzeugung von der Richtigkeit unserer sinnlichen Wahrnehmung sind die Prüfungen, welche wir mittels der willkürlichen Bewegungen unseres Körpers anstellen. Es entsteht dadurch den blos passiven Beobachtungen gegenüber dieselbe Art festerer Ueberzeugung, welche wir bei wissenschaftlichen Untersuchungen durch das experimentierende Verfahren gewinnen. Der eigentliche letzte Grund, durch welchen alle unsere bewusst vollzogenen Inductionen überzeugende Kraft erhalten, ist das Causalgesetz. Wenn wir sehr häufig zwei Naturerscheinungen verbunden haben auftreten sehen, z. B. den Donner immer dem Blitze folgen, so erscheinen sie gesetzmässig aneinander gebunden, und wir schliessen, dass ein gemeinsamer Grund für beide bestehen muss, und wenn dieser Causalnexus bisher immer bewirkt hatte, dass Donner und Blitz sich begleiteten, so werden gleiche Ursachen auch in Zukunft gleiche Wirkungen hervorbringen müssen, und der Erfolg wird auch in Zukunft derselbe sein müssen. So lange wir nun aber auf blosse Beobachtung solcher Phänomene beschränkt sind, welche ohne unser Zuthun von selbst eintreten, ohne Experimente anstellen zu können, bei denen wir den Complex der Ursachen verändern, gewinnen wir schwer die Überzeugung, dass wir alle Bedingungen, welche auf den Erfolg Einfluss haben können, wirklich schon ermittelt haben. Es muss schon eine ungeheuere Mannigfaltigkeit von Fällen existiren, auf welche das Gesetz passt, und es muss das Gesetz den Erfolg mit grosser Genauigkeit bestimmen, wenn wir uns in einem Falle blosser Beobachtung beruhigen sollen.
(ibid., 450 *et seq.*)

Diesselbe grosse Bedeutung nun, welche das Experiment für die Sicherheit unserer wissenschaftlichen Ueberzeugungen hat, hat es auch für die unbewussten Inductionen unserer sinnlichen Wahrnehmungen. Erst indem wir unsere Sinnesorgane nach eigenem Willen in verschiedene Beziehungen zu den Objecten bringen, lernen wir sicher urtheilen über die Ursachen unserer Sinnesempfindungen, und solches Experimentiren geschieht von frühester Jugend an ohne Unterbrechung das ganze Leben hindurch.
Wenn die Gegenstände nur an unseren Augen vorbeigeführt würden durch fremde Kraft, ohne dass wir selbst etwas dazu thun könnten, würden wir uns in einer solchen optischen Phantasmagorie vielleicht

nie zurecht gefunden haben, [...]
(ibid., p. 452)

Unsere Empfindungen sind eben Wirkungen, welche durch äußere Ursachen in unseren Organen hervorgebracht werden, und wie eine solche Wirkung sich äußert, hängt natürlich ganz wesentlich von der Art des Apparats ab, auf den gewirkt wird. Insofern die Qualität unserer Empfindung uns von der Eigenthümlichkeit der äuſseren Einwirkung, durch welche sie erregt ist, eine Nachricht giebt, kann sie als ein Zeichen derselben gelten, aber nicht als ein Abbild. Denn vom Bilde verlangt man irgend eine Art der Gleichheit mit dem abgebildeten Gegenstande, von einer Statue Gleichheit der Form, von einer Zeichnung Gleichheit der perspectivischen Projection im Gesichtsfelde, von einem Gemälde auch noch Gleichheit der Farben. Ein Zeichen aber braucht gar keine Art der Ähnlichkeit mit dem zu haben, dessen Zeichen es ist. Die Beziehung zwischen beiden beschränkt sich darauf, daſs das gleiche Object, unter gleichen Umständen zur Einwirkung kommend, das gleiche Zeichen hervorruft, und daſs also ungleiche Zeichen immer ungleicher Einwirkung entsprechen.
Der populären Meinung gegenüber, welche auf Treue und Glauben die volle Wahrheit der Bilder annimmt, die uns unsere Sinne von den Dingen liefern, mag dieser Rest von Ähnlichkeit, den wir anerkennen, sehr geringfügig erscheinen. In Wahrheit ist er es nicht; denn damit kann noch eine Sache von der allergröſsten Tragweite geleistet werden, nämlich die Abbildung der Gesetzmäſsigkeit in den Vorgängen der wirklichen Welt. Jedes Naturgesetz sagt aus, daſs auf Vorbedingungen, die in gewisser Beziehung gleich sind, immer Folgen eintreten, die in gewisser anderer Beziehung gleich sind. Da Gleiches in unserer Empfindungswelt durch gleiche Zeichen angezeigt wird, so wird der naturgesetzlichen Folge gleicher Wirkungen auf gleiche Ursachen auch eine ebenso regelmäſsige Folge im Gebiete unserer Empfindungen entsprechen.
Wenn also unsere Sinnesempfindungen in ihrer Qualität auch nur Zeichen sind, deren besondere Art ganz von unserer Organisation abhängt, so sind sie doch nicht als leerer Schein zu verwerfen, sondern sie sind eben Zeichen von Etwas, sei es von etwas Bestehendem oder Geschehendem, und was das Wichtigste ist, das Gesetz dieses Geschehens können sie uns abbilden.
(Helmholtz and König, 1896, p. 586)

So wie diese [die richtige körperliche Vorstellung von dem dargestellten Object] gefunden ist, wandern die beiden Blicklinien mit der gröſsten Sicherheit und Schnelligkeit über alle Theile der Figuren hin. Hier bewährt sich also in der That die Gesammtauffassung der Körperform gleich als die Regel für die Vorstellung, nach welcher man die beiden Blicklinien zu führen hat, um fortdauernd auf correspondirenden

Punkten beider Zeichnungen zu bleiben.
In welcher Weise solche Kenntnisse der Bedeutung der Gesichtsbilder von jungen menschlichen Kindern zuerst gesammelt werden, ergiebt sich leicht, wenn wir dieselben beobachten, während sie mit den ihnen als Spielzeug dargebotenen Objecten sich beschäftigen, wie sie dieselben betasten, stundenlang von allen Seiten betrachten, herumwenden, sie in den Mund stecken u. s. w., endlich sie herunterwerfen oder zu zerschlagen suchen und dies jeden Tag wiederholen. Man wird nicht daran zweifeln können, daſs dies die Schule ist, in der sie das natürliche Verhalten der sie umgebenden Gegenstände kennen lernen, dabei auch die perspectivischen Bilder verstehen, ihre Hände gebrauchen lernen. Ebenso lehrt die Beobachtung jüngerer Kinder, daſs sie in den ersten Wochen ihres Lebens diese Kenntnisse noch nicht haben. Wenn ihnen irgend eine instinktmäſsige Kenntniſs angeboren wäre, so sollte man erwarten, daſs es in erster Linie die Kenntniſs des Bildes der Mutterbrust sein müſste und die Kenntniſs derjenigen Bewegung, durch welche sie sich diesem Gesichtsbilde zuwenden könnten. Aber eine solche Kenntniſs fehlt ganz offenbar. Man sieht, daſs das Kind lebhaft wird, wenn es in die Stellung für das Säugen gebracht wird, und unruhig suchend den Kopf hin und her wendet, um die Brust zu finden, aber es wendet sich in den ersten Tagen ebens oft von der Brust ab, wie ihr zu, obgleich es diese frei erblicken kann. Offenbar weiſs es in diesem frühen Alter weder das Gesichtsbild, noch die Richtung seiner Bewegungen zu deuten. [. . .]
Ich folgere daraus, daſs die Deutung auch einiger der einfachsten und für das menschliebe Kind wichtigsten Gesichtsbilder von ihm erlernt werden muß und nicht durch angeborene Organisation von vorneherein ohne vorausgehende Erfahrung gegeben ist.[. . .]
In diesem Sinne können wir behaupten, die Vorstellung der stereometrischen Form eines körperlichen Objects spielt ganz die Rolle eines aus einer groſsen Reihe sinnlicher Anschauungsbilder zusammengefaſsten Begriffs, der aber selbst nicht nothwendig durch in Worten ausdrückbare Definitionen, wie sie der Geometer sich construiren könnte, sondern nur durch die lebendige Vorstellug des Gesetzes, nach dem seine perspektivischen Bilder einander folgen, zusammengehalten wird.
Daſs eine solche mühelose Anschauung der normalen Folge von gesetzlich verküpften Wahrnehmungen durch hinreichend reiche Erfahrung gewonnen werden kann, habe ich zu beweisen gesucht.
(Helmholtz and König, 1896, 599 *et seq.*)

D Appendix

Acknowledgements

I thank my supervisors, Prof. Dr. Andreas Engel and Prof. Dr. Stefan Wermter, for their advice and generous support. Thank you for letting me pursue my passions.

I thank Dr. Cornelius Weber for valuable discussions and for sharing his deep knowledge and understanding of artificial neural systems with me.

I am grateful to Stephanie Badde for fun and productive afternoons of interdisciplinary work.

I am grateful to Prof. Dr. Steffi Beckhaus for inspiring discussions (not only) during lunch time.

I thank Dieter Jessen and Matthias Siemers for their prompt and skilled technical support.

I am grateful to the members of the Knowledge Technology Group at the Department of Informatics, as well as to the members of the Department of Neurophysiology and Pathophysiology at the University Hospital Hamburg-Eppendorf.

Finally, I would like to thank the Sino-German Research Training Group CINACS, DFG GRK 1247/1 and 1247/2, for their financial support.

E

Appendix

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den

____

Unterschrift