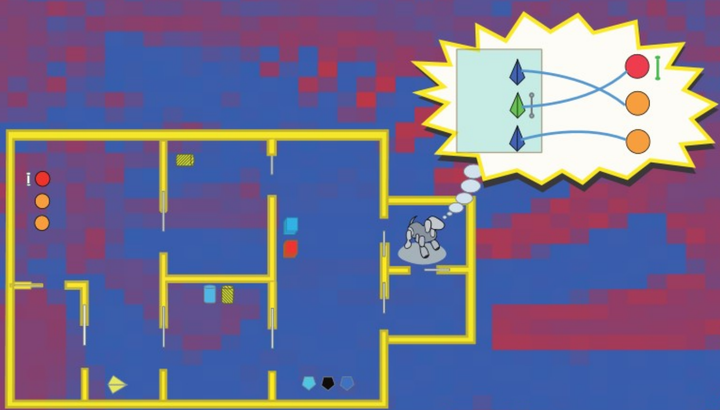
State-of-the-Art Survey

LNAI 5225

Giovanni Pezzulo
Martin V. Butz
Cristiano Castelfranchi
Rino Falcone (Eds.)

The Challenge of Anticipation

A Unifying Framework for the Analysis and Design of Artificial Cognitive Systems



Springer

Lecture Notes in Artificial Intelligence 5225

Edited by R. Goebel, J. Siekmann, and W. Wahlster

Subseries of Lecture Notes in Computer Science

Giovanni Pezzulo Martin V. Butz
Cristiano Castelfranchi Rino Falcone (Eds.)

The Challenge of Anticipation

A Unifying Framework for the Analysis and Design of Artificial Cognitive Systems

Springer

Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada
Jörg Siekmann, University of Saarland, Saarbrücken, Germany
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

Volume Editors

Giovanni Pezzulo
Cristiano Castelfranchi
Rino Falcone
Istituto di Scienze e Tecnologie della Cognizione - CNR
Via San Martino della Battaglia 44
00185 Rome, Italy
E-mail: {giovanni.pezzulo, cristiano.castelfranchi, rino.falcone}@istc.cnr.it

Martin V. Butz
Universität Würzburg
Institut für Psychologie, Kognitive Psychologie III
Röntgenring 11, 97070 Würzburg, Germany
E-mail: butz@psychologie.uni-wuerzburg.de

Library of Congress Control Number: 2008934765

CR Subject Classification (1998): I.2, I.6, F.1.2, G.3, K.8

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN	0302-9743
ISBN-10	3-540-87701-0 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-87701-1 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2008
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Markus Richter, Heidelberg
Printed on acid-free paper SPIN: 12531196 06/3180 5 4 3 2 1 0

Foreword

The general idea that brains anticipate the future, that they engage in prediction, and that one means of doing this is through some sort of inner model that can be run offline, has a long history. Some version of the idea was common to Aristotle, as well as to many medieval scholastics, to Leibniz and Hume, and in more recent times, to Kenneth Craik and Philip Johnson-Laird. One reason that this general idea recurs continually is that this is the kind of picture that introspection paints. When we are engaged in tasks it seems that we form images that are predictions, or anticipations, and that these images are isomorphic to what they represent.

But as much as the general idea recurs, opposition to it also recurs. The idea has never been widely accepted, or uncontroversial among psychologists, cognitive scientists and neuroscientists. The main reason has been that science cannot be satisfied with metaphors and introspection. In order to gain acceptance, an idea needs to be formulated clearly enough so that it can be used to construct testable hypotheses whose results will clearly support or cast doubt upon the hypothesis. Next, those ideas that are formulable in one or another sort of symbolism or notation are capable of being modeled, and modeling is a huge part of cognitive neuroscience. If an idea cannot be clearly modeled, then there are limits to how widely it can be tested and accepted by a cognitive neuroscience community. And finally, ideally, the idea will be articulated and modeled in such a way that it is not a complete mystery how it could be implemented by the brain. Though the idea that the brain models and predicts and anticipates is supported by introspection and a long history of hypotheses, it has largely failed on these latter three counts – especially compared with various theoretical competitors. And this is why the extent to which it has been embraced by cognitive science and neuroscience has been limited.

But there is good news. Mathematical tools from a number of areas, including modern control theory and signal processing, are capable of allowing for very precise mathematical formulations of the basic idea, as well as many specific versions. This allows for the ideas not only to be precisely formulated, but also to be modeled and compared to human behavioral data. And given a number of schemes for implementing these kinds of mathematical models in neural systems, it is possible to see these models as being implemented in the brain. The qualitative idea that the brain models the world is finally being clarified and quantified.

But we are still in the early stages of this process. While there are many proposals and theories that are beginning to take shape, there have been few sustained treatments of the topic that attempt to develop them in detailed and consistent ways. Rather, the applications have largely been piecemeal. In this regard the present volume represents a significant advance in the field. It offers a sustained treatment of various aspects of the general hypothesis, not only in terms of being conceptually clear and consistent, but also in terms of presenting a wide range of particular applications that illustrate the conceptual machinery in action.

It would be an overstatement to say that the idea that the brain is a modeler and predictor is revolutionary, or that the current swell in theoretical interest in the idea represents the initial stages of a revolution in cognitive neuroscience. But while talk of revolution may be overstatement, it cannot be denied that this approach to understanding brain function is beginning to take on an importance comparable to that of traditional artificial intelligence approaches and connectionist modeling approaches. The clarity, detail and quality of the ideas presented in this volume, coupled with the growing importance of this general approach, make this volume a critical contribution to our understanding of brain function, and should be read by anyone with a serious interest in understanding how the brain manages to support cognitive functions.

July 2008

Rick Grush
University of California, San Diego

Preface

Prediction is difficult – especially for the future. *Niels Bohr*

Over the last few decades, it has become increasingly clear that animals most of the time do not simply react in their world based on unconditioned or conditioned stimuli, but rather actively operate in their environment in a highly goal- and future-oriented way, and not just on the basis of current perception, but in part autonomously from environmental stimuli. Psychology now suggests that it is the goal itself that triggers behavior and attention. Learning is highly influenced by current predictive knowledge and the consequent detection of novelty. Behavioral control is most effectively controlled by the help of forward models that substitute delayed or that enhance noisy perceptual feedback. Thus, anticipations come in many forms and influence many cognitive mechanisms.

This book proposes a unifying approach for the analysis and design of artificial cognitive systems: **The anticipatory approach**. We propose a foundational view of the importance of dealing with the future, of gaining some autonomy from current environmental data, of endogenously generating sensorimotor and abstract representations. We propose a meaningful taxonomy of anticipatory cognitive mechanisms, distinguishing between the types of predictions and the different influences of these predictions on actual behavior and learning. Doing so, we sketch out a new, unifying perspective on cognitive systems. Mechanisms, that have often been analyzed in isolation or have been considered unrelated to each other, now fit into a coherent whole and can be analyzed in correlation to each other. Learning and behavior are considered increasingly intertwined and correlated with each other. Attention and action control suddenly appear as very similar processes. Goal-oriented behavior, motivation and emotion appear as related and intertwined.

While the revelation of these correlations is helpful for the analysis and comparison of different learning and behavioral mechanisms, the second benefit of the anticipatory approach is the possibility to modularly design novel cognitive system architectures. The developed taxonomy clearly characterizes which aspects are important for different anticipatory cognitive modules and how these modules may interact with each other. Thus, the second benefit of the anticipatory approach is the facilitation of cognitive system design. Building blocks of cognitive systems are proposed and exemplarily analyzed in diverse system architectures. The interaction of these building blocks then is characterized by their anticipatory nature, facilitating the design of larger, more competent autonomous artificial cognitive systems. We hope that the proposed anticipatory approach may thus not only serve for the analysis of cognitive systems but rather also as an inspiration and guideline for the progressively more advanced and competent design of large, but modular, artificial cognitive systems.

Acknowledgments

This work is supported by the EU project **MindRACES**, from Reactive to Anticipatory Cognitive Embodied Systems, funded under grant FP6-511931 under the "Cognitive Systems" initiative from the EC. Special thanks to our Project Officer, Cécile Huet, and to our Project Reviewers, Lola Cañamero and Deepak Kumar, for their valuable encouragement and advice.

July 2008

Giovanni Pezzulo
Martin V. Butz
Cristiano Castelfranchi
Rino Falcone

Contents

List of Contributors

Gianluca Baldassarre
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: gianluca.baldassarre@istc.cnr.it

Christian Balkenius
Lund University Cognitive Science - Kungshuset, Lundagård SE-222 22 LUND, Sweden, e-mail: christian.balkenius@lucs.lu.se

Martin V. Butz
Department of Psychology, Cognitive Psychology III, University of Würzburg, Röntgenring 11, 97070 Würzburg, Germany, e-mail: butz@psychologie.uni-wuerzburg.de

Cristiano Castelfranchi
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: cristiano.castelfranchi@istc.cnr.it

Rino Falcone
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: rino.falcone@istc.cnr.it

Alexander Förster
Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA) Galleria 2, 6928 Manno-Lugano - Switzerland, e-mail: alexander@idsia.ch

Maurice Grinberg
Central and East European Center for Cognitive Science New Bulgarian University 21 Montevideo Blvd. Sofia, 1618 Bulgaria, e-mail: mgrinberg@nbu.bg

Oliver Herbort
University of Wurzburg, Rontgenring 11, 97070 Wurzburg, Germany, e-mail: oliver.herbort@psychologie.uni-wuerzburg.de

Kiril Kiryazov
Central and East European Center for Cognitive Science, New Bulgarian University, 21 Montevideo Street, 1618 Sofia, Bulgaria, e-mail: kiryazov@cogs.nbu.bg

Boicho Kokinov
Central and East European Center for Cognitive Science, New Bulgarian University, 21 Montevideo Street, 1618 Sofia, Bulgaria, e-mail: bkokinov@nbu.bg

Birger Johansson
Lund University Cognitive Science - Kungshuset, Lundagård SE-222 22 LUND, Sweden, e-mail: birger.johansson@lucs.lu.se

Emilian Lalev
Central and East European Center for Cognitive Science New Bulgarian University 21 Montevideo Blvd. Sofia, 1618 Bulgaria , e-mail: elalev@cogs.nbu.bg

Emiliano Lorini
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: emiliano.lorini@istc.cnr.it

Carlos Martinho
Instituto Superior Tecnico (IST) Lisboa, Portugal,
e-mail: carlos.martinho@tagus.ist.utl.pt

Maria Miceli
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: maria.miceli@istc.cnr.it

Dimitri Ognibene
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: dimitri.ognibene@istc.cnr.it
DIST, Via all'Opera Pia 13 - 16145 Genova, Italy

Ana Paiva
Instituto Superior Tecnico (IST) Lisboa, Portugal, e-mail: ana.paiva@inesc-id.pt

Georgi Petkov
Central and East European Center for Cognitive Science, New Bulgarian University, 21 Montevideo Street, 1618 Sofia, Bulgaria, e-mail: gpetkov@mail.nbu.bg

Giovanni Pezzulo
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: giovanni.pezzulo@istc.cnr.it

Michele Piunti
Istituto di Scienze e Tecnologie della Cognizione-CNR, Via S. Martino della Battaglia 44 - 00185 Roma, Italy, e-mail: michele.piunti@istc.cnr.it

Vin Thorsteinsdottir
Lund University Cognitive Science - Kungshuset, Lundagård SE-222 22 LUND, Sweden, e-mail: vin.thorsteinsdottir@lucs.lu.se

Part I
Theory

Chapter 1
Introduction: Anticipation in Natural and Artificial Cognition

Giovanni Pezzulo, Martin V. Butz, Cristiano Castelfranchi, and Rino Falcone

The purpose of brains is to produce future. *Paul Valery*

1.1 Introduction

What will artificial cognitive systems of the future look like? If we are asked to imagine robots, or intelligent software agents, several features come to our mind such as the capability to adapt to their environments and to satisfy their goals with only limited human intervention, to plan sequences of actions for realizing long-term objectives, to act collectively in view of complex objectives, to interact and cooperate with us, with and without natural language, to take decisions (also in our place), etc.

Currently these capabilities are far beyond the possibilities of robots and other artificial systems. In the next years a huge effort will be required for scaling up the potentialities of the artificial systems that we are able to build nowadays. One way to overcome these limitations is to take inspiration form the functioning of living organisms. A large body of evidence, which we review in this chapter, indicates that natural cognitive systems are not reactive but essentially anticipatory systems. We do not think that this is a mere coincidence. On the contrary, we claim that anticipation is a crucial—and foundational—phenomenon in natural cognition. Individual behavior is guided by anticipatory mechanisms that are used for behavioral control, perceptual processing, goal-directed behavior, and learning. And also effective social behavior relies on the anticipation of the behavior of other agents. We argue that anticipation is a key ingredient for the design of autonomous, artificial cognitive agents of the future: *Only cognitive systems with anticipation mechanisms can be credible, adaptive, and successful in interaction with both the environment and other autonomous systems and humans*. This is the challenge that we anticipate for the future of cognitive systems research: the passage from reactive to anticipatory cognitive embodied systems.

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 3–22, 2008.
© Springer-Verlag Berlin Heidelberg 2008

1.1.0.1 From Reactive to Anticipatory Cognitive Embodied Systems

Overall, we propose an integrated approach to the study of anticipation that encompass empirical, theoretical, and computational approaches. The study of anticipation has a long history in the empirical literature, that we will review extensively in the next section. In Chapter 2, we focus on the investigation of anticipatory functionalities from a conceptual point of view as well as from a computational one. Despite some studies available on this topic, we believe that a unitary approach to the study of anticipation is still missing. This book intends to address exactly this challenge.

1.1.0.2 Overview of This Chapter

The first part of this chapter is conceptual in nature. It consists in finding 'conceptual keys' to understand the phenomena of cognition and behavior, and to use them to inspire our design methodology. For this reason, in Section 1.2 we propose theoretical arguments for assessing the relevance of anticipation and anticipatory behavior in cognition. We start from a discussion of the role anticipation has played in cognitive science and artificial systems research. Particularly, we highlight why living organisms endowed with anticipatory capabilities are able to develop cognitive capabilities—from simpler to more complex ones—including those based on representations. In conclusion, we propose that *cognitive minds inevitably have to be anticipatory devices*. We also argue that studying anticipatory mechanisms in the brain in isolation does not suffice. Rather, it is also necessary to understand the nature of future-oriented behavior, and of future-oriented representations: which are their specific advantages, why do cognitive agents need to anticipate, etc.

Since this book aims to offer design principles for endowing artificial systems with anticipatory capabilities, it is essential to analyze in detail the specific roles of anticipation in several cognitive functions, including sensorimotor control, attention, internal preparation to action, emotional regulation, learning, exploration, curiosity, and decision making. For this purpose, in the second part of this chapter we focus on how anticipatory mechanisms actually work in living organisms: Section 1.3 gives a review of psychological and neuroscientific theories and models of anticipation. This systematic exploration of natural anticipatory systems is meant to be a source of inspiration for the sake of designing artificial anticipatory systems that have the same levels of adaptivity, flexibility, and autonomy.

1.2 The Path to Anticipatory Cognitive Systems

Before we delve into the different aspects of predictions and anticipatory capabilities, we first sketch out the research path that lead to explicit studies of anticipatory cognitive mechanisms such as the formulation of the theory of *anticipatory behavioral control* (Hoffmann, 1993; Hoffmann et al., 2004), the study of *anticipatory behavior* (Butz et al., 2003b, 2007b), the proposition of *the mind as an anticipatory device* (Castelfranchi, 2005; Pezzulo and Castelfranchi, 2007), and ultimately the authoring of this book.

1.2.1 Symbolic Behavior, Representation-Less Behavior, and Their Merge to Anticipatory Behavior

Traditionally, artificial intelligence (AI) investigated the functionality of symbol-oriented cognitive mechanisms, such as search, planning, and decision making in well-defined, discrete problems, principles of first-order logic, of learning based on symbolic inputs (Russell and Norvig, 1995). The concept of representation as traditionally defined and used in AI was rather detached from actually available sensory information, yielding impressive performance in well-defined environments, such as the game of chess, but highly unsatisfactory performance in natural environments, such as for adaptive robot control. In the latter case, the traditional AI-based systems suffered from several fundamental problems: (1) The scalability problem restricted the systems to solve only highly simplified and very small toy problems. (2) The symbol grounding problem (Harnad, 1990) prevented them from identifying effective sensory discretizations, so that effective symbolic representations, which may be suitable for planning or decision making, did not emerge. (3) The frame problem (McCarthy and Hayes, 1969) prevented systems from effectively representing action-affected and -unaffected parts of the environment with logic-based representations.

As a consequence, the situated approach to cognition gained popularity (Brooks, 1991; Chiel and Beer, 1997; Pfeifer and Scheier, 1999). The situated approach challenges several weak points of traditional AI methodology such as that all ‘cognitive’ functions (including perception, categorization, etc.) are assumed to be based on (the manipulation of) internal representations and that reasoning is overemphasized in comparison with situated motor activity. The subsequent situated AI approach had a great impact on our current understanding of cognition, learning, and adaptive behavior. For example, it was shown that the effective coupling of brain-body-environment dynamics, without representational processes, can yield very efficient, representation-less behavioral patterns (Braitenberg, 1984; Brooks, 1991). Many of the consequentially realized cognitive functions were previously considered representational in nature. The consequence of these behavioral successes lead to a critical reconsideration of the role of representations and internal processes.

A fundamental side effect of this approach is the emphasis on *reactive mechanisms*, which mainly originates from the necessity to avoid the grounding problem and at the same time to de-emphasize the role of internal representations. This fact has produced a lack of interest for representational processes, which are however widespread in natural cognition, and resulted in skepticism with respect to the need of internal models and representations.

Although reactive and non-representational systems have shown a range of capabilities that were unsuspected, it became also increasingly clear that they will never reach the levels of complexity observable in natural cognitive systems. On the contrary, the important role for anticipations and anticipatory mechanisms in natural cognition is highlighted in several empirical studies. Nature seems to have found a suitable way to overcome the shortcomings of reactive systems by endowing living organisms with anticipatory mechanisms.

1.2.2 The Power of Anticipation: From Reactivity to Proactivity

Reactive systems are those that produce behavior as a response or reaction to (sensed) environmental conditions and internal needs. They do not need to have a complex representation of their environment since it is sufficient for them to sense it. Take as an example a reactive driving rule: *If you see the car in front of you stopping (e.g., you see the red lights indicating the stop), then press the brake*. In normal traffic conditions, a reactive system endowed with this rule is able to avoid accidents most of the time. In artificial systems research, reactive rules (independently of how they are implemented) lead to efficient systems, since the computation they have to carry out is simple and cheap. However, these systems are not versatile, since they tend to have stereotyped responses, and they are not able to prepare for future conditions, but rather they have to wait for the conditions to occur first.

On the contrary, a system endowed with predictive capabilities can use the following rule: *If the car in front of you is close to a crossing, then it is likely to stop, so stop in advance or at least get ready to stop when a crossing is ahead*. Thus, a system endowed with predictive capabilities can take into account (possible) future events to decide on and prepare current behavior.

Predictive capabilities permit even much more subtle behavior. Anticipatory systems can, for example, select an action whose anticipated effect is judged to be positive, prevent dangers before experiencing them, actively search for information that is expected to be relevant, etc. All these capabilities, that are based on processing information relative to the future, are the keys for passing from mere *reactivity* to *proactivity* and *goal-oriented behavior*.

1.2.3 The Anticipatory Approach to Cognitive Systems

By proposing the *anticipatory approach* to cognitive systems, we argue that—now that the criticism of traditional AI from the situated approach is quite well accepted—it is time to reconsider representations. However, these representations need to be integrated into the situated approach to AI. That is, representations may emerge out of representation-less systems and may suitably alter the capabilities of the situated systems. Thus, representations may no longer be detached symbols, but they will need to be grounded in the body's sensory and motor systems and situated in the perceived environment. Our strong belief is that such modern, behaviorally-suitable representations can emerge from the anticipatory approach to cognition. We therefore propose to focus on how anticipations are realized in living organisms, and to investigate how anticipatory representations permit the realization of cognitive functions.

The Mind Is an Anticipatory Device One central tenet of our anticipatory approach to cognition is that a true cognitive mind serves for (and has evolved for) anticipation: *The mind is an anticipatory device* (Castelfranchi, 2005; Pezzulo and Castelfranchi, 2007). Anticipation is not only required for several cognitive functions, but it is an 'ordering principle' of cognition and its development. For this reason, the study of anticipatory phenomena can shed light onto natural cognition—a view that is currently also gaining consensus in the neuroscientific community (Bar, 2007; Frith, 2007; Hawkins and Blakeslee, 2004). For example, after reviewing a number of memory studies and theoretical analysis, Schacter et al. (2007, pg. 660) concludes that:

> Given the adaptive priority of future planning, we find it helpful to think of the brain as a fundamentally prospective organ that is designed to use information from the past and the present to generate predictions about the future.

There are two main reasons for conceiving cognitive minds as essentially anticipatory and future-oriented. First, cognition should be described as an active and productive activity rather than a passive stimuli-processing system. Second, representational and symbolic capabilities were only able to develop due to adaptive advantages of anticipating and dealing with the future. For this reason, the capability to form grounded representations and symbols depends on the capability to anticipate. We illustrate these points in further detail in the two following sections.

1.2.3.1 The Productive View of Cognition

The *productive view of cognition* that we put forward in our anticipatory approach originates from Kant's (1998) idea that, although our knowledge begins with experience, it does not purely arise from experience, since our productive, generative apparatus determines what we know. We do not passively process environmental stimuli, but actively produce representations by means of our categorical apparatus.

> That all our knowledge begins with experience there can be no doubt. [...] But, though all our knowledge begins with experience, it by no means follows that all arises out of experience. For, on the contrary, it is quite possible that our empirical knowledge is a combination of that which we receive through impressions, and [additional knowledge] altogether independent of experience [...] which the faculty of cognition supplies from itself, sensory impressions giving merely the occasion. (Kant, 1998, Introduction)

This idea has been very important in the theory of Piaget (1954), that introduces as an important element of novelty an emphasis on the situated and action-based origin and nature of representations.

> Any piece of knowledge is connected with an action . . . [T]o know an object or a happening is to make use of it by assimilation into an action schema . . . [namely] whatever there is in common between various repetitions or superpositions of the same action. (Piaget, 1971, pg. 6-7)

Central in the Piagetian theory is the concept of schemas (or, better, sensorimotor schemas), which is a highly recognized concept in cognitive science (Arbib, 1992, 2003; Bartlett, 1932; Neisser, 1976). He describes cognitive development in humans as a process of *assimilation* and *accommodation*, in which schemas of increasing complexity are formed to make sense of the world and to operate on it. A central posit, which is generally adopted in schema-based computational modeling (Arbib, 1992; Drescher, 1991; Pezzulo and Calvi, 2007b; Roy, 2005), is that acting on the basis of an action schema also entails the expectation of action effects. Expectations are crucial in the control of action and categorization—two sides of the same coin—since the compliance of action implicitly verifies the expectations and permits the categorization of an object or an event. Expectations are also used for assimilation, that is, learning of a novel schema when current expectations are not met (cf. Pezzulo and Calvi, 2007b for a further discussion on the relations between the pragmatic and epistemic sides of action schemas).

The productive aspect of cognition is particularly important in cognitive, goal-oriented agents. Contrary to the view that cognitive agents can be represented as input-output devices that passively receive inputs for reacting appropriately, we argue that expectations precede stimuli, both factually and conceptually. Factually, anticipatory representations are already there before inputs are received. Conceptually, a true goal-oriented behavior begins with a goal and not with a stimulus.

Recently a great deal of evidence has been accumulated that strengthens this view. Bar (2007) proposes that the mind, thanks to associative mechanisms, is proactive and continuously generates predictions approximating the relevant future—a position that is consistent with the idea of the mind as an anticipatory device. Also in accordance with the view put forward in this chapter, he suggests therefore that anticipation is one (of the few) unifying principle(s) of brain functioning.

With a slightly different emphasis (on memory studies rather than proactivity and goal orientedness), several authors have recently proposed that the essential function of memory is not 'storage' but enabling dealing with the future. Some examples are 'mental time travel' (Tulving, 1983), 'memory of the future' (Ingvar, 1985), 'memory for the future' (Glenberg, 1997), and the 'prospective brain' (Schacter et al., 2007). All these studies highlight complementary aspects of the productive view of cognition and indicate anticipation as a basic, unitary capability of cognition that produces cognitive and behavioral effects.

1.2.3.2 The Mind Originates from the Need to Deal with the Future

We have reviewed evidence indicating that the organization of the perceptual and motor apparatus is biased toward the future. However, here we do one more step and suggest that the cognitive mind's main function is to anticipate and to deal with the future. In describing animals as machines evolved for the survival and propagation of their genes, Dawkins (1989, pg. 59) argues that

> Survival machines that can simulate the future are one jump ahead of survival machines who can only learn on the basis of overt trial and error.

We have already discussed how anticipation permits the channelization of epistemic and pragmatic activity. But there is another, perhaps more fundamental reason for considering the role of anticipation as essential in the development of cognitive minds: internal representations might have emerged from anticipatory mechanisms such as internal models for the sake of dealing with the future.

The Emergence of Representations from Anticipation What distinguishes a cognitive from a merely adapted system is the capability to form internal representations (and in particular expectations) and to work on them internally before, or instead of, operating directly on the environment. Endogenously producing representations, and working on them internally instead of immediately acting (Piaget's *substitution*) is a hallmark of cognition.

However, building an internal model before acting is costly. The knowledge-based approach in AI has received criticisms by the situated approach, exemplified by Brooks's (1991) idea that 'the world is the best representation of itself', and in theory, due to the frame problem that mines the idea that we can formulate a symbolic representation of the world and possible actions before acting –and in fact Dennett's (1984) analogy with the *Buridan's ass* illustrates the problems of model based approaches.

One crucial challenge is then to understand how representations can arise in situated systems and how the cost of representation is balanced by appropriate gains. It is extremely advantageous for a situated system to be able to deal with the future and not only the present. An anticipation of the future implies gains in specific cognitive functions such as attention and motor control as well as the development of completely new capabilities such as planning (we review extensively the benefits of anticipation in Chapter 3).

Thus, representations might have originated thanks to the need for dealing with the future, for which the direct way is to predict and represent the future. In turn, representing the future and detaching one's own representations from the current sensorimotor interaction might have provided several other advantages. Representing the future, including future events, our own actions in the future as well as other's actions, is essential for coordinating one's own acts for a long time span, for coordinating with others, for realizing future states that are desirable (this includes controlling others), or for avoiding dangerous futures.

Another related (and not concurrent) hypothesis is put forward by König and Krüger (2006) who argue that discrete entities (symbols) emerge in the brain in the process of feature extraction as a byproduct of data compression for the sake of permitting better predictions:

> In the process of mutual optimization of features and predictions, symbols emerge as condensed entities on which predictions are performed. (König and Krüger, 2006, pg. 14)

In a similar vein, several researchers stress the role of anticipation in the development of cognitive capabilities, from the simplest sensorimotor ones to more complex and symbolic ones such as language (Clark and Grush, 1999; Gardenfors and Orvath, 2005; Grush, 2004).

Several studies in cognitive robotics are now beginning to investigate anticipation from a situated perspective. Although they mainly focus on the control of action and basilar social abilities so far, this direction seems very promising to scale up to higher level cognitive and social functions. Understanding the nature and functioning of anticipatory behavior in unitary perspective will be, however, a big theoretical challenge since anticipatory representations have a double-sided nature, being both *grounded* and *detached*. According to the situated approach to cognition, representations have to be *grounded* (Harnad, 1990): An agent can only act adaptively if it stays intimately coupled with its environment. At the same time, representations and especially expectations and distal goals are by definition about future states of affairs (sometimes even impossible ones), and they are therefore *detached* from the current sensorimotor engagement. Thus, groundedness and detachment seem to be at odds: how can both be obtained?

The answer can come from an investigation, in a developmental perspective, of the *detachment process* that permits to develop anticipatory representations starting only from sensorimotor engagement: representations are not *born* detached, but *become* detached. The reviewed literature offers some indications about possible stages of the detachment process. Accordingly, several authors (Clark and Grush, 1999; Grush, 2004; Pezzulo and Castelfranchi, 2007) propose that internal models permitting to *emulate* the external reality are firstly developed for the sake of action control. Once established, they are exapted (i.e., used for a function other than that for which it was developed) for bootstrapping increasingly complex functionalities such as simulative planning and pursuing distal goals. The diversity of anticipatory capabilities existing in nature thus depends on a progressive process of disengagement from the current sensorimotor cycle, enabled by progressively detached anticipatory representations. Since this process exploits an 'inner simulation' mechanism, however, anticipatory representations remain intimately related to situated action and maintain its nature (Barsalou, 1999; Damasio, 1994; Grush, 2004; Hesslow, 2002). The same process can also be in place for social cognition, since anticipatory mechanisms for engaging in future-oriented and social-oriented activity share the same neural substrate.

Consistently, Clark and Grush (1999) put forward the challenge of understanding the anticipatory aspects of cognition in a naturalistic framework, and as originating from situated action. They argue that a crucial step toward truly *cognitive* robotics is reframing the concept of representation in a situated and embodied perspective. They propose that anticipation is the key element, and anticipatory mechanisms (in particular simulative mechanisms) are responsible for bootstrapping grounded representations:

> agents [that genuinely cognize their worlds] are able to substitute inner dynamics for ongoing environmental stimulation, and command adaptively valuable inner spaces that they use to sculpt and modulate their more direct engagements with the world. It is these 'Cartesian Agents' we believe, that must form the proper subject matter of any truly cognitive robotics. Clark and Grush (1999, p. 13)

Grush (2004) has put forward this idea and proposed the *emulation theory of representation*: the most comprehensive account nowadays on how representation originates from anticipatory mechanisms that can be used online for action control and re-enacted off-line for enabling a number of sophisticated cognitive capabilities such as visual imagery, reasoning, theory of mind phenomena, and language. A conceptual analysis of the passage from sensorimotor skills to higher level cognitive capabilities based on anticipation can also be found elsewhere (Gardenfors, 2003; Hurley, 2005; Pezzulo and Castelfranchi, 2007; Pezzulo, 2008a).

Representations Are Grounded Because They Remain Related to Prediction If we assume that representations arise for prediction, and continue to depend functionally on prediction, we can understand them in a novel perspective, that is intimately related to situated action. One definition that is compliant with this view is provided by Bickhard and Terveen's (1995) *interactivism*: representations are ways for setting up indications of further interactive potentialities, and thus serve for future interactions. A related view is Smith's (1996) *intentional dance*.

An example of robotic implementation of the Piagetian, constructivist approach to the formation of object representations can help illustrate this point. In the study by Drescher (1991) concepts for objects are developed autonomously on the basis of (actual or expected) interaction effects. Objects (called *synthetic items*) are not provided to the agent but 'discovered', or better postulated, as a common cause of the expected success of a number of actions; objects are then explanations of sensorimotor patterns. If the agent moves its hand to the left and touches (or expects to touch) a surface, or moves its eyes to the left and sees a circular shape, etc., it can postulate that there is a common cause in all these behavioral effects and then 'create' a synthetic item. Later on, it can predicate based upon this item for forming more complex representations, including action representations.

In a similar vein, Roy (2005) has proposed that concepts for objects, which are, for example, reachable or graspable, are grounded by object schemas (similar in spirit to synthetic objects), which regulate actual behavior and at the same time encode predictions on the consequences of an expected interaction. One advantage of this framework is that schemas for actions, objects, and linguistic symbols share the same representational basis and have demonstrated to be successful in complex cognitive tasks such as robot control and linguistic communication.

This approach has implications for the symbol grounding problem as well. Roy (2005) has proposed that it depends on two mechanisms relating agent and environment: causation (from environment to agent) and anticipation (from agent to environment). Representations, including goal representations, are thus grounded thanks to the circular causality of ACTION + EXPECTATION and OBSERVATION + CAUSATION. This circular mechanism could also explain how we attribute *causality*: it is our productive apparatus that permits the reading of events in the world as causally (instead of simply statistically) related, which is essentially the Kantian solution to the problem of causality.

From a philosophical point of view, this approach can solve the problem of how to justify representations without falling in the grounding problem. On the one side we want to highlight the role of representational, anticipatory processes in cognitive agents, but on the other side we need a naturalistic account for representations that is entirely compatible with situated and embodied approaches to cognition. This point is illustrated nicely in Clark's (1998) *minimal representationalism*: "Minds may be essentially embodied and embedded and still depend crucially on brains which compute and represent."

1.2.4 The Unitary Nature of Anticipation

Overall, we have illustrated how several conceptual frameworks have been developed that indicate anticipation as a key element for cognition and for the development from simpler to more complex cognitive capabilities. As a conclusion to this section, we want to stress one of the innovative aspects of the anticipatory approach to cognition. Notwithstanding the fact that anticipation has multiple facets and has multiple realizations in brains and behaviors, we believe that it has to be considered a *unitary phenomenon*, a hallmark for natural and artificial cognition. Consistently with the idea of a cognitive mind as an anticipatory device, we argue that anticipation is inherently involved in—and in many cases a necessary condition for—several cognitive functions. Once a cognitive mind has evolved the power to deal with the future, this opens the possibility of an entirely new set of capabilities and opportunities, and it can realize proactive and goal-oriented behavior.

For this reason, we believe that a real understanding of the phenomenon of anticipation will come from a study of its unitary aspects rather than (or, better, together with) its different realizations, behavioral and neural. Our objective is then to provide a unitary perspective on the study of cognition and its development by focusing on anticipation and, more in general, on the capability and need to deal with the future. In the rest of the book we pursue this objective in several ways: We provide adequate definitions and taxonomies that help highlight the unitary aspects of prediction and anticipation, and we analyze the powers and limitations of anticipation in natural and artificial cognition, also providing a number of examples.

Since our analysis is grounded in biological and psychological evidence, we now proceed with reviewing evidence for anticipatory phenomena in cognition, both in simple and complex organisms, and we illustrate a unifying view of natural cognition based on anticipation.

1.3 Anticipation in Living Organisms

Besides the conceptual perspective on AI research progress, that is, from symbols, to reactivity and situatedness, back to combinations of these mediated by anticipatory mechanisms, there is also a biological, psychological perspective that strongly points toward the ubiquity of anticipatory mechanisms in animals and humans.

Several converging directions of empirical research indicate a crucial role for anticipatory mechanisms in cognitive functions. These mechanisms range from simple, such as sensorimotor coordination, to highly complex, such as decision making in social domains, social imitation and learning, or communication. In this section, we review biological and psychological evidence for anticipatory mechanisms in the brain and the consequent behavioral capabilities of animals and humans.

1.3.1 Anticipatory Natural Cognition

Several animal and human capabilities require an estimation of future states of affairs for compensating the dynamicity of the environment: for example, the motor preparation of the prey-catching behavior of the jumping spider (Schomaker, 2004) or balancing a pole with one hand (Mehta and Schaal, 2002). It has even been proposed that all motor control is mediated by anticipatory information, which is generated by internal predictive models that permit the emulation of the environment (Doya, 1999; Kawato, 1999; Wolpert et al., 1995).

Visual attention is also greatly influenced by expectations, as testified by classic experiments. Yarbus (1967), for example, showed that a visual scene is scanned differently depending on the observer's intentions. This influence of expected stimuli for orienting attention has been reported not only in humans, but also in pigeons (Roitblat, 1980) and monkeys (Colombo and Graziano, 1994). The constructive and active aspects of perception, and in particular the top-down influences, are discussed in detail in Engel et al. (2001). On this basis, models of the visual apparatus including (hierarchical) predictions have been proposed such as *predictive coding* (Rao and Ballard, 1999) and *prospective coding* (Rainer et al., 1999).

More complex anticipatory capabilities, which are referred to as 'simulative', permit the prediction and processing of expected stimuli in advance. For example, Hesslow (2002) describes how rats are able to 'plan in simulation' and compare alternative paths in a T-maze *before acting in practice*. Simulation can also be used for the prediction of danger. Damasio (1994) argues that during decision making humans engage in 'what-if' simulated loops of interaction with the environment in order to evaluate in advance, via *somatic markers*, possible negative consequences of their actions.

Constructivists such as Piaget (1954) have argued that *sensorimotor schemas*, which enable the prediction of action effects, are progressively developed by means of an active exploration and interaction with the environment, leading to understanding and categorization. The view of situated activity as the basis of cognition, including conceptualization, has been recently revitalized and 'motor' approaches are gaining popularity. One piece of evidence that understanding comes from activity and exploration comes from an experiment performed by Held and Hein (1963): Kittens that were unable to move autonomously in the environment (i.e., those being only passively moved) failed to categorize it correctly, being, for example, unable to avoid cliffs, which shows that they did not develop appropriate depth perception. On the other hand, kittens that were raised similarly (they essentially had nearly the same perceptual input) but that had the possibility to move showed successful categorization and depth perception.

1.3.1.1 Anticipatory Human Cognition

We humans are able to perform a plethora of anticipatory mechanisms that seem to go far beyond the capabilities of other species. We are the "symbolic species" (Deacon, 1997), which was able to develop language and rather complex social structures and cultures. Some interesting examples of these capabilities include:

- We can formulate novel goals and plan in view of *future* needs (this includes abstract and distal ones such as having fun or becoming famous). The possibility to anticipate oneself could have lead to the capability to coordinate one's own actions in the present and in the future, and to have a sense of 'persisting self'.
- We can formulate expectations at an increasingly high level of abstraction and can use these to regulate our actions. For example, we can decide whether or not to apply for a job depending on our expectations about the satisfaction it will provide us, the salary, the free time, the success, etc. Not only can we formulate such abstract expectations, but we also can 'match' them with imaginary futures and select among them (albeit often only with a certain degree of success).
- We are capable of *substitution* (Piaget, 1954), that is, to manipulate mentally our representations before or instead of acting in practice. Probably several animal species are able to use their *internal models* of phenomena for making mental manipulations, but we humans are able to use that ability systematically. A mechanic can assemble and dismantle a motor in his mind before doing it in practice. An architect can propose different plans for restructuring a house. Thanks to anticipation it is possible to deal with entities also when they are not present as stimuli: an ability that is crucial for defining an agent's *autonomy* (Castelfranchi, 1995).
- We can heavily modify and adapt the environment to ourselves, not only vice-versa. While other species adjust their representations to fit the actual world, we often act in the world in order to make it fit our representations of what we want, that is, our goals. Several animal species have the capability to adapt their environments, such as building a nest, but typically they do that in a very stereotypic way. We humans do not have this limitation and have heavily modified our environment to fit our present and especially *future goals* (Gardenfors and Orvath, 2005; Pezzulo and Castelfranchi, 2007).
- We can imagine ourself in the future and reason about possible futures. Tulving (2005) has argued that the capability to engage in 'mental time travel' in the past and the future is a uniquely human capability. Although this view has been questioned, and it might be the case that this capability is also available to other animal species to a certain degree (see e.g. Hesslow, 2002), humans can use mental simulation with unchallenged flexibility. Moreover, recent neurobiological studies (see Schacter et al., 2007 for a review) indicate that the process of imagining future events involves the same brain structures that are necessary to form episodic memory traces. This suggests a novel view of memory, whose main adaptive advantage could be providing building blocks for mental simulation and not (only) remembering. This fact could explain the constructive nature of memory: what is needed to imagine the future is the capability to flexibly recombine information from the past rather than simply replaying the past. Although this view is quite novel in psychology and neuroscience, the relevance of mental simulation is highlighted by several research programs, including *prospection* (Buckner and Carroll, 2007), *episodic future thinking* (Atance and ONeill, 2001), *memory for the future* (Ingvar, 1985), and the prospective brain (Bar, 2007).

- Our highly sophisticated social life appears to rely on anticipatory capabilities as well, such as coordination and cooperation, perspective taking, imitation, theory of mind, and language (Knoblich et al., 2005; Frith and Frith, 2006; Gardenfors and Orvath, 2005; Iacoboni, 2003; Rizzolatti and Arbib, 1998).
- We have developed symbols and a symbolic language. Various researchers (Arbib, 2002; Gardenfors, 2003; Gardenfors and Orvath, 2005; Swarup and Gasser, 2007) have recently discussed how anticipation is a precursor to symbolic communication and permits the evolution of symbols, and then the development of humans as *the symbolic species* (Deacon, 1997).
- Gallese's (2001) *Shared Manifold Hypothesis* and Hurley's (2005) *Shared Circuits Hypothesis* describe how anticipation is essential for bootstrapping capabilities in the individual and especially social sphere. Both describe several evolutionary steps necessary for the development of our current cognitive and socio-cognitive capabilities.

This small list could be expanded at will and only intends to point out how ubiquitous anticipatory mechanisms appear to control and guide our behavior and cognition in general. We now proceed with considering concrete neuroscientific and psychological evidence for anticipatory behavior in animals and humans and how such behavior may come about.

1.3.2 Anticipatory Codes in the Brain

What is the neural substrate of these forms of anticipatory behavior? Is there a unique way to predict, or are there many? Are neural substrates for predictions shared among species? Are there specific brain structures that mediate complex forms of anticipation in mammals and in the human brain? Notwithstanding the fact that several aspects are still obscure, the empirical literature is huge, and cognitive psychology and neurobiology continue to unravel mechanisms and processes based on anticipation in humans and other animals. We refer to Fleischer (2007) for a recent, excellent overview of anticipatory mechanisms in the mammalian brain and to Hoffmann (2003) for an extensive survey of the psychological literature.

Here, we instead try to summarize and systematize the empirical findings of anticipations with respect to the neural codes that could be involved in generating expectations—for the sake of discussing the implications of these findings on theoretical models. We can distinguish among two main kinds of anticipatory neural codes that can perform action anticipation and goal prediction. The former focuses on associative links and the latter on generative mechanisms and internal simulation (see Csibra and Gergely, 2007 for a comparison of these mechanisms with teleological reasoning).

1.3.2.1 The Ideomotor Principle and Associative Learning

The first proposal that the brain includes neural codes that relate expectations to action was formulated in the *ideomotor principle* (Herbart, 1825; James, 1890), which has recently received a number of empirical confirmations from psychological studies (Hommel et al., 2001; Kunde et al., 2004, 2007; Prinz, 2005). It has been proposed that an agent can learn to predict the outcomes of its actions and then store ACTION → EXPECTATION (A-E) (a.k.a. ACTION → EFFECT (A-E)) associative links, which may be neurally specified thanks to a *common neural coding* between perception and action (Prinz, 1990, 2003)[1].

What is relevant here is not only that anticipation is deeply integrated with action representation but that expectations can be used for triggering action. ACTION → EXPECTATION (A-E) sensorimotor codes, once learned, can be 'inverted' and become EXPECTATION → ACTION (E-A) links, which permit the activation of an action by its (desired) effects (Hommel, 2004). The relevance of this mechanism consists in its possibility to account for goal-directed actions in a simple and elegant way, since a desired (predicted) effect, and not a stimulus, is responsible for triggering an action.

Another account of prediction based on associative mechanisms is put forward by Bar (2007). Thanks to the similarity between past and novel stimuli, analogies are established that trigger prediction on the basis of associations that capture the most frequent trends in the stimuli. This kind of predictive mechanism is then of the type STIMULUS → STIMULUS and not ACTION → EXPECTATION. These association-based predictions permit the forecasting of what is more likely to happen in the same context, to preventively set up appropriate actions, and to enable priming (perceptual, semantic and contextual, see, for example, Anderson, 1983).

1.3.2.2 Generative Mechanisms: Internal Models

It has been argued that the brain uses *internal models*, which mimic the behavior of external processes, for motor control of action (Doya, 1999; Kawato, 1999; Wolpert et al., 1995). In particular, *forward models* permit the generation of expectations about the next sensed stimuli, given the actual state and motor command. We can further distinguish between *forward sensory models* (STATE + ACTION → SENSORY FEEDBACK) and forward dynamic models (STATE + MOTOR COMMAND → FUTURE STATE).

Inverse models (or controllers) instead take as input actual stimuli and the goal state and provide as output the motor commands necessary to reach the desired state. Taken together, inverse and forward models permit not only the performance of motor plans but also the control of it and in general the regulation of its behavior in noisy and dynamic environments.

Internal models permit an agent to run an 'inner sensorimotor loop' that parallels actual sensorimotor interaction as shown in Fig. 1.1. This inner loop is extremely

[1] Since in principle, several effects could be associated with an action, to be efficient, this mechanism needs a guarantee that those effects stored are the only relevant ones, such as those effects originating systematically from the same action, the effects that are rewarding, etc. Since associative mechanisms do not have internal states, this may come only from simple associative forms of learning such as Hebbian learning—see, for example, Butz, 2002a and Drescher, 1991 for principled design approaches to tackle this problem.

useful for regulating motor control. For example, it can compensate for delays in sensory feedback and cancel the self-produced part of the input from sensory stimuli (Blakemore et al., 1998), etc. Empirical evidence is reported for a role of internal models in visuomotor tasks (Mehta and Schaal, 2002), eye movements (Shidara et al., 1993), imagery (Jeannerod, 1994), motor execution (Wolpert and Flanagan, 2001), and sensorimotor learning (Wolpert and Flanagan, 2003). It is worth noting that internal models could provide support for anticipation at different time scales and granularity, for which hierarchical models of action control and recognition over time have been proposed (Haruno et al., 2003; Johnson and Demiris, 2005a).

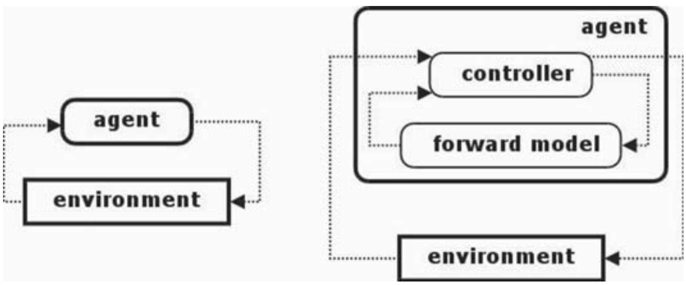


Fig. 1.1 Left: an agent engaged in sensorimotor interaction with its environment. Right: an agent running an 'inner sensorimotor loop' which parallels actual interaction.

One advantage of this model from the computational point of view is that it is rooted in standard control theory, and has parallels with the concepts of Kalman filtering (Kalman, 1960). Another advantage is that it provides a unitary view of anticipation in the brain, suggesting that a unique mechanism could mediate action performance, understanding, and imitation, as well as event understanding.

1.3.2.3 Generative Mechanisms: Compact Coding

Predictive coding is an account of the functional architecture of the brain that originates from Helmholtz's models of perception. It is based on the idea that the sensory brain has a hierarchical structure that has evolved to represent or infer the causes of changes in its sensory inputs (Friston, 2005, 2003; Kilner et al., 2007). It actively does that by means of generative mechanisms that actively predict the input, bias further processing in a top-down manner, and are modulated by bottom-up feedback. This approach has the advantage of being able to integrate cybernetic models of prediction (based on empirical Bayes, Kalman filters, etc.) into a well accepted biological framework.

Besides the further reaching predictive capabilities of such models, though, computational models support this proposition, such as a bidirectional, hierarchical vision architecture, proposed by Rao and Ballard (1999). In this architecture, a higher layer includes 'concurrent perceptual hypotheses' that convey priors and modulate the lower layer. In turn, the lower layer sends back prediction error to higher layers as a result of a match or mismatch of perceptual actions.

Several related generative models, such as Bayesian systems and Boltzmann machines, have been used in vision, speech processing, sensorimotor integration, action execution and understanding, and decision making (cf. Dayan et al., 1995; Friston, 2005; Hinton and Dayan, 1996; Kording and Wolpert, 2006; Weber et al., 2006; Wolpert et al., 2003; Yuille and Kersten, 2006).

1.3.2.4 The Case of Mirror Neurons

Besides such strongly sensorimotor models, recent evidence suggests the presence of a *mirror neuron system* in monkeys and humans. This neural system was originally discovered in the ventral premotor cortex (F5) of macaque monkeys, where goal-oriented actions are encoded that are either performed by oneself or only visually observed while being performed by others (Rizzolatti et al., 1996; Rizzolatti and Craighero, 2004). Recently it has been shown that mirror neurons respond to action goals rather than to their surface characteristics, that is, 'ends' rather than 'means' (Umiltà et al., 2001), also the case with distal goals (Fogassi et al., 2005). This fact suggests a way to understand (and possibly imitate) not only other people's movements, but also actions and intentions (a model along those lines was proposed in Meltzoff and Moore, 1997).

Mirror neurons show the remarkable capability to encode the prediction of the goal of an action, both performed by self and others, using a single neural circuit. This fact has suggested the possibility of breaking the boundaries between the individual and the social spheres (Gallese et al., 2004), and lead to several suggestions that the mirror system may be involved in a number of socio-cognitive functions such as action understanding, imitation, language (Iacoboni, 2003; Rizzolatti and Arbib, 1998), as well as empathy (Gallese, 2001).

For the sake of our analysis, this fact is particularly relevant since it demonstrates that the same anticipatory mechanisms could be used for goal-oriented actions as well as for action understanding and imitation, so that future-oriented and socially-oriented functions can share the same neural basis (Decety and Chaminade, 2003; Iacoboni, 2003; Jeannerod, 2001).

1.3.3 Simulative Theories of Cognition, and Their Unifying Nature

Recent research on anticipatory and in particular generative mechanisms in the brain has revitalized so-called 'motor' or 'simulative' views of cognition, which highlight the role of the motor apparatus in all aspects of cognition, ranging from situated actions to high level cognitive capabilities. Simulative theories of cognition indicate that internal mechanisms used for action monitoring and control can be re-enacted for generating long term expectations and 'covert' simulation of overt behavior (Cotterill, 1998; Grush, 2004; Hesslow, 2002). At the same time, other cognitive phenomena such as understanding and imitating actions performed by others, reasoning, theory of mind, and language can be accommodated within the same theoretical framework (Blakemore and Decety, 2001; Frith, 2007; Gallese, 2001; Gallese et al., 2004; Iacoboni, 2003; Jeannerod, 2001; Kilner et al., 2007; Rizzolatti et al., 2001; Wolpert et al., 2003). Anticipatory representations produced by anticipatory mechanisms can then be used in action preparation, execution, control, and mental action simulation.

Central to this family of models is the concept of internal simulation, emulation, or re-enactment. This productive aspect, which distinguishes simulative theories from similar views based on associative mechanisms, is now gaining relevance in the literature of mirror neurons (Gallese and Goldman, 1998; Oztop et al., 2006; Rizzolatti et al., 2001) and internal models (Doya, 1999; Wolpert et al., 2003).

According to (Gallese, 2000):

> To observe objects is therefore equivalent to automatically evoking the most suitable motor program required to interact with them. Looking at objects means to unconsciously 'simulate' a potential action. In other words, the object-representation is transiently integrated with the action-simulation (the ongoing simulation of the potential action).

Hesslow's (2002) simulative theory of cognition describes thinking as 'covert' behavior. In this sense, anticipatory capabilities permit the re-enactment of motor programs required for situated interaction. For this reason, there is no gap between the sensorimotor and the cognitive mechanisms that enable behavior. Hesslow (2002) suggests the following three aspects:

> (1) Simulation of actions: we can activate motor structures of the brain in a way that resembles activity during a normal action but does not cause any overt movement. (2) Simulation of perception: imagining perceiving something is essentially the same as actually perceiving it, only the perceptual activity is generated by the brain itself rather than by external stimuli. (3) Anticipation: there exist associative mechanisms that enable both behavioral and perceptual activity to elicit other perceptual activity in the sensory areas of the brain. Most importantly, a simulated action can elicit perceptual activity that resembles the activity that would have occurred if the action had actually been performed.

Related views are put forward by Grush (2004) and Barsalou (1999) under the names of 'emulation' and 'simulation' theories of cognition, respectively. These authors suggest two comprehensive attempts to integrate a plethora of cognitive functions, such as motor control, reasoning, theory of mind phenomena, and language, under the same framework that emphasizes the productive aspects of cognition. These aspects are generated by the capability of the mind to construct models of its environment that can be re-enacted and run either on-line or off-line.

1.3.3.1 Kinds of Internal Simulations and Simulative Theories

While several theories have been proposed as "simulative", the authors are referring usually to different aspects of a "simulation" and to different mechanisms for producing simulations. One important distinction is among mental simulation in the sense of off-line processing, or 'covert' behavior, as highlighted elsewhere (Blakemore and Decety, 2001; Grush, 2004; Hesslow, 2002), and mental simulation as the cognitive basis of social skills such as imitation and mind reading (Gallese and Goldman, 1998).

Similarly, Decety and Grèzes (2006) distinguish among several kinds of simulative approaches, which differ in level of access (automatic vs. conscious) and in scope (motor aspects vs. more complex cognitive states). One view is conscious reactivation of previously executed actions stored in memory (Decety and Ingvar, 1990), which can also be chained for producing long-term expectations (Cotterill, 1998; Hesslow, 2002) and 'simulate' overt behavior. Another view stresses the role of unconscious activation of several aspects of action, including its goal, the means to achieve it, and its consequences (Jeannerod, 1999, 2001): all these representations belong to the covert phase of motor preparation and can be reused for observing actions performed by others. The third view is related to the simulation-theory in philosophy of mind (Goldman, 2005) and explains the capability of understanding other's mental states—including beliefs, desires, and feelings—with the capability to "put ourselves into the other one's shoes" by simulation. It is possible that only some of these mechanisms can be used for generating long-term predictions, which can be used off-line, that is, detached from the current sensorimotor context.

Notwithstanding the differences between the approaches, and assuming that the brain could have alternative ways to simulate and emulate, this book focuses on the understanding of the unitary nature of motor and simulative theories of cognition. Simulative mechanisms, in fact, have been claimed to be involved in a plethora of individual and social cognitive functions such as perception, action performance and understanding, decision making, imitation, intentionality, etc. and have the potential to provide a unitary approach.

1.3.3.2 Action Performance, Understanding, and Imitation with a Unique Mechanism

Simulative theories of cognition, which usually stress the role of anticipatory, generative mechanisms, challenge traditional models of cognition, in which perception and action as well as the individual and social spheres are separated domains. Simulative theories provide means to integrate these domains with perception and action. This fact is extremely relevant because both it suggests a suitable engineering methodology and it has a solid biological basis.

Decety and Grèzes (1999) as well as Jeannerod (1999) show that there is a common neural substrate between action production and imagination—evidence that suggests functional equivalence. For example, mirror neurons, shared neural representations, and simulative processes have been argued to be involved in imitation (Iacoboni, 2002; Meltzoff and Decety, 2003), distinguishing the self from others (Decety and Chaminade, 2003), mind reading (Gallese and Goldman, 1998), and language production and understanding (Rizzolatti and Arbib, 1998). These facts have suggested that simulative mechanisms can explain individual and social capabilities in a unique framework. Three related problems that are inferring which action to perform for achieving a desired goal could be realized by the same mechanism. Several authors argue that generative mechanisms for controlling actions can be reenacted endogenously for perceiving, understanding, and imitating actions performed by other agents in order to understand behavior and to infer intentions from observed actions (Blakemore and Decety, 2001; Gallese, 2001; Gallese et al., 2004; Iacoboni, 2003; Jeannerod, 2001; Kilner et al., 2007; Rizzolatti et al., 2001; Wolpert et al., 2003).

According to Rizzolatti and Arbib (1998, pg. 190): "Individuals recognize actions made by others because the neural patterns elicited in their premotor areas during action observation are similar to that internally generated to produce that action". Imitation can consist in the re-enactment of the internal generative models that better fit the observed agent's goal, providing that it is in the motor repertoire of the imitating agent (Demiris and Khadhouri, 2005; Demiris, 2007; Iacoboni, 2002, 2003).

Anticipatory Mechanisms for Generating Different Kinds of Predictions Are there unitary mechanisms in the brain for producing simulations, or not? Predictions of different kinds (rewards, sensory, different state predictions) may require different mechanisms, and the same can be true for predicting one's own actions and external events. However, there could be anticipatory mechanisms that can flexibly learn to generate different kinds of predictions. One example are forward models, which have been proposed to be involved in the prediction of events, both self-generated and external. Schubotz (2007) distinguishes among prediction of events that we can or can not reproduce. Consider as an example of the first case observing the action of walking, and of performing an highly skilled action such as juggling (assuming that we are not able to juggle). In the former case we can use (re-enact) our own sensorimotor system in order to predict possible effects of other's actions. In the latter case, since our behavior repertoire does not include juggling, our internal models are only able to provide us with partial sensory information. However, as Schubotz (2007, pg. 216) claims,

> Forward models for events are not categorically different from forward models for actions. Forward models for events are just a fraction of forward models for actions, a fraction that misses the full-blown interoceptive and exteroceptive description of action models.

1.3.3.3 Anticipatory Action Control and the Sense of Agency

Anticipation must thus enable not only the simulation of individual and social spheres, but also the distinction of the two. If the same neural states are involved both in action performance and in other's action recognition—consequently being engaged in a 'we-space' (Gallese, 2001)—how do we distinguish ourselves from others?

The development of a sense of agency, which permits the understanding the self-attribution of the effects of (our own) actions, has been discussed by Piaget (1954) and Meltzoff and Moore (1997). They suggest that children learn the 'boundaries of their prediction' and thus develop a *body scheme*.

A comprehensive theoretical framework that relates anticipation, and in particular internal models, to agency has been recently proposed by Frith et al. (2000).

They discuss how the failure to access anticipatory signals that are produced during motor control (e.g., efference copies of motor commands) produces deficits in the sense of agency, and thus conclude that awareness of those anticipatory signals is essential to be able to correctly self-attribute an action or an intention. Along the same lines, Pacherie (2007) discusses the phenomenology of first person agency in terms of simpler experiences: intentional causation, the sense of initiation and the sense of control, all dependent on anticipation.

1.4 Conclusions

This introductory chapter has given an overview of different facets of anticipations and anticipatory behavior from a cognitive science perspective. We have introduced anticipation and anticipatory behavior from the theoretical point of view, we have illustrated how anticipatory mechanisms enable a range of anticipatory capabilities in natural cognition, and we have argued that anticipation is a unitary and foundational phenomenon in cognition, and ultimately that a cognitive mind should be conceived as a future-oriented device.

It has been put forward that goal-oriented systems inevitably need to have anticipatory goal representations to be able to control goals flexibly and adaptively. To learn such goal-based control structures, forward and inverse models need to be learned. Forward models allow the anticipation of a reachable goal and its activation triggering suitable motor commands with parallel inverse models. Moreover, forward models give rise to many more predictive capabilities, such as the capability of goal-inference represented in mirror neurons. The inference options strongly depend on the level of abstraction and modularity with which environmental states and circumstances are represented in the brain. The capability of long-term forward model predictions and goal-oriented behavior needs to reach a symbolic level in order to accomplish the flexibility and determination observable in humans.

Besides these strong planning capabilities, forward simulation also allows the development of self, since motor commands lead to the most reliable sensory effects. Thus, motor commands allow an accurate and reliable representation of motor-dependent forward models. In later developmental stages then, these forward models are used and mirrored to understand the agency of others and in effect, their current intentionality and even emotional state, leading to the capability of language and empathy (Arbib, 2002; Gallese, 2001; Gallese et al., 2004).

In conclusion, we propose again that a crucial challenge for cognitive systems research is to understand the passage from reactive to anticipatory natural cognitive systems, and to do the same thing in the realm of artificial cognitive systems. For the design or artificial cognitive systems, however, it does not seem to be sufficient to simply program a simulative system that uses its predictive capabilities for goal selection, imitation, motor control, reasoning, etc. Rather, the discussed different facets of anticipation need to be pinpointed and then modularly structured, as the brain does. Thus, the next chapter proposes an overarching, modular taxonomy of anticipatory mechanisms and their purpose to yield effective, flexible, and adaptive cognitive agents.

Chapter 2
The Anticipatory Approach: Definitions and Taxonomies

Giovanni Pezzulo, Martin V. Butz, and Cristiano Castelfranchi

> The value of a symbol is that it serves to make thought and conduct rational and enables us to predict the future. *Charles Sanders Peirce*

The anticipatory approach that we propose consists in understanding and conceptualizing anticipation and anticipatory behavior in natural cognition and implementing them in artificial systems. We propose that anticipatory systems have capabilities that go far beyond those of purely reactive ones and that anticipation is a prerequisite for several cognitive functions, and in general for goal-oriented behavior.

Consequently, in this chapter we first provide definitions for anticipation and anticipatory behavior (Section 2.1). We then classify and distinguish different types and aspects of predictive and anticipatory mechanisms (Section 2.2). Next, we classify anticipatory mechanisms in goal-oriented behavior (Section 2.3) and finally distinguish anticipatory mechanisms in learning structures (Section 2.4).

The classifications put forward in this chapter are meant to clarify what anticipatory mechanisms are and how they can influence cognitive systems' behavior and learning. Chapter 3 then reviews the potential benefits of anticipatory mechanism and also discusses potential drawbacks. In the subsequent chapters, we then take the artificial systems perspective and put forward the requirements and the capabilities of currently available anticipatory cognitive systems.

2.1 Anticipatory Systems, Anticipation, and Anticipatory Behavior

Although anticipations and predictions are often used nearly as synonyms in natural language, in scientific realms there is a clear distinction between predictive systems and anticipatory systems. Generally, anticipatory systems are those that use their predictive capabilities to optimize behavior and learning to the best of their knowledge. Rosen (1985, ch. 6) might have been one of the first who put this idea into a useful definition. An anticipatory system is:

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 23–43, 2008.
© Springer-Verlag Berlin Heidelberg 2008

> [...] a system containing a predictive model of itself and/or its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a latter instant.

More precisely, he also states that

> An anticipatory system S_2 is one which contains a model of a system S_1 with which it interacts. This model is a predictive model; its *present* states provide information about *future* states of S_1. Further, the present state of the model causes a change of state in other subsystems of S_2; these subsystems are (a) involved in the interaction of S_2 with S_1, and (b) they do not affect (that is, are unlinked to) the model of S_1. In general, we can regard the change of state in S_2 arising from the model as an adaptation, or pre-adaptation, of S_2 relative to its interaction with S_1.

The most peculiar aspect of anticipatory systems is thus their dependence on (predicted) future states and not only on past states. Although the definition provided by Rosen may be too strong (it excludes systems that coordinate with future states without explicitly representing them—we will call this form *implicit anticipation*), it describes the kinds of systems we are mainly interested in: those able to realize behavior mediated by explicitly formulated expectations. In order to produce explicit expectations, anticipatory systems need predictive mechanisms, which may have different realizations, but nevertheless share the common feature of predicting future states.

Thanks to their predictive mechanisms, anticipatory systems are able to anticipatory behavior, which may be defined according to (Butz et al., 2003b, p. 3) as:

> [...] a process or behavior that does not only depend on past and present but also on predictions, expectations, or beliefs about the future.

It is this capability to formulate predictions and to use them for own purposes that distinguishes an anticipatory system from a merely reactive one. For example, anticipation plays a key role in goal-oriented and proactive behavior, since patterns of actions can be selected depending on their expected outcomes and not (only) on stimuli that are available here and now. While reactive systems can be functionally described with STIMULUS → ACTION (S-A) behavioral patterns, anticipatory systems have instead (STIMULUS +) EXPECTATION → ACTION (E-A) behavioral patterns, which is permitted by the explicit prediction of a stimulus or an action effect (STIMULUS → EXPECTATION (S-E), or STIMULUS, ACTION → EXPECTATION (S-A-E)).

However, not all anticipatory behavior has the same functional structure or involves the same mechanisms. Anticipation has multiple facets and functions, powers and limitations. In the next subsections, we introduce a theoretical distinction among *prediction* and *anticipation*. We discuss the most relevant aspects of predictive and anticipatory capabilities and suggest taxonomies that help frame the analysis of actual system implementations and the design of novel anticipatory cognitive systems that realize multiple aspects of the put forward taxonomy.

2.2 Prediction vs. Anticipation

A clear distinction needs to be drawn between predictive systems—those that merely learn to predict—and anticipatory systems. Essentially, it needs to be accepted that predicting is not the same as anticipating.

Prediction	is a representation of a particular future event.
Anticipation	is a future-oriented action, decision, or behavior based on a (implicit or explicit) prediction.

Anticipation, that is the main focus of this book, is based on prediction and is especially important for cognitive systems since it is typically something the agent is concerned with.

To assess the predictive and goal-directed behavior capabilities of adaptive, cognitive learning systems, it is useful to contrast different types of predictions as well as different types of anticipatory behavioral influences. Thus, we now first propose a taxonomy for predictions and anticipations for the realization of goal-directed behavior.

2.2.1 Predictive Capabilities

Prediction has received a considerable deal of attention in several scientific disciplines. Examples are predictions of time series and genetic series, weather forecast, filtering and estimation in controlling plants, etc. Consequently, several methodologies have been developed that are based on a number of different approaches, ranging from (neuro-)biologically oriented to traditional engineering techniques. Some examples are *Kalman filters* (Kalman, 1960), neural networks such as Jordan and Rumelhart's (1992) type RNN or the LSTM (Gers et al., 2003), neurofuzzy methodologies (Tsoukalas, 1998), anticipatory classifier systems (Butz et al., 2003c), or Bayesian approaches (Wolpert and Flanagan, 2001).

While these are all predictive methods, the predictive mechanisms involved are rather different from one another. We now first introduce a taxonomy of predictive capabilities and then discuss a possible role of different sources of information for generating predictions.

2.2.1.1 Taxonomy of Predictive Capabilities

Predictive capabilities can distinguish predictive (learning) systems and help to clarify for which kind of prediction a cognitive system is most effective to generate and learn. Some systems are, for example, very effective to learn artificial grammars but may be very ineffective to work on visual flow predictions, etc. Thus, this taxonomy distinguishes different aspects of predictions. In later chapters, we then use this taxonomy for analyzing the capabilities of different predictive and anticipatory learning systems to actually generate specific types of predictions and anticipations.

Representation As a first criterion we distinguish which types of predictions each system is able to generate. First, it is necessary to distinguish between *discrete* and *continuous* predictions. Some systems are pure symbolic systems that depend on discrete, symbolic input. Others usually handle real-valued inputs or even combinations of both. Moreover, some systems predict (potentially discounted) reinforcement values while others predict (potentially pre-processed) sensory inputs. Besides the predictive input and output processing capabilities, it is necessary to specify if the predictive system can predict exact next input or if it has multiple levels of predictive abstractions available. Moreover, a distinction can be made between systems that are able to selectively predict partial aspects of the successive inputs and systems that predict all successive inputs (and/or generate chains of predictions).

Quality of Predictions Besides the nature of the predictions available, it is necessary to distinguish between different qualities of predictions. The most important question in this respect is if the predictions are concrete predictions of one next state or if they are able to code a number of potential predictions or a range of predictions. Similarly, it is necessary to distinguish systems that are only able to produce deterministic predictions with systems that can generate noisy predictions. In relation to this, predictions may be endowed with a confidence measure. Finally, it needs to be assessed if the system can learn a predictive model successfully given a Markov decision process (MDP) or also a partially observable Markov decision process (POMDP) problem. In the latter case, the system needs to learn internal state representations in order to bridge ambiguous inputs.

Time-Scale of Prediction Prediction can be done at different time scales. Most systems predict the immediate, next input property. However, there are others that can predict long-term dependencies. For example, one can predict the immediate $(t+1)$ and the long-term $(t+n)$ effects of an action. This can be done in different ways: for example, by explicitly learning to predict $t+n$, or by iterating several $t+1$ predictions. Again others are able to generate multiple predictions on various levels of abstraction in space and time. For example, a hierarchical predictive system may be able to predict immediate next sensory input on the lower level, but may predict the abstract flow of the sensory input on a higher level (such as that the flying ball will hit the approaching wall soon and will then bounce back with somewhat reduced velocity).

Obviously, the temporal horizon of prediction influences its accuracy. There is a trade-off between the time scale of prediction and its accuracy: the farther into the future trying to predict, the less accurate (and more uncertain) the prediction will normally be—as we know, for example, from weather forecasts. These limitations make it even physically impossible to predict $t+n$ with bounded accuracy if n is too big. One example is sensory prediction. It is likely that, when I am walking, I can predict the sensory effects of my next step –and there is evidence that such a prediction is unconsciously generated and used for controlling action. It is however very unlikely that I can predict the sensory effects that I will experience after one hundred steps. Another example is state prediction. If I watch somebody walking, I can predict with some accuracy where he will be in a few seconds, but not in several hours.

These limitations depend mainly on two facts. First of all, there are non-linear dynamics that can change the result drastically (for example, nearly unpredictable events such as the walking agent changing his mind). Even assuming that the dynamics do not change drastically, however, the computations are so complex and have an inherently recursive character—so that chaotic behavior is destined to occur eventually—leading any predictive mechanism to eventually produce highly error-prone predictions.

Designing accurate algorithms is not the only way to solve the trade-off between time scale of prediction and its accuracy, though. One alternative way is to generate more coarse-grained predictions. For example, I can predict with great accuracy that a walking man will be 'near the park', or 'in the same city' after one hundred steps. While fine-grained predictions could be inaccurate, coarse-grained predictions can still be very accurate. According to Tsoukalas (1998, p. 576): "Past and present observations are crisp numbers while predictions are fuzzy numbers." Consequently, it has been argued (Pezzulo and Castelfranchi, 2007) that the intrinsic limitations of predictive power could have lead to the development of coarse-grained, abstract concepts.

Generalization Capabilities Yet another highly distinguishing criterion for predictive learners is the capability of generalizing their predictions to similar events, sensory inputs, or other regular input properties. Hereby, similar situations may be characterized by sensory inputs that have a certain amount of stimuli, or stimuli properties, in common. For example, similar sequences may be generated by an identical underlying grammar. Finally, it needs to be assessed if the systems are able to identify abstract environmental properties beyond the generalization over simple stimulus input.

Focusing Capabilities Besides the generalization capabilities with respect to similar inputs, generalization capabilities or rather focusing capabilities are also highly important for the generation of predictions. For example, it might be very easy to predict certain parts of a visual scene (for example, constancy) but much more difficult to predict other parts in the scene (for example, a moving object or agent or the somewhat random patterns of the leaves of a tree on a windy day). Thus, a predictive system should be able to ignore or generate noisy predictions for inputs that are hard to predict whereas concrete near exact predictions for those parts that are easily predictable. Systems can be evaluated on their capabilities of ignoring irrelevant or distracting inputs. Moreover, they can be evaluated if predictions can be generated for parts of inputs only—either spatially or in a more object-related or property-related fashion.

Context- and Action-Dependent Predictions Although it seems very important for cognitive systems to handle:

1. action information and
2. context information

in ways other than bottom-up, sensory information, only few current predictive learners actually induce such distinctions. Thus, it needs to be assessed how systems can handle action information and context information. Similarly, reward information may be handled in different ways. Finally, it needs to be clarified how these different types of information may be merged, if handled separately.

Internal State Predictions Besides the predictions of sensory or pre-processed inputs, it might be highly advantageous to be able to predict own internal states. More importantly, future behavior and future relevancies may give advantages when evaluated for anticipatory behavior. Also, it might be advantageous to predict future motivational and emotional states to be able to exploit opportunities and avoid upcoming danger.

Prediction Related to One's Own Actions or of Other Events It is also very relevant to distinguish between predicting the consequences of one's own actions, and predicting other events that do not depend, or only partially depend, on our actions. The former have the form: ACTION → EXPECTATION (A-E), while the latter has the form: STIMULUS → EXPECTATION (S-E).

It is possible that both kinds of predictions are realized by a shared neural mechanism, such as an *internal forward model*, as proposed by Schubotz (2007). Nevertheless, they need to be distinguished, since the former describe regularities in the effects of our actions while the latter describe more 'objective' regularities in the environment.

This fact has relevant consequences in the process of development. Learning to distinguish what depends on us, and what does not, or what are the limits of our influence, is an important part of cognitive development and the formation of our *body scheme* (Piaget, 1954; Meltzoff and Moore, 1997), as discussed above. If we assume that we can experiment more readily with what depends on us, the effects of our own actions could be learned first, while predictions of external phenomena could depend on a self-other distinction that may come later in development.

2.2.1.2 Different Sources for Prediction

Predictions are typically based on prior knowledge, for example, on our past experience in similar situations or also information acquired by others. However, 'the past' can be represented and processed in different ways. Thus, predictions often depend on different sources of prior information and the exploitation of these to learn suitable predictions effectively.

Statistical Regularities The most popular mechanisms for prediction are based on statistical information that is accumulated. These mechanisms are then often mimicked or bootstrapped with Bayesian prediction or soft computing algorithms, such as neural networks or fuzzy logic. Typically these methods require a lot of information to be trained. Basically, all these mechanisms establish (implicitly or explicitly) a similarity between a set of past experiences and a novel experience, and use this similarity to generate predictions.

Analogy Another source for prediction is analogy, that is based on mapping of knowledge into multiple distinct domains, which permits predictions in a given domain without being trained in it. Analogical reasoners usually assume that a mapping with another domain (in which there has been some training) can be found. In this sense, the similarity is not established within events in a single domain, but in distinct domains. Analogy then focuses on abstract commonalities between domains, such as big dogs are dangerous, thus, also big tigers are dangerous or since I usually find milk close to butter I also expect to find flour close to bread.

In a sense, any categorization is an analogy with a past situation; here we restrict the term 'analogy' to inferences involving mapping between at least two distinct domains and establishing 'profound' or 'structural' similarities and not simply surface similarities.

Inference Predictions can be done on the basis of inference rules. For example, our 'ingenuous physics' model can tell us that all objects fall due to gravity. We can then predict that if we throw an object, it will fall, independently of the object and independently of any observation. On the basis of categorical knowledge (for example, 'all fragile objects that fall down will break'), inferences such as *modus ponens* can be done (for example, 'my glass is falling down' → 'my glass will break') that depend on the structure and not the content of information. Of course, in order to make such an inference it is required that information is first categorized (for example, deciding that a glass is a fragile object).

Prediction Based on the Functional Stance Until now we have mainly provided examples of prediction that are based on information at the level of physical properties of objects. There are however some parts of our environment that can be better understood if we think in terms of their function, and not their physical realization. For example, if we want to understand why birds fly, it is better to think of their wings in terms of their function (as a product of evolution) and not in terms of tissues and bones. In the same way, we can conceptualize artifacts such as cars and computers, whose function depend on the fact that they were explicitly designed for something.

Dennett (1987) introduces the term *physical stance* for models that are only concerned with physical and chemical properties of objects and the term *design stance* for models that are concerned with things like purpose, design, and function. When we assume the design stance, we can form predictions that are much more difficult on the basis of a purely physical stance: for example, we can predict that a car without the engine will not move, while a car without a seat may still move. Of course, the design stance only makes sense for some objects in our environment. It usually produces quite coarse-grained and abstract predictions, which, however, can be often behaviorally much more useful than fine-grained physical predictions.

Prediction Based on the Intentional Stance: The Social Domain Other sources for prediction exist in social behavior that permits the prediction of other's behavior. Using Dennett's (1987) terminology, if we treat an agent as an intentional entity we are assuming the *intentional stance* and we can ascribe to it beliefs, goals, desires, preferences, etc. (regardless if the agent has these mental states or not). We can then use a representation of the other's mind for predicting its behavior: for example, if we know that John likes music, we can predict that he will go to a live music show. This source of prediction is typically not available for entities to which we can not ascribe mental states.

In the social domain the capability to predict the other's behavior on the basis of the intentional stance is called *mind reading*. Two different accounts have been suggested. According to the former, called 'theory theory', an agent can build up a model of the minds of other agents, including their beliefs and goals, and use it to predict their behavior. According to the latter, called 'simulation theory', the agent represents other people's mental states by adopting their perspective and by using it's own mind as a model. In the former case the intentional stance is explicitly requested for building a model of the agent's mind, while in the latter case the intentional stance is implicit, since the further assumption that the agent's and the other's mind are similar is also required [1].

2.2.1.3 On the Complementarity of Different Forms of Prediction

These different forms of prediction are certainly not mutually exclusive. For example, we can predict the trajectory of a ball both on the basis of the 'ingenuous physics' model and on the basis of mechanisms such as internal forward models that predict a series of sensory states. These mechanisms have complementary powers and limitations: the former can be used in absence of perceptual input but arguably produces coarse-grained predictions (for example, the ball will fall quickly), while the latter needs perceptual input and produces fine-grained predictions (for example, the actual trajectory of the ball integrating forward predictions and sensory feedback). One can also predict behavior of an agent by using two models, one based on the physical stance and the other on the intentional stance. Of course, these models will generate different kinds of predictions, which may be complementary and may be integrated to an overall prediction as needed.

The modularity of the brain also suggests that predictive representations are dependent on various sensory inputs and may be combined in various ways dependent on their availability, current estimated reliability, current intentions, etc. (Maravita et al., 2003; Schwartz et al., 2004). Thus, complementarity among different predic-

[1] We do not discuss other possible approaches to simulation theory that are based on resonant states and mirror neurons (Gallese and Goldman, 1998), which may also be interpreted as requiring an intentional stance.

tive mechanisms is very relevant in modular and hierarchical architectures, which we discuss later on in further detail.

2.2.2 Anticipatory Capabilities

The predictive capabilities are mostly an important ingredient for successful anticipatory processes, however, anticipatory mechanisms come also in very different forms and can be useful in rather different circumstances. Thus, we now provide a taxonomy of different anticipatory mechanisms. First, we introduce a distinction between *implicit* and *explicit* forms of anticipation, and between *on-line* and *off-line* uses of anticipation. Next, we discuss the kinds of anticipations related to a variety of an agent's cognitive functions, offering a taxonomy of the involvement of anticipations in different cognitive processes.

2.2.2.1 Implicit vs. Explicit Anticipation

A difference exists between mechanisms that realize anticipation with or without explicit representations of future events. Anticipation can be realized procedurally, without any need for formulating explicit expectations: we call this *implicit anticipation*. But it can also be mediated by explicit representations: we call this *explicit anticipation* (cf. Butz et al., 2003b; Pezzulo, 2007).

Implicit Anticipation In nature there are several functions that require a reference to the future, and not only to current states. Take as an example sensorimotor coordination. What is required in a dynamic world is to coordinate with future events, not current ones—otherwise no one would be able to catch a ball, or to track a flying bird.

Although these functions include in principle an anticipatory aspect, they do not require necessarily an explicit anticipatory representation (although the anticipatory representation may facilitate or even enable some of these tasks). It is an empirical issue whether or not anticipatory behavior in nature is realized by means of representations, or not. One organism can simply learn to coordinate just in time with a dynamic event of the environment, say following a flying bird with its eyes, without predicting its movements. Gibson (1966) has provided excellent examples of sensorimotor coordination in perception that is not mediated by representations, and recently O'Regan and Noë (2001) have put forward this approach by arguing that sensing consists in the skilled exploitation of structure of sensorimotor contingencies by means of the perceptual apparatus. Similarly, evolution may have evolved reactive but implicitly anticipatory structures, such as the efficient behavior of bacteria following some gradients to better food sources or also the morphological intelligence embedded in animal bodies (Pfeifer and Gomez, 2004).

We then define implicit anticipation as a functional structure that is functionally anticipatory but realized by a reactive mechanism. Implicitly anticipatory structures can be described as: STIMULUS AT TIME T → ACTION THAT IS LEARNED/EVOLVED FOR TIME T+1. Almost all successful actions, also of the stimulus-response kind, require that the action is adapt to respond to the world as it is *after* the stimulus, otherwise it will often be too late. That is, it is in the nature of living organism's learning systems to be responsive to *signs of the future* (for example, a shadow or a noise as a sign of danger) that are perceived now, even without forming explicit representations of the future.

Explicit Anticipation However, it is not the case that all behavior in natural cognition systems work this way. Accumulating evidence, which we review in the next section, indicates the presence of anticipatory mechanisms that permit the generation of explicit expectations, and to use them for initiating, regulating, controlling, and selecting action. This view is also popular in AI and cybernetics. For example, according to Craik (1943, pg. 61), anticipation can generate imaginary experiences, that is, mental simulations of external realities:

> If the organism carries a small-scale model of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it.

Anticipatory mechanisms, however, can be diverse. They can have different realizations and permit the anticipation of different things. Not all explicit representations of future states are of the same kind: accordingly, Butz et al. (2003b) introduced a taxonomy that distinguishes among implicit, payoff, sensorial, and state anticipatory mechanisms:

- *Payoff anticipatory mechanisms* predict the payoffs of an action and base action selection on the payoff predictions.
- *Sensorial anticipatory mechanisms* produce sensory expectations and support perceptual processing.
- *State anticipatory mechanisms* produce more complex forms of expectations, such as event anticipations, that support decision making and execution.

2.2.2.2 On-Line and Off-Line Uses of Anticipation

An important distinction is among on-line and off-line uses of predictive mechanisms for the sake of enabling anticipatory behavior. Several cognitive functions, such as action control, require predictive mechanisms that are used on-line for generating predictions (for example, of the sensory consequences of an action's effects) that serve for regulating behavior. In this case, we refer to on-line mechanisms since the predictions are coupled to the current sensorimotor cycle: this guarantees that a prediction can be matched against a sensory input.

However, predictive mechanisms can also be used differently, for example, for the sake of generating long term predictions. We call this use, that is unrelated to the current sensorimotor cycle, an off-line use.

On-Line Uses of Anticipation Anticipation can be used on-line. For example, the prediction of the next sensory input can be used for the sake of action control and monitoring. Figure 2.1 provides an illustration of an agent that simply interacts with its environment (left), and an agent that produces expectations and runs an 'inner sensorimotor loop' that parallels actual interaction (right) (see also Grush, 2004). The agent on the right has an anticipatory mechanism, called *forward model* (see Chapter 1), that permits the prediction of next sensory input. This predictions can be used in several ways. For example, the agent can compare the predicted and sensed action effects for the sake of monitoring and adjusting its execution.

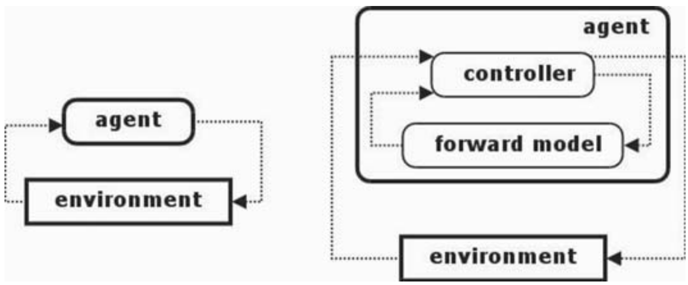


Fig. 2.1 Left: an agent engaged in sensorimotor interaction with its environment. Right: an agent running an 'inner sensorimotor loop' which parallels actual interaction.

Off-Line Uses of Anticipation Anticipation can be used off-line, as well. For example, the *simulation* of the long term effects of one's own behavior produces representations of a (possible) future that can be stored and/or internally manipulated. Processing information about the future produces striking adaptive advantages and opens the possibility to develop so-called high-level cognitive capabilities. Examples of specific advantages of conceiving the future now, such as comparing opportunities and possible action outcomes, simulative planning (Hesslow, 2002), forecast and avoid possible dangers (Damasio, 1994), are illustrated in detail in Chapter 3. As an example, Figure 2.2 illustrates a case of simulative planning in which possible trajectories are generated, evaluated (with respect to the compliance with current goals) and then selected.

Recently several artificial systems have been proposed van Dartel (2005); Hoffmann (2007); Pezzulo (2008b); Tani (1996); Vaughan and Zuluaga (2006); Ziemke et al. (2005) that generate long-term predictions by chaining short-term predictions, and use this capability for simulative planning (a virtual exploration of multiple possible plans before—or instead of—attempting them in practice) or for evaluating the outcome of their actions in advance.

A possible neural substrate for this mechanism is indicated by Middleton and Strick (2000): a 'loop' between the *cerebellum*, that produces sensory predictions (Blakemore et al., 2001), and the basal ganglia, that selects the action to perform and initiates movement (Redgrave et al., 1999).

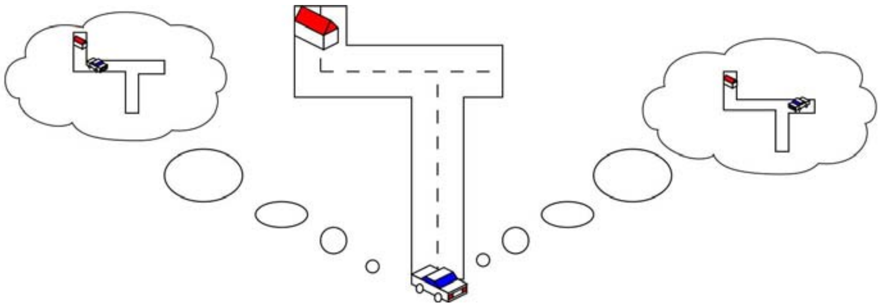


Fig. 2.2 Simulative Planning. Before deciding where to go, a driver anticipates the effects of his two choices by simulating driving in the two directions. He then selects the best option.

2.3 Anticipation and Goal-Oriented Behavior

Several theoretical frameworks in the philosophical, biological, and psychological literature are based on the idea that living organisms act purposively, and it is impossible to understand behavior without referring to its teleological structure (Rosenblueth et al., 1943). For example, the fact that an action can be defined only on the basis of its goal is a central tenet of the *ideomotor principle* (Herbart, 1825; James, 1890); similarly, Arbib and Rizzolatti (1997) argue that ACTION = MOVEMENT + GOAL. However, when one comes to the specification of what are the goals of a living system, and what is the functional structure of goal-oriented behavior, she faces a variety of different possibilities.

Clearly, living organisms of different complexities can have different teleological structures. Goals can be either explicitly represented and pursued, or embedded in the organism's design, which is shaped by evolution. This also means that in some cases 'purposively' has to be intended in the stronger sense (that is, a mechanism exists that includes an explicit goal, which triggers action), while in other cases it has to be intended in a purely functional way (that is, the organism is designed/evolved for realizing a goal that is not represented, such as 'surviving').

Castelfranchi (1998) introduces a distinction between strong and weak forms of goal orientedness, that are called *goal-oriented* and *goal-directed* respectively:

> In other terms, the agent's behavior is aimed at producing some result: thus we are talking of a goal-oriented action and of a goal-oriented agent. Among goal-oriented systems I will consider in particular goal-directed systems. In these systems not only action is based on perception, but the latter is also the perception of the action's effects and results, and the agent regulates and controls its actions on such a basis. The agent is endowed with goals, that is, internal anticipatory and regulatory representations of action results.

Here we are mainly interested in goal-oriented behavior in the strong sense (that is, goal directedness), and we claim that anticipation is required for acting truly purposively in that sense. One could say that implicit forms of anticipation are sufficient for acting purposively in the weak sense, while explicit forms of anticipation are needed for acting purposively in the strong sense.

In order to be goal-oriented in the strong sense, it is not sufficient that the behavior of an organism is implicitly oriented toward a (non represented) goal that is selected by evolution. On the contrary, the organism's actions are selected, executed, and monitored thanks to an explicit, anticipatory goal representation (and not simply with a functional equivalent that is selected by evolution). The fact that actions have such a teleonomic structure has recently been acknowledged in several psychological and neurobiological studies. For example, Jacob and Jeannerod (2005) argue that "An action is a goal-directed sequence of bodily movements initiated and monitored by what we shall call a *motor intention*", and Gallese and Metzinger (2003) highlight that, "In the monkey brain microcosm so far explored, the goal of grasping an object is still almost completely overlapping with the action-control strategies. Action control actually equates to the definition of the action goal: the goal is represented as a goal-state, namely, as a successfully terminated action pattern", see also Hommel et al. (2001) for a theory of action representation based on anticipatory and ideomotor codes.

Overall, truly goal-oriented behavior in living organisms is realized by functional structures that include anticipatory mechanisms and representations. In fact, enabling goal-oriented behavior may be one of the most important functions of anticipations in cognitive systems.

2.3.1 The Anticipatory Structure of Goal-Oriented Behavior

A comprehensive theoretical framework for understanding the relations between anticipation and goal-oriented behavior is that of *anticipatory behavioral control* (Hoffmann, 2003; Hoffmann et al., 2007), which has its roots in the *ideomotor principle*:

> Departing from the premise that almost all behavior of humans and higher animals is goal-oriented, the framework proposes that (1) a voluntary action is preceded by a representation of the to-be-attained effect(s), (2) learning of such effect representations is triggered by the comparison of predicted and actual effects resulting in the primary learning of action-effect relations, (3) situational context is integrated secondarily, (4) action-effect representations are activated by the need or desire of an effect-related goal, and (5) conditioned action-effect relations can also be activated by contingent stimuli.

Three points are worth noting. The first is that in this theoretical framework any behavioral structure is essentially goal-oriented and not simply stimulus-response. Action is purposive and not reactive. The second is that anticipation is required in all the phases: learning actions (their effects and how to execute them), deciding which motor action to execute, controlling and monitoring actual execution. The third is that, since the *ideomotor principle* indicates that goal-oriented behavior depends essentially on anticipation of the action's effects, it suggests an *'inversion' of the direction of causality*: goal-oriented behavior proceeds from effects to causes, from the future to the past, and not vice versa.

2.3.2 Not All Anticipatory Behavior Is Goal-Oriented

We have argued that goal-oriented action is anticipatory in nature. However, the converse is not always true: predictions and anticipations can have several roles, that we will review and illustrate in Chapter 3. For example, a system that is able to anticipate the effects of its own actions in the future could use anticipation (only) for learning and not for activating and controlling its current actions. According to Millikan (2004, p. 191):

> In the case of an animal that predicts and represents the results of its own purposive action, it is easy to slip into thinking that in representing the future it is guided, no just *by* the future it represents, but *toward* the future it represents. That is, it is easy to confuse anticipating future events for which it itself is purposively responsible with using representations of those future events as guiding goals.

2.3.3 Which Anticipations Permit Goal-Oriented Action?

According to Hoffmann (2003), two main kinds of anticipations are required in a goal-oriented behavioral structure (see Figure 2.3).

1. The first anticipation type is *effect anticipation*. It relates to the goal: the intended effect of the action. In order to specify a goal-oriented trajectory, a representation of the intended effect has to precede action. For example, the representation of the hand in contact with a pencil (including proprioceptive and exteroceptive input) should precede the initiation and selection of the actual grasping movement.
2. The second anticipation type is *start anticipation* (or *trigger anticipation*). It relates to the environmental, contingent conditions that signal a good opportunity to successfully produce the desired effect. For example, the action to grasp a pencil can only be started when a pencil is there. These opportunities have to be picked in time in order to successfully produce the goal-oriented action, otherwise there is the risk of being unable to do it any more. However, this does not mean that an opportunity triggers directly an action, unless it satisfies an agent's goal –this is one difference between goal-oriented and merely reactive agents.

The first anticipation type ensures that the action can be purposively selected and its effects controlled. The second anticipation type, in a sense, reverts the stimulus-response paradigm: it is not a stimulus that triggers an action, but the presence of an opportunity that is anticipated. In this way, the opportune context for an action is not passively experienced, but actively searched for and anticipated—for this reason, this kind of anticipation can also be seen as a form of context dependency.

However, anticipations can be of different kinds and can involve different representation formats: this depends on the fact that anticipatory goal-oriented action is

Fig. 2.3 Two kinds of anticipation

organized hierarchically, and each level of the hierarchy requires different kinds of anticipation. We henceforth investigate the issue of hierarchical goal-oriented action and its anticipatory components.

2.3.4 The Hierarchical Organization of Anticipatory Goal-Oriented Action

There is large consensus nowadays that in cognitive agents the control of action is multilevel and that the brain includes hierarchies of increasingly complex motor representations, that operate at different levels of representational abstraction and time granularity (Hamilton and Grafton, 2007; Fries et al., 2007; Wolpert et al., 1998, 2003). Different taxonomies have been proposed and we distinguish four levels: intention, goal, action, and movement. One can intend to fulfill an action, say *eating something*, even without specifying the means to do it. This is possibly the highest level, that we might call *intention*. The action can be therefore situated and anchored to the present situation, for example *eating this ice cream right now*: this is a *goal*. In order to execute the *action* to eat the ice cream, a further specification and instantiation is needed in terms of the appropriate sensorimotor representations that can realize the goal. Lastly, the appropriate motor programs and movements have to be actuated (and different patterns of actions can realize the overall action)[2].

An important related aspect is the realization of serial behavior, such as a sequence of actions appropriate for realizing a distal goal. Lashley (1951) firstly noticed motor plans do not consist of simple chains of stereotyped behaviors triggered by feedback. On the contrary, whole sequences or plans are established in advance as high-level motor plans that unfold into lower-level actions and movements—again, a hierarchical structure (MacKay, 1987). Evidence comes from numerous sources. For example, patterns of errors in motor behavior reveal knowledge of successive actions (for example, slips of tongue) (Lashley, 1951). Moreover, the time to initiate a series of movements increases with the number of movements to be produced (Sternberg et al., 1978).

Figure 2.4 provides an example of hierarchies of action organization in which action plans (for achieving an intention, goal and action respectively) are specified at an increasingly detailed level from high to low; see also (Pacherie, 2007; Wolpert and Flanagan, 2003). Hommel (2003) provides a comprehensive review of how planning is organized hierarchically, too, and how this provides several advantages, such as to prepare plans at a rather abstract level while letting details be filled in from lower level control components, effectively delegating actual control to already established sensorimotor loops. Hierarchical models of action control and recognition were extensively studied from a computational perspective, too (e.g., Bakker and Schmidhuber, 2004; Dehaene and Changeux, 1997; Paine and Tani, 2005; Schmidhuber, 1991b; Stringer and Rolls, 2007; Wiering and Schmidhuber, 1997).

[2] A different terminology is introduced in Bratman (1987) and also Pacherie (2000): future directed intention, present directed intention, motor intention, movement. Hamilton and Grafton (2007) suggest yet another action specification hierarchy: the intention level (the long-term goal of an action); the goal level (the short term goals for achieving the intention); the kinematic level (the body posture and movements in space and time); the muscle level (the pattern of muscle activity).

2.3.4.1 The Role of Expectations at the Different Levels

Expectations, arguably of different kinds, can have roles in action selection and monitoring that depend on the level they belong to.

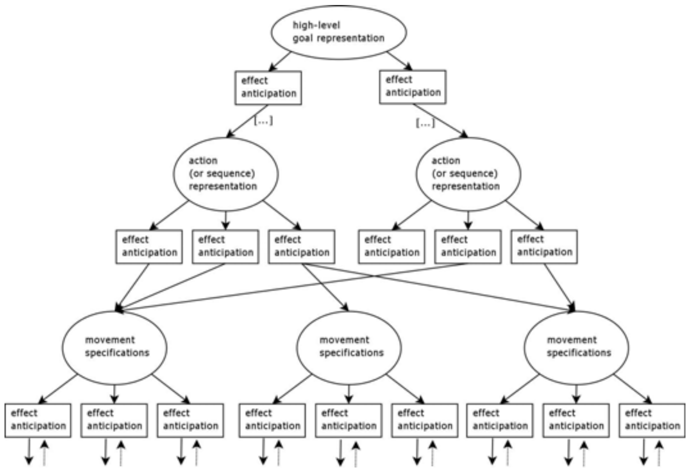


Fig. 2.4 Hierarchical organization of action.

At the higher level, expectations can be used for selecting the intention (among all the possible agent's goals) and for monitoring its achievement. At the level of goals, expectations can be used for selecting the action to execute and to monitor it. At the level of actions, expectations can be used for selecting the appropriate movements to actuate and to monitor their performance. However, the three kinds of expectations differ with respect to their contents, representational formats, and degrees of accessibility, and are generated by different mechanisms. Expectations formulated at the intentional and goal levels have declarative content that can be used, for example, for practical reasoning and deliberation (Bratman, 1987). Expectations at the action level have instead a sensorimotor format that can be directly matched with sensory and proprioceptive information (cf. Frith et al., 2000; Pacherie, 2007 for a discussion of consciousness access to these expectations).

2.3.5 Additional Elements of True Goal-Oriented Behavior

Are the anticipations of one's own effects, and the selection of an action based on these effects, the only ingredients of goal-oriented behavior? Traditional models of goal orientedness in cybernetics and AI such as the TOTE (test, operate, text, exit) model (Miller et al., 1960) indicate additional functions based on anticipation, which complement the ideomotor approach.

We have already seen that an important distinction is among goals that are or are not explicitly represented. Another important distinction is the degree of awareness and accessibility of the goal for internal manipulations. Differently from the ideomotor principle, TOTE is based on an explicit goal representation that is also available outside the action itself. Such a goal representation permits the performance of a number of on-line and off-line operations. One of them consists of evaluating the world and in particular matching the desired with the current state; in this respect, the test sub-process has the function of both the action trigger and the stopping condition. More precisely, the mismatch serves to select and trigger the rule whose expectation minimizes the discrepancy. Another effect of the explicit and available goal representation is that, as opposed to the ideomotor principle, the TOTE "knows" if and when a goal is achieved. Lastly, an explicit goal representation detached from actions can serve to deliberate and to compare goals for their selection on the basis of reasons and beliefs. Another distinction among mechanisms of goal orientedness is that while in the ideomotor principle desired results (motivating the action) are not distinguished from expected results of actions, the latter including the former, in the TOTE they are distinguished. We refer to Pezzulo et al. (2007) for a comparison of the TOTE model with the ideomotor principle, in which we conclude that they are highly complementary despite contrary views. They merely focus on different aspects of goal orientedness, which can be successfully integrated. Chapter 5 describes concrete artificial learning architectures that are capable of efficient anticipatory goal-directed learning and behavior.

2.4 Anticipation and Learning

L'intelligence organise le monde en s'organisant elle-même *Jean Piaget*

Anticipation does not only shape behavior at the time scale of action execution, but also at the time scale of evolution and learning. In order to satisfy its goals, a goal-oriented system needs to develop a behavior repertoire that reliably reproduces desired expected outcomes. For doing so, it needs two capabilities: (1) to learn to predict the effects of its own actions, and (2) to develop autonomously a behavior repertoire that is adapt to satisfy its goals. Both are enabled by predictive mechanisms.

2.4.1 Learning to Predict

All adaptive organisms direct their behavior toward effects that are rewarding, and at the same time they avoid (possibly) dangerous or punishing states. For doing so, learning to predict future events and future outcomes of their own actions on the basis of experience is a presupposition of any learning system.

Several learning algorithms permit the implicit learning of anticipations of rewards and sensory states. However, here we are mainly interested in learning explicit anticipatory models of the environment. We have in fact argued that in order to act really purposively, an agent has to develop anticipatory codes for predicting its environment, and in particular the effects of its actions. Explicit anticipatory mechanisms are then needed for flexible goal-oriented behavior. As we have reviewed, anticipatory codes in a living organism's brains can have multiple realizations, but for the sake of simplicity we indicate them as STIMULUS → EXPECTATION (S-E) or ACTION → EXPECTATION (A-E).

Different anticipatory mechanisms could require different learning mechanisms, possibly implemented in different brain areas (see, for example, Doya, 1999). However, expectations about future salient events, such as rewards and punishments, have a major role in learning, and the relevance of prediction in many if not all forms of learning is widely acknowledged. For example, one central tenet of theories of classical conditioning (Rescorla and Wagner, 1972) is that organisms only learn when events violate their expectations. The neural correlates of this phenomenon are starting to be understood. Neurobiological studies (Schultz et al., 1997; Schultz, 1998) illustrate that prediction of reward and selection of rewarding actions in primates could be mediated by error signals carried by dopaminergic neurons, and Doya (2002) describes meta-learning strategies based on these neuromodulatory mechanisms. These neural mechanisms that signal error in (reward) prediction and surprise can be used for learning the value, positive or negative, of an agent's actions. Other examples of learning mechanism that depend on predictions are discussed elsewhere (Doya, 1999; Kawato, 1999; Wolpert et al., 1995), in which *internal models* of the environment, including sensory and state predictions, can be learned by minimizing the sensory/state prediction error. Another example comes from the psychological domain: it has been suggested (Hommel et al., 2001; Kunde et al., 2004, 2007; Prinz, 2005) that the behavior repertoire of an anticipatory agent includes ACTION → EFFECT pairs that are developed by means of associative mechanisms.

2.4.2 Bootstrapping Autonomous Cognitive Development: Surprise and Curiosity

Another crucial problem for a goal-oriented system is to autonomously develop a repertoire of actions that can appropriately satisfy its goals, or which actions are worth learning. As recently acknowledged in neuroscience and in reinforcement learning research (Dayan, 2002; Schultz et al., 1997; Singh et al., 2005), two mechanisms depending on anticipation, surprise, and curiosity, are deeply involved in the autonomous cognitive development of actions.

Basically, surprise is a measurement of mismatch between prediction and actual sensation. In order to be useful for a learning system, it has to signal *novelty* and *relevance* at the same time. Novelty can be assimilated to unpredictable, at least in the sense that the learning system has not (yet) developed a repertoire of actions to reliably produce the intended effects, or an internal model that correctly predicts the event. However, not all unpredicted events are relevant for a learning system, since some of them are not salient to learn, or cannot be learned at all. In this sense, habituation mechanisms are needed to balance surprise and to cancel stimuli that are not useful for learning. Moreover, mechanisms for deciding which parts of the environment can be learned are needed as well since the system can otherwise get stuck in hopeless attempts.

If an agent has these mechanisms, it can autonomously direct its attention and decide which parts of its environment to explore and learn—which is referred to as "curiosity". The idea is that internal forms of reward could signal which actions are 'intrinsically interesting' to learn and that an agent can be 'curios' about parts of the environment that are surprising (signaling that they are novel) and predictable (signaling that they can be learned, that is, are believed not to be simply random) (Singh et al., 2005; Schmidhuber, 1991a, 2002). Surprise and curiosity lead to the exploration of novel and predictable parts of the environment and may thus bootstrap cognitive development. They may be considered a sources of 'internal motivation' to explore the environment and consequently expand one's own knowledge and related action repertoire. One hypothesis on the neural correlates of these mechanisms is put forward by Redgrave and Gurney (2006), who indicate that dopamine could signal the unpredictability of actions and may be required to learn novel actions.

2.4.3 From Willed to Automatic Control of Action and Vice Versa on the Basis of Surprise

There is another form of learning, often referred to as skill learning (Fitts and Posner, 1967), that consists in 'automatization', 'routinization', or 'chunking': constructing a compact representation of action sequences that can then be triggered and executed automatically, without willed effort. We refer to Luria (1966) and Norman and Shallice (1986) for a distinction between *willed* and *automatic* control of action. Basically, the difference is that while in the former case attentional/conscious resources are spent for triggering and monitoring the action execution, in the latter case they are not required. Typically, willed actions are considered to be the result of a deliberation, or related to a process of reasoning, while automatic ones are not.

If we assume that goal-oriented actions, to be executed in a willed or automatic way, could have the form 'action-effect', routinization could consist in the formation of ACTION → EFFECT, ACTION → EFFECT, ACTION → EFFECT sequences with two main peculiarities: (1) while the whole 'routine' can be executed with willed effort, the intermediate action-effect pairs are triggered automatically (each action-effect pair is triggered automatically if it belongs to a sequence and the previous element in the sequence is active); (2) while in the willed modality all the effects of actions are explicitly under attention, tested and monitored during performance, in the case of routines some of them are skipped.

What is the meaning of skipping a test? There are at least two possibilities

- tests can be performed automatically but without attention and conscious access. In this case the sequence retains the structure ACTION → EFFECT, ACTION → EFFECT, ACTION → EFFECT, but the tests on effects are performed at the automatic level and are not accessible.
- some tests are really skipped. In this case the sequence functionally changes and becomes ACTION → ACTION → ACTION → EFFECT. Although the effects can be checked only at the end of the whole sequence, there are other ways to assess the success of an action. Typically each action creates the appropriate preconditions

for the next one to be executed. If one action can not be executed, this means that the previous action has not created the appropriate preconditions.

Evidence for the first form of 'test skipping' comes from studies of the difference between novices and expert performing actions (cf., for example, Chi et al., 1981), which indicates that there are several differences in the error rates as well as in the degree of conscious effort needed. In our perspective, during learning (for example, to drive a car) internal models are usually unreliable and the error prediction signal is needed for training them (for example, press the brake pedal, expect that the car will slowly decelerate). This means that several tests are performed to assess if the predictions are correct, and some of these tests have to be executed under attentive control. After learning, internal models are so reliable that complex sequences of actions can be safely run automatically, without conscious effort, which might explain the fact that monitoring is no more under attentive control.

Evidence for the second form of 'test skipping' comes from errors in routine action performance, for example, when the agent fails to follow a 'script' (Schank and Abelson, 1977) because one of its actions has been performed but has not produced the appropriate preconditions for the next one. For example, one can unlock a door with a key, and then try to open the door, without acknowledging/testing that the door is still locked (because there are two locks, while usually there is only one). In this case, the action to unlock has been successfully executed but it has not produced the appropriate preconditions for the next action, that is opening the door, to be executed.

Overall, we might say that routinization can be safely performed when a certain pattern of actions is *reliable* enough, and the amount to prediction error in the individual actions as well as in the chaining of actions is minimized. But, if during action execution something goes wrong –a surprising event occurs, such as the car failing to stop when the brake pedal is pressed, or the door failing to open– the control of action can become again willed, since attention is needed to understand what goes wrong. This means that surprise has a crucial role in the passage between the two levels of action control, willed and automatic, in the two senses: from willed to automatic via routinization, and from automatic to willed when surprising events arise[3].

One related view on action routinization is put forward by Hommel (2003, abstract). According to his view, routinization can be done on-line when a plan is formed: "planning an action turns the cognitive system into a kind of reflex-machinery, which facilitates the proper execution of the plan under appropriate circumstances". The facilitation consists essentially in the context-sensitivity of sensorimotor structure that subserve motor action also without attentive and willed control. Although the structure of actual action execution can be described as reflex-like, the action is still planned in terms of anticipated goal states or events, and is triggered by appropriate (and anticipated) contextual conditions. This kind of action and

[3] This passage has implications for the availability of information to attention and consciousness, too, since different kinds of information could be under attention control during willed and automatic control of action. We do not discuss this issue in further detail here.

action plan thus retains the two basic requisite of goal-oriented action introduced in sec. 2.3.3, *effect anticipation* and *start anticipation* (see also Hoffmann, 2003).

2.5 Conclusions

In this chapter we have provided definitions for anticipation and anticipatory behavior, we have discussed different types and aspects of predictive and anticipatory mechanisms, we have classified anticipatory mechanisms in goal-oriented behavior, and finally we have distinguished anticipatory mechanisms in learning structures.

Overall, we have illustrated a comprehensive conceptual framework of prediction, anticipation, and anticipatory behavioral components. In the successive chapters, we provide several examples of artificial learning systems and analyze them by means of this framework. Before doing so, however, we discuss specific potential benefits of anticipations in artificial systems, providing several examples of different cognitive capabilities, individual and social, such as motor control, attention, decision making, and social cooperation.

Chapter 3
Benefits of Anticipations in Cognitive Agents

Martin V. Butz and Giovanni Pezzulo

> Man shoots an arrow into the future with a chord attached. The arrow fixes itself in an image, and he hauls himself toward it. *Paul Valery*

This book proposes the anticipatory approach for the design of cognitive agents. That is, we propose to implement anticipatory mechanisms in artificial cognitive systems to generate highly flexible and adaptive systems that can efficiently cope with dynamic environments. While the first two chapters identified several different facets of anticipatory mechanisms, this chapter focuses on the benefits and potential drawbacks of using these types of anticipatory mechanisms in cognitive systems. We focus hereby on the explicit forms of anticipation defined in the previous chapter. Thus, this chapter intends to answer questions such as: "What are the specific advantages of anticipatory mechanisms and capabilities?", or "Why should we endow artificial cognitive agents with anticipation?", as well as, "Are there any disadvantages in doing so?"

This chapter first identifies three general potential beneficial aspects of anticipations and then provides concrete examples of the impact of anticipation in several cognitive functions such as motor control, learning, attention, and social interaction. Section 3.3 points out that anticipations might not always be advantageous and may sometimes interfere with effective behavioral execution—especially when the involved predictions are inaccurate or not synchronized well. Finally, we summarize and draw conclusions.

3.1 Potentials for Anticipatory Systems

The previous chapters have pointed out that anticipation is a fundamental aspect of cognitive agents. Anticipations distinguish cognitive agents from adapted reactive systems, which do not form explicit predictions, for example, of their environment, of the effects of their actions, or of the behavior of others. Before describing the advantages of anticipation in single cognitive functions, here we provide an overview, distinguishing between three kinds of advantages:

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 45–62, 2008.
© Springer-Verlag Berlin Heidelberg 2008

- Multiple representations
- Future-oriented capabilities
- Bootstrapping complex cognitive capabilities

3.1.0.1 Multiple, Detached Representations

Anticipation enhances adaptivity of living organisms and artificial systems. While reactive agents are certainly able to adapt to some environments, there are limitations in the complexity of the behavior they can develop. In this chapter and throughout the book we provide several examples of behavioral capabilities that are enhanced thanks to anticipatory mechanisms in the fields of action selection, control, input selection, reflection, reasoning, and others. It is shown that based on the representation of future expectations, anticipatory cognitive systems are able to control themselves more effectively by means of predictive control principles (cf. Schlesinger and Barto, 1999; Camacho and Bordons, 1999). Moreover, they are able to distinguish changes caused by themselves from changes caused by others, and use that knowledge to adapt their representations, predictions, and behavior accordingly. These capabilities and others will be discussed in more detail below.

Another important feature of anticipatory systems that enables several cognitive functions is the necessary capability to represent multiple states in parallel. Systems that have implicit forms of anticipation can certainly coordinate with their environment, but it is unnecessary to represent more than one (the current) state. Anticipatory systems, by definition, need to represent the present state and at least one potential future state. The consequent capability to represent (and act on) multiple representations and to engage in mental simulation of alternative perspectives is required for several cognitive tasks. One tenet of purposive action is that a representation of at least two states, the present and the desired one, is necessary for goal-oriented behavior. It is also important that these states share the same representational format, since they need to be compared to be able to decide if the goal is already achieved or, at least, if progress is being made. Representing future alternatives explicitly (such as an upcoming event, the existence of an object, etc.) opens up the possibility to perform complex mental operations: for example, it allows the comparison of multiple options for action decision making, the comparison of available affordances in the present and in the future, the reasoning about the state of mind of other agents (theory of mind, Buckner and Carroll, 2007). All these capabilities share the prerequisite of imagining or simulating what is not here-and-now, that is, being able to detach potential future states from the current state as well as to detach other agents' states of mind from the current own state of mind.

3.1.0.2 Future-Oriented Capabilities

Some capabilities are eminently future-oriented, since they are carried out for the sake of future and not present needs. Specific mechanisms for predicting and representing the future can facilitate them. In general, the capability to conceive the future enables the selection of actions to establish future and not (only) present immediate outcomes, selecting among multiple candidate futures, coordinating one's own actions in the present and the future in order to realize intentions and goals that go beyond satisfaction of immediate needs, or producing and selecting future affordances instead of simply exploiting present ones.

Taken together, these capabilities lead to the possibility to *coordinate with the future* and not only the present—which is fundamentally different from the capability to efficiently coordinate behavior dependent on the present. This fact is of paramount importance for explaining goal-oriented behavior. By explicitly representing a goal state, an agent can coordinate behavior based on that representation, which is the idea put forward in control theory by Adams (1971), who argues that goals serve as reference signals from the future (in Chapter 2 we also discussed the fact that goal-oriented behavior needs to realize an 'inversion' of the direction of causality, that is, from the future to the past). A teleonomic structure based on an explicit representation of the future goal state can explain how an anticipatory agent coordinates its actions in the present. It may also serve to study how the agent represents time and how the inversion from goals to means may be realized. Moreover, it may be suitable to study how goals that can only be accomplished in the more distant future may influence current behavior (which seems especially relevant to explain intricate human behavior).

It should be noted that several researchers, nonetheless, argue that teleonomic systems can also be accomplished by intelligently evolved behavioral patterns. For example, on the basis of a parallel between self-organizing systems and morphogenesis, Keijzer (2001) sketches out an explanation of goal-oriented behavior that avoids explicit representations of the future. He essentially suggests that an evolved behavioral structure may implicitly encode teleonomy (an implicitly anticipatory system, as defined in Chapter 2). However, the listed benefits below should point out that although implicit anticipatory mechanisms may accomplish some of them, the adaptivity and flexibility of the potential benefits can only apply if the future is represented explicitly. That is, only actual representations of the future allow the flexible adaptation of behavior based on both the generated representations of the future and the encountered inaccuracies of the generated representations.

3.1.0.3 Bootstrapping Increasingly Complex Cognitive Capabilities

Anticipatory capabilities may not only be useful by themselves, but they may also serve as the neural basis for the formation of increasingly complex cognitive capabilities and abstract representations. Based on simple sensorimotor representations, hierarchical, abstract representations may emerge that could be highly useful for the development of grounded, symbolic representations of the environment, such as object representations or event representations (such as touching an object). Recent machine learning literature indicates that anticipation-based mechanisms that depend on estimations of current prediction error, such as curiosity and surprise mechanisms, can play an important role in the autonomous development of a repertoire of increasingly sophisticated, hierarchical representations and behaviors (Butz et al., 2004b; Oudeyer et al., 2007; Schmidhuber, 2002, 1991b; Singh et al., 2005). Overall, anticipation might play a role not only in enhancing individual cognitive functions, but also in extending a cognitive agent's capabilities to learn more complex, abstract, and symbolic concepts.

The more abstract cognitive capabilities face an important detachment problem. In Chapter 1, we have proposed that simulative theories of cognition offer a bridge between present- and future-oriented capabilities. The key mechanism is the possibility to produce expectations by endogenously reenacting the sensorimotor structures that are used for the control of action, decoupling these expectations from the online action. In this way, representations can be generated that are *grounded* (since they originate from online control and monitoring of action) but that are detached from their original online use. To be able to generate predictions that are independent of the current sensory and motor state and other current facts, the systems face a (symbol) detachment problem (Pezzulo and Castelfranchi, 2007).

3.2 Potential Benefits of Anticipatory Mechanisms on Cognitive Functions

We now illustrate the benefits of the incorporation of anticipations in several cognitive mechanisms, illustrating them with examples from natural and artificial cognitive systems. We start bottom-up from the hypothesized lowest level anticipatory processes to the highest, most abstract ones.

3.2.1 Effective, Context-Based Action Initiation

As discussed in Chapter 2, two kinds of anticipation can be distinguished in goal-oriented action selection: *effect anticipation* and *start anticipation* (Hoffmann, 2003). Effect anticipation refers to the decision of which goal to fulfill (the *end*) given several anticipated possibilities and a current motivational state that prioritizes the options. Start anticipation refers to the anticipation of the starting conditions that lead to the selection of that action (that is, the *mean*) that is most likely most effectively going to satisfy the end given the current context. Certainly, an actual action selection process, however, is not really that sequential. Goal selection must consider the current context to ensure that the goal is achievable—seeing that we usually do not come up with a current goal that is absolutely unachievable (although we can dream up such goals of course). However, the advantage of goal-based action selection is that context can be flexibly accounted for during goal selection (which goals are currently achievable considering the context) and during the suitable action selection (which is the best action or action sequence to achieve the goal considering current context-based constraints). The sensorimotor redundancy resolving architecture SURE_REACH, for example, illustrates this flexibility very effectively on the control task of a redundant arm (Butz et al., 2007a; Herbort and Butz, 2007).

An approach that combines multiple experts and consequently achieves great behavioral flexibility was introduced in Wolpert et al. (1995) and Wolpert and Kawato (1998). In those works, forward and inverse models are coupled. The coupled forward-inverse models specify alternative motor plans, which entail alternative hypotheses about the context (for example, lifting a *full glass* or an *empty glass*), and they compete for gaining motor control. The prediction of the right sensorimotor flow in a forward model is used for the evaluation of the hypotheses and the determination of the current best forward-inverse couple, whose inverse model then gains control. Combinations of forward and inverse models have been used in robotics both in distributed approaches (Tani, 2003; Tani et al., 2004) and in localist ones (Demiris and Khadhouri, 2005; Mohan and Morasso, 2006; Pezzulo and Calvi, 2006a; Tani and Nolfi, 1999; Wolpert and Kawato, 1998).

While expert selection and context-based adjustments may be based on current sensory states, they are often based also on future expected states. Psychological experiments have shown that current behavior always adjusts to both the current circumstances and their expected consequences (Fischer et al., 1997; Kunde et al., 2004). The anticipatory behavior adjustment is then usually termed *preparatory behavior*. For example, in the anticipation of a heavy load, we adjust our grip and muscle tension to still be able to hold and lift the object in question. Thus, behavior is adjusted not only based on the currently observed context, but also on predicted, relevant context.

3.2.2 Faster and Smoother Behavior Execution

Actions may not only be selected (and initiated) on the basis of anticipations of their effects, but anticipatory mechanisms (e.g., *Kalman filters*, Kalman, 1960) can also be involved in their control and filtering on the basis of a systematic comparison of expected and actually sensed inputs.

According to Adams (1971, pg. 132), anticipations can be used as a reference signal for the control of voluntary acts:

> Beginning the movement brings an anticipatory arousal of the [perceptual] trace, and the feedback from the ongoing movement is compared with it.

The availability of this perceptual trace *before* actual sensory feedback is very suitable to stabilize dynamic systems. For example, Mehta and Schaal (2002) have shown that forward models must be available to humans when balancing a pole with their finger. They conclude that the most likely model available to humans while balancing a pole must be a Kalman-filter-like system, since successful control required the availability of a forward model and even shortly interrupted sensory inputs could be compensated without any problems, which indicates a Kalman-gain based combination of sensory and forward model information. This result is also in line with state estimation experiments investigated in Wolpert et al. (1995). A review on various control tasks also supports such a model, which continuously updates the unfolding control strategy based on sensory and forward model feedback and differences between the two (Desmurget and Grafton, 2000).

Besides the filtering aspect of a Kalman filter, it has also been shown that we use forward models to cancel out self-generated stimulus aspects. *Smith predictors* were previously used for this purpose, in which a forward model mimicked the plant and canceled out predictable parts of sensory feedback. This permits the use of (only) the unpredictable parts of the feedback for correcting errors within the feedback loop (Miall and Wolpert, 1996). Although Mehta and Schaal (2002) suggest that Smith predictors cannot be applied for the control of unstable equilibrium points (such as pole balancing or also keeping the balance while standing), canceling out own sensory effects still appears necessary in various cases, as strikingly illustrated by the fact that we are not able to tickle ourselves (Blakemore et al., 1998).

In robotics, there are many examples of anticipation for the control of action. For example, Mel's (1990) robot, Murphy, can exploit efference copies of motor commands for generating simulated perceptual inputs and is thus able to maneuver its arm robustly even in the partial absence of sensory stimuli.

Besides the capability to filter sensory input or to substitute delayed or missing input, these anticipatory mechanisms also enable an easier detection of unexpected inputs. Consequently, also the processing and the resulting behavioral adjustment should be possible more quickly and more effectively. Several cognitive mechanisms come into mind when unexpected inputs occur such as surprise (and, for example, consequent quick [protective] reactions), contemplations of why the unexpected feedback occurred, or the search for reasons and consequent behavioral adjustments based on the drawn conclusions. We discuss these aspects in further detail below.

3.2.3 Improving Top-Down Attention

Attentional processes are typically classified into bottom-up attention and top-down attention. Bottom-up attention refers to attentional processes that are induced by sensory events, such as sudden changes or movements in a scene (Pashler, 1998).

Top-down attention can focus the limited epistemic resources (the bottleneck in cognitive processing) and gather (only) the information that is expected to be salient (relevant for the current tasks in mind), which is often called *selective attention*. This includes ignoring currently irrelevant stimuli and stimuli changes as well as actively searching for goal related information and actively confirming or refuting hypotheses and expectations. Top-down attention is the one that is likely to be mostly controlled anticipatorily.

A simple form of such top-down attention can be found in *priming*. Studies in different domains (see e.g., Anderson, 1983) reveal that the perception of a certain stimulus co-activates related representations. This seems to be a rather diffuse effect, which concerns perceptual, semantic, and contextual dimensions. Several authors have then attributed this effect to widespread low-level associative mechanisms in the brain that "[. . .] take advantage of frequent trends in the environment to help interpret and anticipate immediate and future events" (Bar, 2007, pg. 280). According to the terminology introduced in Chapter 2, this kind of associative-based priming is *preparatory* in nature and consists in STIMULUS → STIMULUS and not ACTION → EXPECTATION brain codes.

However, neuroimaging studies have shown that cognitive facilitation (e.g., in perceptual tasks) also depends on associative predictions (Bar, 2004). Thus, priming is not limited to simple stimulus associations but often has a predictive component as well. Linguistic experiments have shown that this predictive associative capability can be guided by various modules in the brain including language-based predictors (Griffin and Bock, 2000; Bock et al., 2003). These and various other experiments show that the capability of top-down attention leads to typical anticipatory behavioral effects that can be very task specific and are usually highly goal-directed. Hereby, the phenomenon of inattentional blindness may serve as the best example. Inattentional blindness is most clearly illustrated in the infamous "gorilla experiment" (Simons and Chabris, 1999). Asked to count the number of completed passes of a basketball within a small group of players, most participants do not detect other interesting events in the scene, such as the (slow) movement of a human in a gorilla costume passing through the basketball players.

First implementations of such attentional effects are emerging. Along these lines, Balkenius and Hulth (1999) implemented several systems in which attention is conceived as a mechanism of selection for action. Thus, attention is oriented proactively for the sake of gathering information useful for action. In a schema-based architecture presented in Pezzulo and Calvi (2006a), sensorimotor schemas for acting and gathering action-related information are coupled. Perceptual schemas orienting attention toward relevant inputs, whilst motor schemas are responsible for executing the most appropriate motor action given the sensory and motivational context. Chapter 4 discusses further architectures and other facets of anticipation involved in attentional processing.

3.2.4 Improving Information Seeking

Top-down attention is thus often guided by (associative) anticipatory mechanisms. This is also highly necessary, since we are not able to process more elaborate information in parallel, termed the "bottleneck of attention" Pashler et al. (2001). Thus, a cognitive agent needs to select amongst the available information to be able to process the selected information in more detail. Due to this bottleneck, we are destined to seek the current most useful and relevant information in our environment, that is, we are inducing epistemic activities 'querying' the environment for the sake of testing hypotheses and predictions.

Kirsh and Maglio (1994) have called this kind of information seeking activity *epistemic actions*, which consist in probing, controlling, and testing the environment for the sake of knowledge gain. Epistemic actions are performed by means of pragmatic actions, which have, however, an epistemic finality instead of the usual pragmatic one. For example, linguistic experiments have shown that given the task of telling the time from a clock, eyes scan the clock differently dependent on which language the participants use to tell the time (Bock et al., 2003). Thus, participants are searching for the time information displayed by the clock in anticipation of which aspect of the time they have to utter first, dependent on the structure of the language used. Similarly, rats were shown to use information seeking sensors, like whiskers, which probe the environment before "entering" it (e.g., floor test before putting the foot down, Mitchinson et al., 2006). This shows that top-down attention does not only bias sensory processing but also directs actions towards gaining further information.

Information seeking is generally necessary in any search scenario in which the agent perceives only partial information about its environment. In this case, it is destined to execute actions to gain further, currently hidden information (such as moving a magazine to see if the missing keys are located under it). *Active vision* research has studied this phenomenon in detail (Ballard, 1991; Thomas, 1999). It is also the most obvious case of necessary information search because visual information is maximally accurate only when perfectly in focus.

Anticipatory agents seem to be particularly well prepared to execute effective epistemic actions since they can estimate the expected information gain for several potential goals and thus choose that goal (and consequently necessary action) that is expected to yield the highest information gain. We may call such a processing method a *goal-oriented perception* (or *constructive perception*), since the information is searched in the anticipation of an effective current state perception—or construction of a state representation. Each action can be considered an *epistemic action*, since it has the epistemic implication of unraveling hidden information. Formally, this process may be described as empirical Bayes performed by hierarchical architectures (e.g., predictive coding architectures Kilner et al., 2007; Rao and Ballard, 1999), in which expectations (priors) channelize epistemic activity.

Similarly but more generally, *curious behavior* may be induced in the general search for, for example, food but also for general knowledge gain, which can also result in learning improvements. To be more precise, curiosity may lead to the selection of actions that are anticipated to improve the knowledge of the perceived world, that is, to gain information about the world. When implementing curiosity, we consequently need to design a scheme where those actions are selected that lead to an optimal information gain, that is, actions from which we expect the results to improve our own predictor or environmental representation most effectively. In this case, the most important aspect seems to trigger those actions that lead to improved predictive capabilities that are behaviorally useful (cf., Butz, 2002b; Schmidhuber, 2002, 1991a).

3.2.5 Improving Decision Making

Another relevant aspect is an investigation of the systems' capabilities of exploiting predictions for effective action decision making. Actually, the discussed epistemic actions may already be considered as one potentially beneficial influence of anticipations on action decision making. However, there are other benefits besides biasing action decision making for information seeking.

In general, planning and deliberation, as defined in several AI systems, can be considered an anticipatory mechanism. However, in the last years consensus has grown about the fact that deliberation (and rationality in general) are bounded. This means that only few options can be considered at a time and these options consequently need to be chosen heuristically. Thus, while exhaustive planning appears inapplicable, the capability of generating partial future predictions and of adjusting behavior according to these predictions can be a strong benefit for anticipatory behavioral systems. The problem is only how to select which parts of the future to explore before acting.

One way to bound anticipatory mechanisms is what was previously called *preventive anticipation* (Davidsson, 1997, 2003), that is, anticipation that only occurs if the predicted usual behavior leads to an undesired situation. In general, simulating the expected cause of events may lead to the activation of memory traces that indicate unpleasant events and thus cause the agent to prevent this event from occurring in the first place. For example, a child may stop her hand long before it reaches a fire thanks to the anticipated feeling of pain that has been previously experienced and stored as a memory trace of a bodily sensation—a *somatic marker* according to Damasio (1994). This mechanism permits the prevention of possible dangers by stopping actions whose predictions activate negatively marked bodily states.

Besides the prevention of danger, though, other motivational influences from predicted futures may positively influence behavior decision making. To accomplish this, utility information may be merged with predictions to be able to generate opportunistic behavior, shape action preparations, and thus predispose action decision making. Essentially, any motivation or emotion may be linked to certain environmental properties. The activation of these properties may then induce the activation of action patterns that lead to the activated goals and consequently bias action decision making towards executing actions that lead to positive, motivationally linked states.

In the robotics community, Shanahan (2005) has explored the possibility to generate long-term predictions related to current active courses of action in order to receive 'feedback from the future', for example, with a mechanism similar to the somatic markers. Davidsson (1997) showed that preventive anticipations can also be helpful in competitive as well as cooperative multiagent scenarios. Similarly, Hesslow (2002) suggested to run actions 'in simulation' to be able to compare their possible outcomes before really executing them, which he called simulative planning. Hesslow's (2002) *simulation hypothesis* has been recently tested in simulated and robotic settings (Stephan and Gross, 2003; Weber et al., 2006; Ziemke et al., 2005). In these studies, multiple candidate long-term plans are generated, for example, by chaining forward models (Hoffmann and Möller, 2004), and are compared in order to select the best one. Gross et al. (1999) use internal simulation of the sensory consequences of multiple possible motor actions to perform robust planning in the presence of noise. Yet another similar methodology for using simulative planning in order to select the action to perform in the real world is described in Vaughan and Zuluaga (2006).

Another interesting possibility provided by anticipatory and in particular simulative mechanisms is to work offline, that is, to decouple from the current sensorimotor cycle and simulate interesting past and potential future episodes in memory. The anticipatory representations provided by available forward models can not only be used online for the selection of action, but also for more complex functionalities such as offline planning. Possible outcomes of events can be simulated and compared offline by exploiting the same machinery involved in online visual and motor planning, but without sending commands to the effectors. During this operation the expected stimuli replace actual ones and serve as inputs for generating predictive chains. Interestingly, Butz and Hoffmann (2002) have shown that similar action decision making influences can be induced by short-term planning or by offline simulative mechanisms. Both mechanisms yielded simulated behavior in an anticipatory behavior system that was comparable to the behavior of rats in advanced conditioning experiments (Colwill and Rescorla, 1985, 1990), which showed that also rats use internal predictive models of their environment for anticipatory action decision making.

3.2.6 Object Grounding, Categorization, and Ontologies

According to constructivists like Piaget (1954) and interactivists like Bickhard (1998), objects are an autonomous discovery of cognitive agents. The anticipation and verification of object behavior can lead to the insight of object permanence and to autonomous construction of cognitive reality. Agents can interactively enlarge their ontology by learning new *synthetic items*, which are conceived as the common cause of a set of related interactions. Learning new synthetic items opens the possibility to learn more and further abstract concepts or develop conceptual ontologies.

Drescher (1991) has done pioneering work on concept formation by agent-organism interaction and its prediction. Another account of concepts formation, and the development of hidden states, is proposed by Morrison et al. (2001); here the central idea is that concepts support accurate prediction.

Roy (2005) has proposed a concept of *symbol grounding* (Harnad, 1990), which depends on two mechanisms relating agent and environment: causation (from environment to agent) and anticipation (from agent to environment). According to this idea, concepts for objects which are, for example, reachable or graspable are grounded by schemas, which regulate actual behavior and, at the same time, encode predictions of the consequences of an expected interaction. Roy et al. (2006) have used anticipatory sensorimotor representations that are able to ground conceptual knowledge such as the words 'red', 'heavy', and 'cup'. They have built the robot named Ripley, which is able to build up representations of objects based on their sensorimotor structures and to exploit (e.g. associate) those representations in order to manipulate (e.g. pick up) the objects, on the basis of verbal instructions, which are understood in sensorimotor terms. Although this anticipatory approach is very promising, it currently lacks fundamental learning and adaptive capabilities. Nonetheless, the currently observable behavior of Ripley is very impressive and based on anticipatory, sensorimotor grounded representations.

Hoffmann and Möller (2004) and Hoffmann (2007) conducted a series of robotic experiments that illustrate how internal simulation of possible trajectories can be used to ground concepts related to navigation. For example, distance from obstacles ('far', 'close') is grounded and estimated by running simulations until they encounter the obstacle. Dead-ends are recognized through simulated obstacle avoidance, while passages are grounded in successfully terminated simulations of navigation. These experiments illustrate that objects can be conceptualized as (expected) interactions with the environment by reenacting (really or in simulation) internal forward models that serve for motor control.

In the ART (adaptive resonance theory) framework (Carpenter and Grossberg, 1988), bottom-up signals and top-down expectations are used to self-organize categories and resolve the stability-plasticity dilemma. On the basis of the theory of perceptual symbols (Barsalou, 1999), Pezzulo and Calvi (2006b) show how perceptual and abstract categories can be evolved by a situated agent that interacts with its environment. Schenck and Möller (2006) use visuomotor anticipation for navigating visual scenes: objects and shapes are recognized by means of their compliance to certain sensorimotor transformation produced by the movements of the camera.

Finally, anticipation can be used for distinguishing self-produced motion from sensory stimuli, which are caused by interacting with objects in the environment. Developmental psychologists have shown that children at early stages fail to distinguish events that are under their control from events that are not. They accomplish this task only after the development of a full fledged *body schema*: it is only by understanding the boundaries of one's own predictive capabilities that it is possible to discriminate self from others and from the environment (Maravita et al., 2003; Piaget, 1954). The capability of distinguishing self from others and the environment is also very important to enable efficient (anticipatory) social interactions, which are further discussed below.

Robotic studies (Bongard et al., 2006a,b) illustrate the evolution of body schemas by means of interactions and their anticipation. In a robotic experiment, Mohan and Morasso (2006) illustrate how body schemas can be extended (for example, by holding a stick in the hand) and adapted to new situations by adapting its predictions.

3.2.7 Social Abilities

While the aforementioned benefits focused on individual behavioral aspects, anticipations appear to play also a highly important role in the development of efficient social interaction, such as imitation, perspective taking, joint attention, intentional action understanding, and even language. In the following paragraphs, we distinguish between simpler forms of social interaction, in which we include imitation, perspective taking, and joint attention, and more complex forms of interaction, including trade and language.

Imitative Social Interaction Social capabilities, such as imitation, perspective taking, and joint attention are certainly crucial for the development of efficient social interaction and are supposed to rely on a neural substrate, which is also involved in other future-oriented activities. Meltzoff and Moore (1997) suggest that children learn to imitate in four phases: (i) motor babbling, and the formation of a body schema, (ii) understanding and imitation of body movements, (iii) understanding and imitation of actions on objects, and (iv) understanding and imitation of intentions.

Many computational studies exist on these themes (Breazeal and Scassellati, 1999; Dautenhahn and Nehaniv, 2002; Demiris and Hayes, 1996; Kaplan and

Hafner, 2006; Oztop et al., 2005; Scassellati, 1999) and in many cases common anticipatory mechanisms have been developed for the control of action, imitation, and attention. For example, Johnson and Demiris (2005a) have developed Hammer, a system which uses several coupled inverse and forward models for generating actions and for imitating actions performed by others. In this case, the perceptual input is firstly processed by the forward models. The forward model that currently yields most accurate predictions activates its related inverse model, which is thus able to generate a comparable behavior (that is, a behavior which would have produced the same perceptual input).

Johnson and Demiris (2005b) demonstrate how to implement perspective taking by simulating the other's perspective. The ingredients are similar to those described in the previous system, but an allocentric map of the scene is added. By simulating the other point of view, the system is able to take its perspective and know, for example, that while from its own perspective two objects are placed one in front of the other, from the other point of view they are side by side. Demiris and Hayes (1996) and Demiris and Khadhouri (2005) have studied intentional action performance and recognition as well as imitation in real robots, with a specific reference to the mirror neuron system.

The robot Ripley (Roy et al., 2006), using a similar simulation mechanism that is realized in real time, runs an internal abstracted simulation of the scene, is able to understand how objects appear from its own as well as from another's perspective, and to fulfill requests such as "touch the one on *my* left" or "touch the one on *your* left" by building up abstract scene representations. The running internal model of Ripley is also used for maintaining the state of known objects in the environment in its memory. The approach is very similar to the typical small-scale models in cognitive science apart from the fact that objects are understood by the robot through anticipatory schemas describing (expected) sensorimotor consequences of possible interactions and are not directly encoded as a collection of attributes or properties.

Complex Social Interaction The so-far discussed interactions are rather simple forms of social interaction that require the derivation of the simple immediate current intentions of another individual. Nonetheless, they all already require the capability to distinguish self from other, which seems only possible by anticipatory mechanisms, as discussed above. Certainly, though, anticipations are also an essential ingredient to accomplish more complex social interactions. Such more complex interactions are necessary, for example, for the development of more complex forms of communication including language (Arbib, 2005) and more sophisticated trading interactions (Gomez et al., 2007).

While simple forms of coordination, cooperation, and competition may be possible without any explicit anticipation of longer term effects, it has been recently shown in the neuroscientific literature that resonant neural structures, such as mirror neuron systems, are deeply involved in the creation of a social space, in which one's own and other's actions can be recognized and interpreted intentionally and separately (cf. Arbib, 2002; Gallese, 2001). Anticipatory mechanisms enable the recognition and understanding of other individuals and, essentially, also enable the derivation of their intentions and consequently necessary re-actions (such as, flee or fight when aggressive or hostile, interact or trade when friendly).

Another capability that requires anticipation in the social domain, specially in the philosophical literature, is *mind reading* and the *theory of mind*. In Chapter 2 we have discussed how taking the *intentional stance* (Dennett, 1987) permits the generation of predictions about other's behavior on the basis of one's own knowledge about other's beliefs and goals and not simply on the basis of the observation of their current actions. Cooperation on common tasks (e.g. hunt) is significantly enhanced by mutual understanding and anticipation. More complex forms of cooperation, such as trade, also depend not only on the anticipation of immediate actions. For example, the recognition of another individual and the memory of past interactions with that individual enable the expectation of the other individual's future behavior, which allows more complicated ('fair') trade interactions, building a trust model.

Several researchers have discussed how anticipations may permit the development of language (Arbib, 2005; Gardenfors, 2004; Gardenfors and Orvath, 2005; Swarup and Gasser, 2007). In relation to trade, the hypothesis is that trade can extend to the trade of (more abstract) information, such as the location of water, prey, hiding places, or enemies. Moreover, anticipation enables the further teaching of individuals, not only through better learning by imitation but also through more abstract learning and student-teacher like interactions.

As of now, no computational system exists that is able to pair the complexity of most social abilities described above. Most work in social robotics focuses on simpler forms of collective behavior, for instance those exhibited by social insects—some examples of which can be found in evolutionary robotics (Nolfi and Floreano, 2000) and swarm intelligence (Kennedy and Eberhart, 2001). Anticipatory mechanisms have received little attention in the realization of these forms of coordination and cooperation: the emphasis is on embodiment, dynamic agent-environment coupling, and self-organization. It remains to be demonstrated how these methodologies can scale up to more complex forms of social life that exist in nature—for which we expect that the usage of technologies based on anticipatory representations and mechanisms will be inevitable.

3.2.8 Learning

Besides the various anticipatory influences on behavior, also learning may very well be influenced. Anticipations may improve behavioral learning as well as the learning of (further) predictive representations. We distinguish the following potential influences:

- indirect learning influences,
- direct model learning influences, and
- direct behavior learning influences.

In the first case, predictions may shape internal learning structures in certain ways that may be advantageous for future learning. In the second case, model learning may be improved by focusing learning, ignoring irrelevant or unpredictable inputs, delegating learning responsibilities more effectively, estimating feedback reliability, and biasing learning to explain unexpected events. Finally, the predictive model may be used directly to structure behavioral patterns offline or use it as a planning mechanism. Witkowski (2003) provides a review of the anticipatory and predictive elements inherent in the most known forms of learning.

Indirect Learning Influences While indirect learning influences may occur and are expected to suitably organize hierarchical predictive and representational structures, and control programs in particular, to the best of our knowledge, concrete examples of such effects are sparse. An example of implicit anticipatory shaping can be found in the robot Alvinn (Pomerleau, 1989), for which it was shown that adding a predictive component to the control structure, which consequently co-structures internal network representations, improves the supervised learning of a control task.

Model Learning Influences Schmidhuber (2002) discusses possible strategies to be adopted by artificial systems in order to focus learning efforts on the *predictable* components of spatio-temporal events. The capacity to predict and to discriminate between predictable and unpredictable elements in the environment is also fundamental for meta learning mechanisms such as curiosity (Oudeyer et al., 2005; Schmidhuber, 1991c).

Anticipation can also be used for learning 'hidden structures' in data series. Elman (1990) firstly showed that a connectionist network can learn the formal structure of language, its syntax and semantics, by learning to predict the next word, without any innate knowledge. A similar methodology has been applied to a series of stimuli sensed by robots, which also have a rich structure. As shown by Pierce and Kuipers (1997), by learning to anticipate the effects of its actions, a robot can learn a rich hierarchical model of the robot's sensory and motor apparatus, such as position and type of its sensors and the degrees of freedom of its effectors. This knowledge can bootstrap the representation of own body schema.

Other approaches that learn predictions to further guide learning describe a hierarchical RNN system in which novelties (for example, a corner at the end of a corridor) are detected thanks to predictions generated at multiple levels (Nolfi and Tani, 1999; Schmidhuber, 1992a). König and Krüger (2006) illustrate how discretization and symbol formation may depend on anticipation-based feature selection and data compression. Further methods may be found in the most recent post-workshop proceedings on anticipatory behavior (Butz et al., 2007b).

Attention can also have highly important learning influences and thus also, in particular, anticipatory top-down attention. Attention has been shown to be mandatory to enable more elaborate and detailed learning. Similarly, the lack of predictive capabilities disables filtering and thus better learning. In the mean time, attention enables and facilitates the understanding of relevant environmental changes and particularly unexpected changes.

Expected and unexpected changes seem also crucial for learning and autonomous development. Unexpected changes in different dimensions should be associated since they are most likely correlated for the sake of learning novelties. Surprises, or the detection of unexpected events, should lead to curious behavior and consequent epistemic actions, which in turn can lead to the detection and learning of new correlations and interdependencies in the environment.

Faster Learning by Internal Simulation and Hypothetical Thinking Not only anticipatory and imaginative capabilities can be used for action selection and control, but also for learning from 'imaginary' self-generated experiences. Tolman (1932) indicates *vicarious trial and error*, that is, learning as if the experience had really happened, as an essential trait of purposive behavior. Jordan and Rumelhart (1992) have explored these ideas with computational systems. Computational systems such as DYNA (Sutton, 1990), XACS (Butz and Goldberg, 2003), and NN-based internal RL learning (e.g. Baldassarre, 2002b, 2003) indicate that learning by internal simulation is faster than to execute all in the environment. Bakker et al. (2006) have realized a reinforcement learning vision-based robot that learns to build a simple model of the world and itself. To figure out how to achieve rewards in the real world, it performs numerous 'mental' experiments using the adaptive world model. Butz and Hoffmann (2002) have shown that internal simulations or online behavioral predictions are necessary to simulate the behavior of rats. In the case of internal simulations, internal motivationally-based reward triggers were necessary to allow effective behavioral adaptation.

In a sense, all learning includes some form of prediction. Reinforcement learning, for example, can be used to learn a predictive model of the world dynamics and calculate expected reward. In their seminal paper, Sutton and Barto (1981b) introduced a framework for *temporal difference learning*, which was successively developed (Sutton, 1988; Doya, 1996). However, most forms of learning, including several reinforcement learning techniques, include only implicit forms of anticipation. More recently, the RL community has introduced the concept of *predictive state representation* (Littman et al., 2001; Wolfe and Singh, 2006), which has been shown to be effective in learning (Wolfe et al., 2005) and planning (James et al., 2004). Predictive state representations are also similar to Observable Operator Models (Jaeger, 2000, 2004). Both approaches represent states based on currently valid predictions.

Model-based RL algorithms (Sutton and Barto, 1998) use explicit predictive models of their environment enabling the adjustments of behavior offline as well as online during environmental interaction. The most prominent architecture of that type is the Dyna-Q system (Sutton, 1991), which originally used a tabular representation of the encountered environment to be able to adjust reinforcement values more effectively. While Dyna-Q was not able to generalize over states and actions, the *anticipatory classifier system* autonomously learns and generalizes situation-action-effect rules (Stolzmann, 1998; Butz, 2002a) and can use those for effective behavioral adjustments—online or offline (Butz and Hoffmann, 2002).

3.3 Arising Challenges Due to Anticipations and Avoiding Them

Despite the potential advantages of anticipations, possible disadvantages and novel challenges due to anticipatory processes should not be neglected, either. Certainly, the generation of predictions always comes at a cost. If prediction is highly inaccurate or even impossible, it is most likely better not to bother and rather to act reactively based on the current knowledge of state. In the following, we discuss various possible disadvantages mainly stemming either from cases in which anticipations are not helpful so that it is disadvantageous to generate (useless) predictions, or from cases in which predictions are not sufficiently accurate and consequently misleading in the resulting anticipatory process.

One aspect of a potential disadvantage of anticipatory processes is the fact that automatic control processes can be much more effective than anticipatory control processes. That is, blindly executing a certain motor command sequence can be expected to be much faster than an effect-guided execution—simply due to the necessity to process and combine more sources of information. Thus, when speed is a crucial factor, automation seems necessary and, anticipatory behavior may become increasingly automatized by means of behavioral training.

Early AI research has been criticized for the use of internal models for the selection and guidance of actions as well as typically very symbol-oriented representations. Building a complete model of the environment can be very costly and even completely impractical. Using a model of the environment for decision making by means of planning techniques can lead to an exponential scale-up and consequently is often intractable. Thus, anticipation in the sense of planning requires useful model representations that focus on behaviorally useful aspects of the environment. These aspects may then be modularly and hierarchically exploited to induce maximally effective action decision making. Plain first-order logic based reasoning or symbolic planning, on the other hand, are bound to be ineffective in natural environments.

Given a suitable predictive model representation of the environment, anticipatory attention mechanisms need to be employed for effective sensory processing and knowledge inference, as discussed above. However, even if this process is maximally effective in some sense, it might prove ineffective in other tasks—as exemplified in the overlooked gorilla when counting basketball passes (Simons and Chabris, 1999). Thus, attention and anticipatory focus control can be advantageous to re-act more quickly and more effectively to relevant incoming stimuli, but it can also cause distractions so that important information may be misinterpreted or totally overlooked. Such attentional errors can have negative impacts on immediate behavior, action decision making, reasoning, and learning.

Also, anticipatory mechanisms may cause unwanted interferences when the predictive information is inaccurate or ill-timed. Especially in control problems it is highly important that timely information is available. Given predictive information is delayed or inaccurate, control processes might undergo interferences due to the anticipatory information source—actually a counter-effect of why the anticipatory information was incorporated in the first place, that is, to stabilize control.

Such unwanted interferences are certainly not restricted to causing instabilities in immediate control processes. Also action decisions may be strongly and negatively influenced when the predictive model is inaccurate. Just imagine the decision to still cross the road with a car approaching and underestimating the car's speed. Since our predictive models are certainly not perfect, it may come as a surprise that we still cross the street with cars approaching (in the distance). As has been argued by many researchers, the brain is a Bayesian processing systems that also represents the certainty of its predictions (Deneve and Pouget, 2004; Knill and Pouget, 2004), which can help to prevent such decision errors and may prevent unwanted interferences also in control processes. Bayesian information processing also underlies, essentially, the idea of Kalman filtering, which integrates information based on its estimated reliability (Kalman, 1960).

Such anticipatory errors can also extend into the social sphere. If we misinterpret the behavior of others as a (not necessarily physical) threat, we can sometimes react absolutely inappropriately being unnecessarily mean or excessively friendly. On an immediate cooperative level one might remember an occasion where one person thought the other person would hold the glass (still, or by now), realizing the mistake only when the glass already smashes on the floor. Also on a trading level issues like treachery seem to be an effect of ill-guided predictions. A person may pretend to be your friend making you believe in a certain fact and later on may use your false believe, effectively "pulling the wool over your eyes".

Thus, while anticipations may be advantageous in several respects, it should be clear that not all processes are best accomplished with anticipations involved. Once things have become automatized, anticipatory execution of certain actions may be much less effective and much slower than the execution of the (learned) automatic process. Planning processes can only be employed when effective representations are available. Anticipatory attention needs to be employed with care to avoid the ignorance to certain important environmental facts. Similarly, other anticipatory interferences may need to be prevented possibly employing rather pessimistic reliability estimates of predictive information sources. Finally, also during social interaction anticipations need to be employed with care to avoid misunderstandings, ineffective cooperation, or being fooled.

3.4 Conclusion

This chapter has pointed out several advantages but also potential disadvantages of anticipations in cognitive systems. Generally, we distinguished between advantages stemming from the possibility to form and maintain multiple representations of states and thus potential realities, the capability to act goal-oriented in various senses, and the potential to bootstrap complex cognitive capabilities based on simpler, predictive representations. The necessary capability to represent multiple realities is a requirement to be increasingly anticipatory and also to be able to distinguish self from other during social interaction. Future oriented capabilities are those that make the cognitive system better prepared for future events due to (limited) knowledge about the future. Moreover, better adaptivity is enabled due to better detection and processing of unexpected events, the enabled distinction between self and other, and the possible substitution of delayed feedback information. Finally, anticipatory representations appear to be the useful basis for the representation and for the development of increasingly complex cognitive processes.

The potential of such enhanced cognitive capabilities was then discussed and detailed. We started from very immediate potential benefits of anticipation including immediate improvements on action initiation, flexibly adjusting to context and state information, as well as benefits for action control, such as predictive control and sensory filtering. Next, we highlighted potential positive influences on attentional processing and related information seeking behavior. This lead us to potential advantages in actual action decision making and goal selection. While these advantages focused on individual decision making, we then also noted that anticipations are highly important to accomplish simple and complex forms of social interaction. Last but not least, we noted that anticipations can also help shaping and efficiently learning predictive structures themselves.

Besides all the potential advantages, we also pointed out that anticipatory mechanisms may have unintended effects, such as very slow processing, instabilities in control processes, misguided attention, inappropriate decision making, or destructive social interaction. These stem mainly either from the cost of generating predictions or from unexpected inaccuracies of the generated predictions. Despite such potential drawbacks, it should be remembered that in most cases the disadvantages emerge only out of the additional capabilities due to the involved anticipatory process. Thus, while drawbacks may emerge, they are expected to be minor compared to the system capabilities gained due to the involvement of anticipatory mechanisms.

In the subsequent chapters, we proceed with various case studies of artificial, partially anticipatory systems and their potential with respect to the different anticipatory mechanisms identified. These studies are intended not only to highlight certain system capabilities but also to provide a first step toward the design of modular, hierarchical, and highly interactive cognitive system architectures, where each module is realized with the maximally suitable representation, learning, and behavior approaches available.

Part II
Models, Architectures, and Applications

Chapter 4
Anticipation in Attention

Christian Balkenius, Alexander Förster, Birger Johansson, and Vin Thorsteinsdottir

4.1 Introduction

Although attention can be purely reactive, like when we react to an unexpected event, in most cases, attention is under deliberate control anticipating events in the world. Directing attention and preparing for action takes time, and it is thus useful to be able to predict where an important event will occur in the environment and direct attention to it even before it happens. Another reason for the need for anticipation is the processing delays in the visuomotor system. In the human system it takes at least 100 ms to detect a visual target (Lamme and Roelfsema, 2000) and to just look at a moving object, we need to anticipate its movement to control the muscles of the eyes to move our gaze to the location where the target will be (von Hofsten and Rosander, 1997).

The role of anticipation in attention can also be seen in the close connection between attention and action (Balkenius, 2000):

Attention-as-Action Shifting the attention and gaze from one object to another can be considered an action (Posner, 1980). The principle of attention as action implies that attention can be controlled in a way similar to actions and it thus becomes natural to use similar learning methods for attention as for action. Just as it is necessary control action with anticipatory methods, is it also necessary to anticipate the location of objects that will become the target of attention. The reason for this may be that the object (1) is not yet visible, (2) is not in the visual field, or (3) is moving.

Selection-for-Action The object in the focus of attention, and gaze is used to set the parameters for motor actions (Allport, 1990; Balkenius and Hulth, 1999). For example, to reach for an object, we first direct our attention toward that object. Its position in space is used to control the movements of the arm while the form of the object is used to preshape the hand in aniticipation of the grasp (Castiello et al., 1992; Castiello, 1999).

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 65–83, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Deictic Reference The focus of attention is used as an implicit argument for the action. The focus of attention refers to the fixated object without explicitly representing all of its properties (Agre and Chapman, 1987; Ballard et al., 1997). Instead, the attentional system sets up a sensory-motor transformation that controls action. Since this transformation must be in place before the start of the action, this implies that attention must anticipate action. This implies that fixation of attention can be seen as an epistemic action that collects information for the next action.

Attention-as-Inhibition Attention involves inhibition of distractor objects and previous targets as much as activation of new targets (Anderson and Spellman, 1995). A learning attention system must learn not only what to look at, but also to ignore stimuli that do not predict any rewards.

In this chapter, we will describe a number of areas of attention control that require anticipation. The different areas all include learning as an important component. The simplest form of anticipation in attention involves learning *what* to look at to receive a future reward. The learning, in this case resembles classical conditioning in that a stimulus itself predicts a reward. Alternatively, a particular stimulus may predict *where* a rewarding stimulus will occurs. The required learning is this case is instrumental learning and may be implemented using reinforcement learning algorithms. The anticipatory process is typically called visual cueing. A spatial prediction may also include a temporal component that anticipates *when* a target stimulus will appear or how it will move. Such predictions are necessary for successful smooth pursuit. Finally, any of the above predictions must be influenced by the current *context* or *task*.

4.2 Learning What to Look at

One area where anticipation plays a role in attention is the prediction of which visual target will be rewarded. To attend to such targets it is possible to use a saliency map that suggests the location of visual targets that predict a reward. This results in a bottom-up system for attention selection since the flow of information goes from the visual image to the final selection step.

A saliency map combines a number of visual feature maps into a combined map that assigns a saliency to every location in the visual scene (Itti et al., 1998; Itti and Koch, 2001a). Each feature map is typically the result of applying some simple visual operator to the input image. For example, a feature map could consist of an activity pattern that indicates all the vertical edges in an image. Other types of feature maps may code for intensity, color, motion or some other visual feature. The result of summing the different feature maps is that the saliency map will code for locations in the image with many features. For example, a region with many edges and bright colors will be more salient than a location without any such features.

The necessary feature extraction resembles processing in the visual area V1 where cells react to oriented visual contrast or to specific colors. In technical applications, this processing is implemented as two-dimensional convolution of the input image with different filter kernels that correspond to the receptive fields of area V1.

Although not intended as a model of attention, different methods for extraction of interest points (key points) in images play a similar role to such filters. The Harris detector (Harris and Stephens, 1988) is often used in image processing to quickly find regions of interest in an image. Subsequent processing is performed only in these regions. An alternative method which is invariant to scale and rotation was described by Lowe (2004). It uses differences of gaussians at different scales to extract extrema points which are then tested for stability. These type of methods are good candidates for technical applications where real-time performance is required. It is also possible to use even simpler filters to extract oriented contrast at a single scale in many cases (e.g. Balkenius, 1998; Kopp, 2003).

In principle, it would also be possible to also have feature maps that code for more complex aspect of the scene. For example, Balkenius (2000) suggested that a sensory buffer with much the same role as the saliency map could consist of a visual code at many different levels ranging from individual features to objects.

An early system for directing attention by bottom-up means used a generalized symmetry operator (Reisfeld et al., 1995). It was demonstrated that the operator would find regions in images that naturally attracts attention such as faces, eyes and other symmetrical objects. The model demonstrates that more complex properties such as symmetry can be useful in directing bottom-up attention. It is possible that a bottom-up attention system would be more useful if this type of features could also be detected at a lower level.

Since the concept of a saliency map was introduced, it has been incorporated in a large number of models and theories (Itti and Koch, 2001a; Vijayakumar et al., 2001; Rao et al., 2002; Balkenius et al., 2004; Hoffman et al., 2006; Singh et al., 2006; Walther and Koch, 2006; Chen and Kaneko, 2007; Shi and Yang, 2007; Siagian and Itti, 2007). Saliency maps have shown to be useful both as models of human attention and for technical applications.

4.2.1 A Learning Saliency Map

We will here formulate a general saliency map framework that allows a saliency map to learn what visual features predict reward. The model is based on the idea that it is possible to weigh the different feature maps differently depending on what target the system is looking for. To look for vertical lines, the feature map that detects such lines would have a larger gain than maps for other features. Here we will use a formulation of a saliency map $S(x,y)$ as a linear combination of a number of feature maps F_m convolved with a smoothing function G (Balkenius et al., 2004):

$$S(x,y) = G(x,y) * \sum_m \theta_m F_m(x,y) \tag{4.1}$$

The feature maps F_m can be the result of simple visual operators such as line and edge detectors but can also be generated by more complex image processing algorithms as described above. The gain values θ_m describe the weighing of each feature map. The smoothing function G is typically a gaussian or a box filter. This formulation differs from alternative models that may only use the linear combination without the smoothing (e.g Itti and Koch, 2001b; Navalpakkam and Itti, 2007) and has the advantage that it is less sensitive to the exact location of the target features.

After calculating the saliency map it is necessary to select the next location to attend to. One possibility is to do this deterministically by selecting locations in sequence according to their salience (Itti et al., 1998). Alternatively, selection can be based on a probability density function defined over the salience map. For example, the location can be selected using the Boltzmann distribution where a temperature parameter that determines how random the selection should be. When the location has been selected in the image, a gradient ascend is performed on S to find the closest local maximum in the saliency map. This partitions the position space into a finite number of regions, each corresponding to a local maximum of S. Although not strictly necessary, this makes the selected locations more stable which improves visual processing in subsequent steps. Alternatively, an ε-greedy method can be used where the maximum location is selected except at exploratory trials, which occur with probability $1-\varepsilon$ (Sutton and Barto, 1998). At an exploratory trial, one of the possible fixation points may be selected with equal probability.

These selection methods can be seen as computational simplifications of the biased competition that has been found in the nervous system (Desimone, 1998; Desimone and Duncan, 1995). It is also possible to implement this form of selection using the interaction of competition and cooperation in a field of artificial neurons (Erlhagen and Schöner, 2002). When a target has been selected it results either in a covert shift of attention or in an overt saccade movement that reorientation the eyes from one location to another.

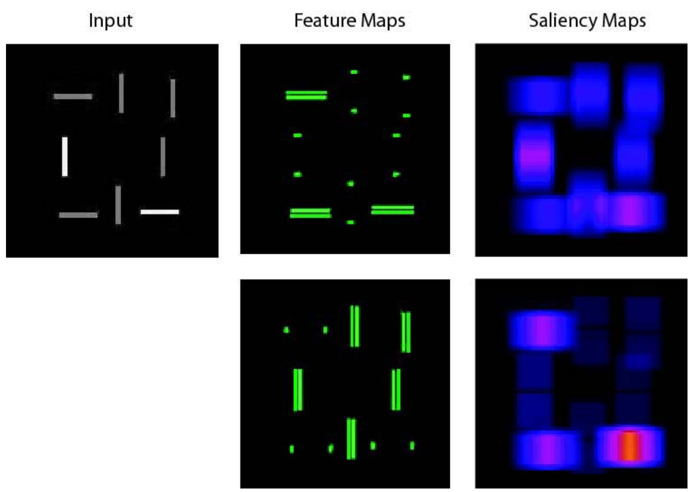


Fig. 4.1 The saliency maps before learning (top) and after learning (bottom)

Fig 4.1 shows the saliency map for a simple scene based on this model. Two simple feature maps are used to detect horizontal and vertical lines in the image and the saliency is a smoothed sum of these two feature maps using a box filter. In most application of saliency maps a much larger set of feature maps are used for a larger set of visual features. It is also common to have feature maps that operate at different scales.

Although most models have not considered learning in the saliency map, it is easy to see that by setting the gain parameters appropriately, it becomes possible to tune the saliency map to different features and several models have used various models to calculate the gain parameters directly (cf. Lee et al., 1999; Vickery et al., 2005; Martinez-Trujillo and Treue, 2004; Treue and Martinez-Trujillo, 1999; Navalpakkam and Itti, 2007).

Here we will describe a novel approach based on reinforcement learning. The saliency map S can be seen as an approximation of a value function for reinforcement learning. $S(x,y)$ is thus an estimation of the reinforcement that will be received if location $\langle x,y \rangle$ is attended. Unlike the standard action-value function in reinforcement learning (Sutton and Barto, 1998), there is no state in this formulation. Instead, each location in the image corresponds to an individual action that directs attention to that location.

At each time step, the error in the value function is calculated as

$$\delta_t = [r_t(x',y') - S_{t-\tau}(x_{t-\tau}, y_{t-\tau})] \tag{4.2}$$

where τ is the delay between the fixation of a stimulus and the time when the corresponding reinforcement is received. This delay is necessary since there will typically be a substantial delay between the time when the salience map selects a particular location and the time when reinforcement is received which may only occur after a slow object recognition phase. This gives the model a predictive component.

The gain coefficients θ_m are updated to reflect the actual reinforcement received and the value at the attended location using the learning rule

$$\theta_{m,t+1} = \theta_{m,t} + \alpha \delta_t F_m(x_{t-\tau}, y_{t-\tau}). \tag{4.3}$$

A result of this learning rule is that gains corresponding to features that are included in the target stimulus will increase in size while features that are part of the distractors will become less salient. This makes the resulting gain vector θ sensitive to both the features of the target and the environment. This is more efficient than simply using the parameters of the target to set the weights, and also more in line with human performance (Navalpakkam and Itti, 2007). It differs from approaches that only take the target features into account (Lee et al., 1999; Vickery et al., 2005; Martinez-Trujillo and Treue, 2004; Treue and Martinez-Trujillo, 1999).

This form of learning is generally referred to as classical conditioning, which is used by animals to learn a *predictive relationship* between a neutral stimulus and a subsequent reinforcing stimulus (Dayan et al., 2000a). More specifically, the training of the gain parameters implements discrimination learning, where the model learns to discriminate between the target and the distractors. The equations 4.2 and

4.3 are identical to the Rescorla-Wagner model of classical conditioning except for the addition of the temporal delay (Rescorla and Wagner, 1972). It should be noted that many other learning rules are also possible and will lead to similar results (see Balkenius and Morén, 1998).

The saliency map at the bottom in fig 4.1 illustrates how the saliency map has changed as the result of reinforcing attention to the horizontal lines but not to the vertical ones. The saliency is now much stronger for the reinforces stimuli than for the non-reinforced distractors.

The learning saliency map described here is the first step toward a learning attention system. Although it is able to speed up visual search considerably compared to a total search of the scene (Balkenius et al., 2004), it only uses immediate visual information that may be available in a scene. It is clear that a target is more easily found if we also take into account spatial and temporal relations in the scene and this is the topic of the next two sections.

4.3 Cue-Target Learning

Cue-target learning in humans is most often studied in a variant of the Posner task (Posner, 1980). In this task, a cue first appears on a screen and is followed by the presentation of the target stimulus. A cueing effect is considered to be found when the time needed to find the target is lower after the presentation of a valid cue.

In *direct cueing*, the cue is located at (or near) the location where the target will appear while in *symbolic cueing*, the cue typically appear at the center of the screen and its identity indicates the location where the target will appear (Solomon, 2004). For example, an arrow pointing to the right could be used as a symbolic cue that indicates that the target will appear to the right. It is also possible for the whole visual scene to act as a cue as is the case in *contextual cueing* (Chun and Jiang, 1998).

4.3.1 Cueing by a Single Stimulus

Several models have used reinforcement learning as a basis for cue-target learning. Schmidhuber and Huber (1991) built an artificial fovea controlled by an adaptive neural controller. The fovea had high resolution in the center and low resolution in the periphery. Without a teacher, it learned trajectories causing the fovea to find targets in simple visual scenes, and to track moving targets. The controller used an adaptive input predictor (a limited kind of world model) to optimize its action sequences. The only goal information was the shape of the target - the desired final input. Since this reinforcement learning task is of the ‘reward-only-at-goal’ type, it involves a complex spatio-temporal credit assignment problem. The latter was solved using a recurrent network training algorithm.

Q-learning was used in the model of Goncalves et al. (1999) to control attention based on multimodal input and reinforcement signals. The model includes subsystems for long term memory, what and where processing and attention control. An interesting feature of the model is that it integrates attention and action in a straight forward way.

Another model that uses reinforcement learning to control visual attention was described by Minut and Mahadevan (2001). The Q-learning algorithm is used to select the target locations in the image. Reinforcement learning was also used by Shibata et al. (1995) to control the movement of a visual sensor over an image. The goal of the system was to find the optimal location of the sensor for object recognition. The same neural network was used both for object recognition and to produce the sensory motion output.

Here, we will use Q-learning as the basis for the description of a cue-target learning system. Let us assume that we already have a bottom-up attention system of the kind described in the previous section. That system will produce a sequence of fixations in the visual scene and the idea is to use these actions as training for a reinforcement learning system.

The current state s for the reinforcement learning system consists of a description of the visual stimulus at the focus of attention and may include anything from a simple set of primitive features to a complex object description. Through learning, this state is associated with a number of actions a by learning the function Q(s, a) which predicts the reinforcement that will be received if action a is taken in state s. Like the stimulus coding, these actions can be at different levels of complexity and here we will only consider a few basic cases.

Absolute Attention Shifts When the system is trained with absolute attention shifts, each action represents a movement of attention to an absolute position in the visual scene. This means that no matter where the visual cue appears, it references a specific location. An example would be a verbal cue like "BOTTOM" which could indicate that a rewarding stimulus will appear in the lower part of the scene. Like in the saliency map, each location in the scene corresponds to a specific action. There will thus be as many actions as there are potential target locations.

Relative Attention Shifts In most cases, it makes more sense to use relative actions. These are coded in relation to the current fixation point. An example would be an arrow pointing down. The predicted target location in this case would depend on the placement of the arrow. If it is to the left, it would indicate a target down to the left, while the same cue to the right in the scene would indicate another target location.

Timed Attention Shifts It is also possible to have different shift actions for different time lags between the cue and the target. For example, if we have divided time into discrete steps, a single action a is replaced by a whole set of action $a_{t+1}, a_{t+2}, \dots a_{t+m}$, where m is the horizon of the possible predictions. To parallel the behavior of humans, the system can be trained to perform the attention shift just before the target is expected to appear. This will results in anticipatory saccades.

It is clear that the above notion of attention as actions will result in an enormous number of potential actions and the normal training methods for reinforcement learning would not be tractable. The solution is to use the bottom-up saliency map as a training system. Since it produces a small set of states and actions to train on, it immediately makes this framework viable. Furthermore, if the trained target locations for an action are selected from the finite set defined by the gradient ascend process in the saliency map, the space is further reduced.

4.3.2 Contextual Cueing

The visual context can be used both to predict the location of a target object and for disambiguation if the visual pattern of the target does not contain sufficient information for identification. Torralba (2003) shows an example of a blurred visual scene where it is clear from the context that the blob at the street is a car although it is impossible to identity the blob out of context. This is a striking demonstration of the importance of context in visual processing and a problem that often occurs in artificial vision systems.

Siagian and Itti (2007) use low level gist (one glance) information to improve the result of a saliency competition by reducing noise. The model uses background features more than foreground to localize features retrieved by saliency competition. In this way it was able to categorize different scenes with good accuracy. This model is intended to cooperate with a model of selective attention using saliency maps.

Balkenius and Morén (2000) developed a model that differs from the others in the respect that the context representation is assumed to be built up over time as the visual scene is scanned. This model has been successfully applied to contextual cueing (Balkenius, 2003) and habituation (Balkenius, 2000).

The model of Pomerleau (1994) also uses a form of contextual cueing. Saliency maps are created using expectation of future input. A hidden layer in a neural network, trained to perform a temporal sequential task, is used to predict what the next inputs will be. The expectation of what the features will be in the next frame determines which portions of the next visual scene are attended to.

4.3.3 Fovea Based Solution

The fovea based solution uses a multiresolution fovea module in combination with a long short-term memory (LSTM) recurrent neural network (Hochreiter and Schmidhuber, 1997). The fovea module divides the given image into parts of different resolutions and different fields of view (see Fig 4.2), including a central fovea.

Fig 4.2Composition of a fovea image by dividing the original image in into different sub-images and reduce the resolution with respect to the position of the sub-image. This data is given as input to the controller. The composed image in the right side is only given for illustration.

Beside the biological inspiration, the fovea model has another advantage: It highly reduces the amount of data without losing much information, because the "interesting" part of the image is in the center of the fovea where the resolution is

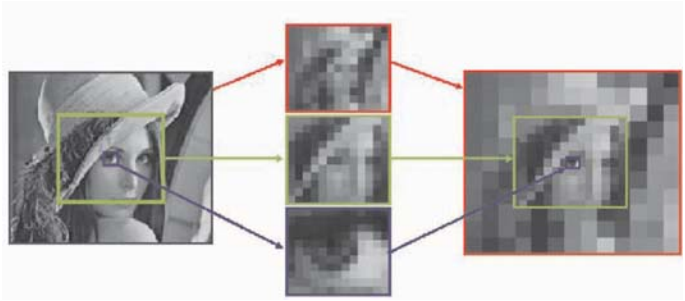


Fig. 4.2 Composition of a fovea image by dividing the original image into different sub-images and reducing the resolution with respect to the position of the sub-image. This data is given as input to the controller. The composed image in the right side is only given for illustration.

very high. The drawback of this model is that the full information is not available if the fovea is not focused on the target. In this situation, the fovea has to do saccadic movements until it finds the target position. Memory capabilities of the controller, in our case a neural network controller, are essential in this situation to anticipate present peripheral objects.

Fig 4.3 shows the performance of a neural network which accomplishes saccadic movements with memory capabilities. A LSTM network is trained by a Stochastic Policy Gradients Reinforcement Learning method (Wierstra et al., 2007). The task is to find the triangle on a screen with a correct orientation. The scenario is inspired by a work of Schmidhuber and Huber (1991). A correct (with an angle up) and an incorrect (with an angle down) oriented triangle are randomly placed on the visual field. The network looks for the correct triangle by generating saccadic movements and stops if the correct triangle is in the center. Fig 4.3 shows a typical episode. The saccadic movements consist of orthogonal small (1 pixel width) and large (10 pixels width) jumps. The saccades are slightly noisy in consequence of the stochastic nature of the network.

The fovea based solution combines many of the properties of a contextual cueing with the advantages of a fovea that is more sensitive at the fixated location. Since the fovea approach allows the attention system to receive input from a large part of the visual scene, this blurred input can serve much the same way as the context or gist in the systems described above.

4.4 Attending to Moving Targets

The attention systems described above all deal with the environment as if it consisted of stationary scene that may change from time to time. However, in reality, everything is nearly always in motion, either because objects in the world move or because the observer itself is moving. To direct attention toward moving objects it is necessary to anticipate how the objects will move relative to the observer and this problem must be solved by both biological and technical systems.

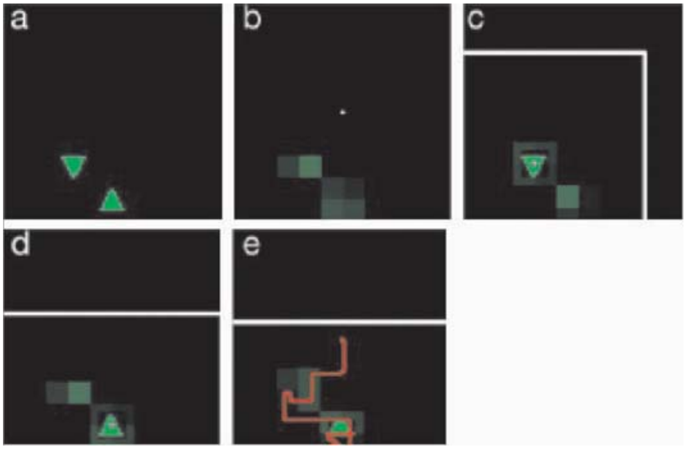


Fig. 4.3 Finding the triangle task. (a) shows the randomly generated picture with two different triangles. (b) is the input - the fovea view - for the network. The fovea is centered in the middle of the image at startup. (c) shows the input with the focus on the incorrect triangle and (d) with the focus on the correct triangle. (e) illustrates the saccadic path for finding the correct triangle.

What are the requirements of a system that needs to predict the motion of a visual target? Consider a system that attempts to predict the position of a target object based on a sequence of its previous positions. Such a system should learn a function from a number of observed positions $p_{t-n}, \ldots, p_{t-1}$ to the estimated position $\hat{p}_t$ at time t. Any of a number of learning algorithms could learn such a function by minimizing the prediction error $e_t = p_t - \hat{p}_t$. The learned function constitutes an anticipatory model of the target motion.

We now add the constraint that the perception of the target, including its localization, takes τ time units. In this case the problem translates to estimating $\hat{p}_t$ from $p_{t-n}, \ldots, p_{t-\tau}$, since the rest of the sequence is not yet available. In addition, this means that the system only has access to the prediction error e_t after τ additional time steps, that is, learning has to be set off until the error can be calculated and the estimate of $\hat{p}_t$ has to be remembered until time $t + \tau$ when the actual target location p_t becomes available.

The important point here is that a system of this kind will never have access to the current position of the target until after a delay. Any action that is directed toward the current target position will thus have to depend on the predicted location rather than the actual one. This is further complicated by the fact that any action directed toward the predicted location will also take some time to execute. For example, if an action is performed with constant reaction time ρ, an action directed at $\hat{p}_t$ at time t will miss the target, since once the action has been performed the target will be at position $p_{t+\rho}$. Consequently, the system needs to anticipate the target position $\hat{p}_{t+\rho}$ already at time t when the action is initiated.

In summary, the system needs to keep track of the target at three different times. The first consists of the currently observed set of positions $p_{t-n}, \ldots, p_{t-\tau}$ that can be called the *perceived now*. The second is the *anticipated now*, that is, $\hat{p}_t$. This is the actual position where the target currently is, but this is not yet accessible. Finally, any action must be controlled by the *anticipated future*, that is, $\hat{p}_{t+\rho}$.

Although this looks like a very complicated way to handle time, unless the delays τ and ρ are negligible, the use of some form of prediction is unavoidable. The delays in the human brain are long enough to necessitate anticipatory models and this has important consequences for how we learn to pursue a moving object with our eyes.

4.4.1 Models of Smooth Pursuit

Smooth pursuit eye movements are used to tracking moving targets and are stimulated by the movement on the retina. Unlike saccades, smooth pursuit eye movements are not under voluntary control and does not involve a shift of attention. Instead, smooth pursuit appear to be more related to fixation of a stationary target in that the same target is attended to all the time. For smooth pursuit to keep the target stationary on the retina, it is necessary to anticipate how the target will move and to place the focus of attention where the target will be at each point in time.

Singh et al. (2006) describes a model based on Itti and Koch's model of selective visual attention (Itti, 2000; Itti and Koch, 2001a) that includes motion processing. They first computed motion saliency features using optical flow models (Lucas and Kanade, 1981), but with poor results. Then they used the pyramidal Lucas-Kanade sparse feature tracking algorithm (Bouguet, 2002) to detect corners in the scene (as their targets had corners on them). They use Gestalt principles that fit the properties of their targets to group features into objects. They model the fixation on moving objects/smooth pursuit by encoding the tracking error obtained from foveal feature update and finite optical flow window into a motion error salience map, called a confidence map. There is an inverse relationship between the saliency of an object and confidence of its position.

The model of smooth pursuit eye-movements of Shibata et al. (2001) uses a predictive model based on linear or non-linear regression networks to predict the behavior of the target stimulus. A cascaded control scheme is used in which the predictor sets the desired path for the eye controller.

The model clearly illustrates that anticipation is absolutely necessary for any system using active vision in a realistic environment. Since the visual processing is delayed, it is not possible to control the movements of the camera using traditional feedback control. Instead, the direction of the gaze must be controlled by a model that predicts where the target is now.

One limitation of this model is that it only allows a single model of target motion although it would not be too complicated to allow several such models that could be chosen depending on the circumstances. An open question is whether there should be many models for different targets or if a single simple model is sufficient for all cases. For example, a model which simply extrapolates the motion of the target, possibly taking acceleration into account, may be sufficient. It is also possible that much more complex models should be learned when the system has to observe complicated but regular motion.

An alternative model of smooth pursuit eye-movements was proposed by Balkenius and Johansson (2005). The main strength of this model is that it includes both continuous and discrete target predictions. The model assumes that a linear predictor is sufficient for normal pursuit of the target, but when certain events occurs, the model learns specific expectations of the behavior of the target.

The model is thus a compromise between a continuous and a discrete model of target motion. The main strength of the model is that the hybrid approach allows it to learn complex scenes where the target may disappear and reappear or change motion abruptly at certain locations. It also acknowledges that the modeling of target behavior must occur in global and not egocentric coordinates.

4.4.2 Engineering Approaches

Here we describe the tracking problem in the task of tracking a moving ball. Ball tracking is a very common, but challenging task for different applications like video streaming, sports game automatic annotation, robot soccer etc. The ball is usually small, has no special marks or characteristics which can be easily detected. It can be mistaken for parts of the player's bodies or is occluded by other objects. Even the color and shape of it are unreliable, since they get blurred from the distance and the fast movement of the ball. Yet another complication comes with the real time requirement of most of the applications.

The classical engineering method to predict a moving ball are Kalman filters and variants (Welch and Bishop, 2006). Kalman filters work great, if the ball is rolling slowly on a flat ground. Prediction limitations become significant when the ball has multiple states, e.g. jumping, slipping, or if the environment is more complex, e.g. equipped with reflecting walls or a rough ground.

In addition to the time consuming model design of the environment a preprocessing step is necessary to convert the real world sensor data, e.g. from a camera image, to the abstract representation of the model: The ball position and speed has to be detected, the positions of walls and other obstacles have to be computed as variables of the model.

Some methods cover the strong nonlinear behavior of the tracked object by the use of multiple model Kalman filters, e.g. Interacting Multiple Model (Blom, 1984; Mazor et al., 1998) or Generalized Pseudo Bayesian (Blom and Bloem, 2004). Other researchers are using particle filters as an alternative to Kalman filters. Particle filters are more general and they allow non-unimodal, non-Gaussian and non-linear problems. Kwok and Fox have successfully used a Rao-Blackwellised particle filter (RBPF) to predict the movements of a ball in a highly dynamic environment with reflecting walls and other obstacles which influence the behavior of the ball (Kwok and Fox, 2004).

A lot of research has emerged in the last years, all dealing with different methods for ball detection and tracking under different environment properties and application requirements. In general a ball tracking process can be divided into 4 interacting steps: (1) Preprocessing of the visual scene, (2) visual ball model, (3) ball detection and (4) ball tracking. Another related field is ball movement prediction, where the common problem of system delay and reaction time is overcome by predicting the ball's position in the future given some history data. This also can be used to support the tracking of the ball, since the ball in the video input has to conform to its physical model of movement.

In the next sections, different methodologies for each of the ball tracking sub-problems will be presented and discussed. The works described here usually cope with more than one of these sub-problems, but rarely with all of them.

Isard and Blake (1998b) present a general object tracking approach, applied to a simple example of a bouncing ball on a table. Kwok and Fox (2004) deal with real time ball detection and tracking for the RoboCup AIBO league. Basketball tracking is described by Ross et al. (2006). The RoboCup middle size league is the test bed of Treptow et al. (2003), who detect and track a black-white soccer ball in real time. Yet another ball game, tennis, is used as application scenario by Yan et al. (2005), who present a ball tracking algorithm. In summary, the most important current works span ball tracking and the different ball models (both physical and visual), which will be presented in detail in the next sections.

Preprocessing Video input is a very unreliable data source, which has to be carefully pre-processed before a small and fast moving object like a ball can be detected and tracked. Traditional approaches include background extraction (Yan et al., 2005; Ross et al., 2006), color enhancement, color segmentation (Kwok and Fox, 2004), sub-sampling, edge detection etc. A major drawback of all preprocessing methods is their high computation requirements, since they need to be applied to the whole video data. Therefore the algorithms need to be simple and fast.

Visual Ball Model Visual models can be roughly divided into color-based, pattern-based and shape-based. Color-based approaches are dependant on the environment and on previous knowledge about them. For example, Kwok and Fox (2004) use a purely color-based ball model for the RoboCup AIBO league, where the ball is expected to be bright orange and no other objects have the same color.

Ross et al. (2006) use a mixture of different classifiers to build a robust visual model of the ball. Six different ball prototype patterns or patches, each of size 19x19 pixels, are learnt from examples. Additionally, the three main components of a PCA over these patterns and background extraction are used to identify the ball. Treptow et al. (2003) use Haar wavelet like features to learn the visual model of the ball from examples. Adaboost is used to create a robust classifier using the single features.

Isard and Blake (1998b) apply B-spline shape fitting of the contour of the ball model, without color information. Quality is measured by intensity gradient normals on the contour. Yan et al. (2005) use a mixed shape and pattern-based approach. Eclipse fitting and gradient vectors of the intensity of the pixels are used for training a SVM to classify the foreground blobs.

Ball Detection Ball detection is done based on the visual model of the ball. There are two different ball detection problems: initialization of the system, when the ball has to be found for the first time, and re-detection of the ball, when it is occluded by other objects. The second case can be usually handled by the ball tracking application itself by using a robust physical model of the ball, which describes its movement.

Treptow et al. (2003) uses a particle filter based approach. When the ball has to be detected for the first time, particles spread through the whole visual field and slowly concentrate on the most promising regions based on the visual model of the ball (see above).

Ball Tracking Treptow et al. (2003) uses a simple particle filter to track the ball, without any supporting physical model. A particle filter is also applied by Yan et al. (2005). However, a dynamic model of the ball is used, where the expected ball movement is dependent on the distance to the players in the tennis match.

State-based approaches are used by Isard and Blake (1998b); Kwok and Fox (2004); Ross et al. (2006). They all describe different states in which the ball can be (rolling, bouncing, kicked, etc.) and either hand-code the state features and their transition probabilities (Isard and Blake, 1998b; Kwok and Fox, 2004) or learn them from examples (Ross et al., 2006). This powerful physical model is then used for tracking the ball, supported by different probabilistic state-space models. The Condensation algorithm was introduced by Isard and Blake (1998a), much resembling a particle filter with weights. A Rao-Blackwellised particle filter is taken by Kwok and Fox (2004).

4.4.3 The State Based Approach

The anticipatory component of both the engineering approaches and in the cognitive models either uses a state based approach or an approach based on prediction. In this and the following sections, we introduce the basic assumption of these two approaches with a simple example.

The following equation shows a toy example of a state based description of a falling ball in one dimension. The variable p_t represents the position of the ball over the ground at time t and v_t is its velocity. The state representation is made in homogenous coordinates - hence the additional 1 in the state description. The transition matrix describes the change of the state of the ball between time t and $t+1$. The parameter g is the gravitational acceleration of the ball.

$$\begin{pmatrix} 1 \\ p_{t+1} \\ v_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ g & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ p_t \\ v_t \end{pmatrix} \tag{4.4}$$

f we denote the state of the ball at time t with s_t and the transition matrix with M, the general state based approach can be formulated as,

$$s_{t+1} = Ms_t. \tag{4.5}$$

The beauty of this formulation is that if we want to predict the position of the ball n time steps into the future, all we need to do is repeat the application of the transition matrix. Thus,

$$s_{t+n} = M^n s_t. \tag{4.6}$$

Now, let's assume that the ball hits the ground, At this point in time, the above state transition matrix has to be replaced by one that describes the bounce:

$$\begin{pmatrix} 1 \\ p_{t+1} \\ v_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ g & 0 & -e \end{pmatrix} \begin{pmatrix} 1 \\ p_t \\ v_t \end{pmatrix} \tag{4.7}$$

Here e corresponds to the coefficient of restitution and the minus sign indicates that the ball will change direction.

With the two transition matrices in place, it is possible to predict the position of the ball at any point in the future given that we know when to select which transition matrix. One possibility is to again use the idea of a context. In this simple example, there are two contexts. Either the ball is in free fall or it is in contact with the ground. The appropriate transition matrix must be selected for each context. If c_t is the context at time t, the state based model translates into

$$s_{t+1} = M_{c_t} s_t, \tag{4.8}$$

where c_t selects the appropriate model at each time step depending on the context.

An obvious problem here is that this formulation assumes that the context at time t is known beforehand. A more appropriate formulation would be to make the context a function of the current state as expressed by the following formula,

$$s_{t+1} = M_{c(s_t)} s_t. \tag{4.9}$$

In this case, the context can be defined as a function that describes the local environment of the ball, for example, whether there is a surface where the ball will bounce or not.

So far we have only considered a situation in which the models are already known. We will now extend the state based framework to an unknown situation. Let us assume that the ball has been observed for some time resulting in the time series $p_0 \dots p_n$. We want to construct the state transition matrices from this sequence. This can be done using a large number of methods, but they all derive from the fact that the state based model defines a linear system. This means that the least squares method is applicable to the solution.

4.4.4 The Prediction Approach

An alternative to the state based approach is to instead view the problem as a prediction task where the next position of the ball has to be predicted based on a sequence of previously observed position. Let us assume for simplicity that we have access to two consecutive positions of the ball p_t and p_{t-1}. To calculate the next position we can use the equation,

$$p_{t+1} = \begin{pmatrix} g \\ 2 \\ -1 \end{pmatrix}^T \begin{pmatrix} 1 \\ p_t \\ -1 \end{pmatrix} \tag{4.10}$$

and for the bounce, we can similarly write

$$p_{t+1} = \begin{pmatrix} g \\ 1-e \\ e \end{pmatrix}^T \begin{pmatrix} 1 \\ p_t \\ -1 \end{pmatrix} \tag{4.11}$$

For the linear prediction approach, the coefficients for the prediction matrix can be estimated, for example, using the least squares method. Various iterative methods are also possible. For example, Balkenius and Johansson (2007) used the least mean squares method to fit the parameters for a linear predictor of this kind in a model of smooth pursuit.

As in the state based approach, it is necessary to use different matrices for the different contexts and the main problem is to partition the different training points into sets that can be used to train the different predictors.

The predictor need not be linear as in the example above. In many cases it would be more suitable to use a non-linear predictor since it may potentially be able to find the two contexts from the data with the use of additional input variables. For the non-linear prediction, some form of learning function approximator such as artificial neural network has to be used.

4.4.5 The Fovea Based Approach

The following example shows a fovea based ball tracking implemented with a supervised trained LSTM network. The environment is based on an Ikaros (Balkenius et al., 2007) plug-in by Christian Balkenius. A ball falls down from a random position with an random angle at the top of the screen and is reflected by the floor, the sides or a horizontal bar (see Fig 4.4).

The basic idea of the task is to demonstrate a ball tracking method without an explicit model of the physical behavior of the ball. Furthermore, also the non-linear model switching, which is normally hand-coded (see Section ball tracking related), is learned automatically. The setup is intentionally kept simple to focus the task on the ball tracking part and not on the visual ball detection process.

Setup The visual scene is a 8 bit grayscale image with a size of 128 by 128 pixels. The fovea has 2 levels of detail, each level has a size of 10 by 10 pixels. The centered high resolution part has the same resolution as the image itself. In the surrounding low resolution part each fovea pixel integrates the information from a patch of 8 by 8 image pixels.

The LSTM network has an input size of 200 values for the 2x10x10 fovea pixels , 20 hidden LSTM cells and 4 output connections. Each of the two pairs of the output values represents an offset to the target position of the ball relative to the current fovea position. One pair delivers the offset to the current ball position and the other pair predicts the position of the ball for the next frame.

The LSTM network is trained with 100 episodes of the falling and reflecting ball. Each episode has a length between 67 and 179 frames. The mean length is 119 frames. This results in 11934 training samples. The test set consists of 2334 frames split into 20 episodes with lengths between 80 and 162 time steps.

Fig. 4.4 Ball simulation environment. The normal environment data is seen on the left image. The ball is represented by a small square. The reflecting objects are gray. The middle image illustrates the fovea view of the same situation of the left image. The blue area shows the fovea's invisible area. The right image shows a typical episode. The green path represents the ball positions and the red path the predicted positions.

Results In the beginning of the learning process the network is only able to follow the ball for short time, before it loses it again. During this phase, best tracking performance is achieved for strictly vertically falling balls and worst results occur at the side walls of the environment. However, after some more training, the fovea-based network is able to track also balls falling at any angle and reflected by the side walls.

After the training is finished, the network has anticipated the behavior of the ball in the particularly given environment and develops two useful properties: tracking the ball when it is visible and imitating its trajectory when the ball disappears. In the second case, the fovea is moving together with the assumed ball trajectory and tracks the ball again as soon as it appears in its view. In summary, the network is able to learn the behavior of the ball and use it for both tracking it and imitating or predicting it.

4.5 Combining Bottom-Up and Top-Down Processes

An obvious shortcoming of bottom-up attention models is that they do not adapt to the current situation, since they always direct attention to the same visual properties. This limitation has been addressed in various ways by different models by including also top-down influences on attention. Note that most top-down models also includes a bottom-up component which is often the larger part of the system.

Westin et al. (1996) describes an early system using both bottom-up and top-down processes for the control of attention. The bottom-up processing uses normalized convolution to filter the input image. This method has the advantage that it makes it possible to include the certainty of an estimate in the calculations thus avoiding, for example, border effects and allow sparse or heterogeneous sampling of the image. It also allows top-down influences to modulate the bottom-up processing in a simple way by changing the certainty of each measurement or signal.

Another form of simple top-down influence was described by Balkenius and Hulth (1999). A bottom-up attention system was enhanced with a simple learning ability that would make the system direct attention to the type of stimulus that had previously been rewarded. The system used a simple form of reinforcement tuning to learn the desired target stimulus. The system was limited in that only a single stimulus type could be learned and the system would thus not adapt to different tasks or situations, but the idea of using reinforcement learning to tune a bottom-up attention system can be useful also in more complex models.

A minimal form of top-down influence in a basically bottom up model was included in the model of Choi et al. (2004). In the model, a human operator can train the system to ignore certain features based on top-down inhibition. The top-down system consists of a modified ART network that receives input from the saliency map and may categorize it as uninteresting and subsequently inhibit it. The top-down part of the model is very simple as essentially categorizes the independent components used for the saliency maps. A more useful form of top-down inhibition requires contextual control (see Balkenius, 2000).

Many models of top-down attention have been inspired by the human neurophysiology. A typical model of this type is the one proposed by Corchs and Deco (2002). It includes a model of the primary visual cortex as well as visual area 2-4 and an object recognition system corresponding to inferiotemporal cortex. It also includes a spatial coding system that works in parallel to the object recognition pathway resulting in one "what" system and one "where" system. Both systems receives top-down biases from a module corresponding to prefrontal cortex.

The model aims at reproducing data at a neuronal level, and as is usual with this type of model, it is not clear how it could be extended to more realistic inputs and outputs. Nevertheless, it suggests many ideas for the architecture of a biologically inspired attention system. In particular, the disassociation into identification and spatial processing is probably a required component of any attentional system as is the local competition within each module of the system.

Deco and Zihl (2001) describe a model that combines bottom up and top-down processing that addresses the problem of selecting both the scale and area of the attended location in an image. It consists of a number of modules that process the input at different spatial levels and can be viewed as hierarchical predictors at the different spatial levels. This results in a sequential coarse to fine analysis of the image. The main contribution of the model is that it shows that it is necessary to process visual input at several simultaneous scales to be able to direct attention either to global patterns or to local features.

Rao and Ballard (1999) introduced the idea of the visual cortex as a hierarchical predictor. A number of processing levels make up a hierarchy of Kalman filters where the feedback from the higher levels act as prediction for the lower levels. The model attempts to explain the behavior of classical and non-classical receptive fields in the visual cortex. In their method, uncertainty about anticipations by components of other components is estimated, affecting the amount of influence levels have on each other (when the higher level is very certain that an event is going to happen in a lower layer, it has more control over the lower layer than when it is not certain at all).

Prediction is thus a fundamental aspect of these models. However, the top-down influences are very direct and it is not clear how the model can handle complex task-requirements since there are no temporal aspects in the basic model. Nevertheless, the framework of hierarchical predictive systems appears to be very fruitful for future attention systems, Especially if anticipation over time could be included.

Spratling and Johnson (2004) developed a neurobiological model of attention in cortex which is based on levels of interacting cortical modules. There is a biased competition within each module which is influenced by top-down and bottom up inputs. The model can reproduce some instances of foreground/background segmentation and contextual cueing.

Navalpakkam and Itti (2003, 2005) extended the previous bottom-up model of Itti and Koch with two types of top-down influences. One is a task specification that is used to decide on which features are important in the scene and also on what spatial locations should be attended to. The other is a system that attempts to find the "gist" of the scene that can further guide the selection of a particular image region. The model also includes a symbolic component that models long-term and working memory.

These additions are necessary extensions of the original model, but it appears that some of the additions follow a very different design philosophy than the original model. However, the symbolic top-down systems suggests how more complex task related control systems could interact with a top-down attention system.

In the model by Balkenius et al. (2004), a number of pre-attentive processing stages were selected based on their computational efficiency and utility in finding targets for the attention system. These bottom-up systems consisted of horizontal and vertical contrast at a single scale, curvature detection using the Harris operator (Harris and Stephens, 1988), and a foreground estimator (Stauffer and Grimson, 1999).

In addition, the model included a number of attentive, or top-down, systems that learn to predict image regions where targets are likely to occur or relations between different features in the image. For example, on attentive system learns associations between target stimuli and visual cues that predict the location of target stimuli. A number of different relations between cues and target were studied including absolute and relative spatial and temporal relations between a single or multiple cues and a target.

Chapter 5
Anticipatory, Goal-Directed Behavior

Martin V. Butz, Oliver Herbort, and Giovanni Pezzulo

> As Man is a reasonable Being, and is continually in Pursuit of Happiness, which he hopes to find in the Gratification of some Passion or Affection, he seldom acts or speaks or thinks without a Purpose and Intention. He has still some Object in View; and however improper the Means may sometimes be, which he causes for the Attainments of his End, he never loses View of an End, nor will he so much as throw away his Thoughts or Reflections, where he hopes not to reap any Satisfaction from them. (Hume, 1748, pg. 33-34)

David Hume may be one of the first who thought about the causes that actually enable us to act goal-directedly in our pursuit of happiness. Besides having usually an end, or goal, in mind, Hume realized that the end must elicit those means that were learned to *correlate* with the end. Such correlation knowledge, according to Hume (1748), was based on three types of connecting "ideas": resemblance, contiguity in time and place, and cause and effect. Knowledge of correlations and cause-effect relations alone, though, do not directly lead to effective behavior. Thus, not only the question how we learn correlations in the environment needs to be addressed, but also how we can exploit the obtained knowledge, if learned properly. While Hume did mainly address the former question, the latter question was acknowledged by Hume only in so far that the acquired knowledge may be used to pursue our goals.

Another related line of research on causation is put forward by Kant (1998). Sloman (2006) contrasted a 'Humean' and a 'Kantian' view of understanding correlations and causations in particular. The former is evidence-based, probabilistic, and statistical. The latter is structure-based and deterministic. Kant highlights the role of concepts and necessity in contrast with the Humean emphasis on observation and correlation. The Kantian notion of causation is more complex and requires understanding of spatial structures and relationships as well as the capability to reason about what happens when they change. While humans are usually seen as explorers that learn correlations in the world based on experimentation—and thus more 'Humean' evidence-based—they rely on and inevitably detect the 'Kantian' a priori structures in our world given due to time, space, and physical constraints.

Just like humans, all other anticipatory systems, both biological and artificial, need to learn and store knowledge about themselves and the world. Only this knowl-

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 85–113, 2008.
© Springer-Verlag Berlin Heidelberg 2008

edge enables them to predict future events, gauge the consequences of one's own actions, and finally interact competently and intelligently with objects or other agents. Thus, one of the questions addressed in this chapter is how knowledge about the world or the self may be represented.

However, regardless if we acquire knowledge based on 'Humean' or 'Kantian' principles, or both, just having gathered knowledge about the world does not mean per se that our decisions will be wise and our actions will be appropriate. It is not even clear how predictive knowledge may be turned into actual decisions in general. The *ideomotor principle*, which dates back to the 19-th century (James 1890; Herbart 1825; cf. Hoffmann et al. 2004), suggests that actions are bi-directionally linked to the effects they usually produce. Thus, once a goal is chosen and activated, the bi-directional links point to those actions that previously caused the goal to come about. While this still does not clarify the actual mechanism of selecting the appropriate means, it implies that an inverse mechanism is necessary that stores means to achieve current goals.

With the ideomotor principle as the basic principle of goal-directed behavior in mind, this chapter analyzes related predictive and anticipatory systems that learn predictive representations of their environment and can use those to act goal-directedly. Predictive systems are systems that are able to predict sensory inputs or pre-processed, more abstracted perceptual input. From an adaptive behavioral perspective most important are systems that *learn* such predictive representation. These predictive capabilities are an important part of any goal-directed behavioral system that is explicitly anticipatory. However, as suggested in the comparison of the insights put forward by David Hume and the ideomotor principle, predictive capabilities are only the first step toward an anticipatory behavioral system. Thus, the second question that this chapter addresses refers to the structures and processes that enable the selection of actions or decision making, based on the acquired knowledge.

We identify two fundamental classes of approaches that realize action selection based on predictive representations of sensory-motor correlations. First, *schemas* form a mental (internal) predictive world model, which encodes all kinds of properties, independent of possible tasks and goals. Although the representation might correspond to an exhaustive internal world model, the schemas alone cannot be used directly for decision making or action selection. Before a decision and action is made, internal processes are required that evaluate possible means in the light of current behavioral goals and desired states. Even more so, to be able to make complex decisions or execute meaningful actions, many schemas may have to be combined. Thus, schema approaches generally build a forward model that is used inversely for action selection.

Second, *inverse models*—in contrast to schema approaches—encode direct connections between behavioral goals and actions. Thus, they may be directly used for action decision making without any further processing. Inverse models can be seen as the result of abstracting or aggregating schemas because they focus on a generalized, inverted representation of properties in the world. In this sense, they can also be considered a world model, which is however rather limited compared to the world models realized in schema approaches.

Accordingly, this chapter first gives an overview of several kinds of schema approaches and inverse modeling approaches. We classify each system from an anticipatory behavior perspective discussing how knowledge is represented and which processes are necessary to turn anticipatory knowledge into behavior. As we cannot review all architectures that have been proposed to date, we exemplify each class of anticipatory behavioral systems with a representative model. Examples are chosen to provide details of state-of-the-art models of goal-directed behavior and to cover a broad range of approaches, including symbolic, subsymbolic, and neural models as well as supervised, unsupervised, and reinforcement learning approaches.

In the next section, we first give a brief history of schemas and provide a definition. We then distinguish different schema system classes and give system examples. Similarly, we discuss inverse model approaches, combinations of both approaches and other advanced techniques. In the second part of the chapter, we assess weaknesses and strengths of the architectures in learning and representing predictions and in using those predictions for the generation of anticipatory cognitive functions. Finally, we contrast the systems' capabilities and give an outlook on potential future macroscopic organizational structures for anticipatory systems, especially highlighting hierarchical and modular structures.

5.1 A Brief History of Schemas

Originally, psychology and cognitive sciences suggested that knowledge about our world is represented by schemas. Drescher (1991) concisely defined a schema as a representation of a triple that links a situation or condition, an action, which may be carried out in this situation, and subsequent effects. How the knowledge is turned into actual behavioral decision making and control remains unspecified. However, any schema representation may be considered as a structure that represents sensory-motor correlations, that is, how motor activity usually affects the perceived environment.

Historically, the term 'schema' may have been firstly introduced by Bartlett (1932) referring to a map or structure of knowledge stored in long-term memory. Successively, Piaget (1954) described schemas in a more operational sense, roughly as mental representations of some physical or mental action that can be performed on an object or event. He considered schemas the building blocks of thinking and as the basic structure underlying behavior and cognition (in a process that he described as 'assimilation and accommodation'). Schmidt (1975) proposed complex schema structures, which encode generalized motor programs for a variety of tasks and internal models of the sensory inputs that accompany movement execution.

Also many approaches in the field of artificial intelligence are based on the schema notion, including *frames* (Minsky, 1988), *scripts* (Schank and Abelson, 1977), *schemas* (Arbib, 1992, 1989; Drescher, 1991; Neisser, 1976; Norman and Shallice, 1986; Pezzulo and Calvi, 2007b; Shapiro and Schmidt, 1982), *anticipatory classifiers* (Butz, 2002a; Butz and Hoffmann, 2002; Gérard and Sigaud, 2001; Gérard et al., 2005; Stolzmann, 1998), *neural schemas* (Mccauley, 2002), *semiotic schemas* (Roy, 2005), and *behaviors* (Brooks, 1991; Maes, 1990). Architectures including distributed and competitive functional units are often referred to as ‘behavior-based’ or ‘schema-based’. Several integrated frameworks have been proposed for designing them; among the most popular ones we can mention the behavior-based approach proposed in Arkin (1998), the *NSL/ASL* in Weitzenfeld et al. (2000), and the *Robot Schema (RS)*—a formal language for designing robot controllers proposed in Lyons and Arbib (1989), which includes perceptual and motor schemas. Drescher (1991) was one of the first who implemented a functional schema-based approach showing simple goal-directed behavioral capabilities.

Schema theories are strongly motivated by biological and ethological models—some of the first implementations intended to replicate the behavior of the cockroach or the praying mantis in robots (Arkin et al., 2000). In general, a schema links conditions, actions, and result components, which are sometimes also called (sub-) schemas. This enables the control system to execute those actions whose conditions are currently satisfied (that is, schemas that apply) and whose result components appear currently desirable. Often, schema theories stress the importance of procedural knowledge, that is, a schema constitutes the long term memory of perceptual or motor skills, or the structure coordinating such skills. Schemas are especially well-suited for parallel and distributed systems, since they can be seen as concurrent computing units.

While several researchers have described the usage of schemas in the perspective of reactive and behavior-based robotics (Arkin, 1998; Brooks, 1991), schemas embed a predictive component that is used for action selection. Moreover, perceptual schemas are often shaped for motor actions. That is, schema representations are usually essential for motor actions. Thus, although not necessarily explicitly anticipatory, schemas serve for the control of behavior with a predictive component. Which schema representations exists and how these representations may be turned into behavior is discussed in the following section on the basis of various system examples.

5.2 Schema Approaches

Schemas integrate situations, actions, and their effects, mostly independent of upcoming tasks, potential goals, or any constraints. However, as these schemas are not related or specific to a certain task or goal, they cannot be directly used for decision making or action selection. For example, consider a schema that specifies that when holding a cup (condition) and drinking out of it (action), thirst will decline (effect) and another schema that specifies that when in a kitchen and grasping a cup, the cup will be held by the hand. Now, given the goal of wanting to quench one’s thirst, and given further the fact of being currently in the kitchen, then both schemas my be integrated suggesting that when grasping the cup and then drinking from it may quench the thirst. However, note that there is no schema that directly specifies what to do given the goal of quenching thirst, rather, little pieces of schema-based information need to be integrated into a more complex decision or action.

In the following sections we review several classes of schema approaches that differ in both, the form of the knowledge representation and the processing of the knowledge. We begin with a review of symbolic knowledge representations, from which actions may be derived by comparing expected results, planning, or the generation of a behavioral policy. Then, neural network approaches are discussed, which ground schemas on simple perceptions and derive actions from planning processes or the preparation of a controller by dynamic programming.

5.2.1 Symbolic Schemas for Policy Learning

An approach for using world model representations to improve policy learning and effectively generating an action policy is the *tabular DYNA-PI* model (Sutton, 1990), which may be considered as one of the simplest schema-based approaches. As in all reinforcement learning approaches, the core is an actor-critic architecture (Sutton and Barto, 1998). The critic implements an "evaluation function" and the actor an "action policy". The evaluation function assigns a reward or reinforcement value to each possible state-action pair. The action policy determines which action to take in a specific state. In addition to this actor-critic model, DYNA-PI learns a predictive world model. This model is composed of two functions, a state transition function and a reward function (both functions may be stochastic). Both functions are learned by initially random interactions with the environment. The world model is used to predict the consequences of actions, in terms of reward and future states of the world. The action policy, due to a reward-backpropagation mechanism, realizes the inversion process. Behavior is triggered, by choosing that action that is expected to yield the highest reward in the long run, which is effectively a form of payoff anticipation.

The key idea of the DYNA-PI model is that an agent endowed with a world model can produce "simulated experiences", besides the experiences gathered during actual environment interactions. Thus, the evaluator and actor can be further trained on simulated experiences. If the learned world model is accurate enough, this "mental training" will speed-up the improvement of behavioral performance in the real world. Thus, besides payoff anticipations, DYNA-PI uses internal simulations of anticipated events to improve its behavior—a form of state anticipation.

Reinforcement learning approaches were recently also carried-over to logic-based representations, in which case they are often referred to as relational reinforcement learning. Kersting et al. (2004) applied reinforcement learning ideas to a logic-based, relational world model framework. Using reward propagation techniques and a matcher mechanism, desired goal states were activated and propagated through the logic-based relational world model. The first-order logic-based abstractions in the world model showed to improve behavior and planning capabilities significantly, also enabling generalizations to similar contexts dependent on the relational logic-based representation available. Thus, besides tabular representations, generalized relational schema-like representations can be applied effectively combining world model representations with policy learning. Another challenge, however, is to learn a suitable generalized schema representation from experience alone, which is addressed in the subsequent sections.

5.2.2 Symbolic Schemas and Prediction for Selection

A prominent example of an online generalizing world model learner are Anticipatory Learning Classifier Systems (ALCSs, Stolzmann, 1998). These learning systems are inspired by the psychological principle of anticipatory behavior control (Hoffmann, 1993; Hoffmann et al., 2004), as well as by the schema approach of Drescher (1991) and DYNA-PI. ALCSs learn a *generalized* predictive model of an environment online. Predictive knowledge is stored in condition-action-effect rules, called classifiers, that represent a schema-based world model. The ACS2 system (Butz, 2002a) combines heuristic search with genetic mechanisms to generalize the predictive world model online.

As DYNA-PI, ACS (Stolzmann, 1998) originally included reward values directly in the schema representations. Given a generalized schema representation, however, reward aliasing can occur in which case the schemas may be sufficiently accurate to predict action effects but may be over-general to represent an optimal behavioral policy (Butz, 2002a). Consequently, XACS (Butz and Goldberg, 2003) was developed, which separates state value and schema learning. XACS is a combination of ACS2's model learning capabilities with the evolutionary online generalizing RL mechanism XCS (Wilson, 1995). The system learns online a generalized state value function, which is represented by a set of condition-value tuples, using XCS-based techniques. Moreover it learns a generalized world model according to the model learning techniques of ACS2. In reinforcement learning terms, XACS learns generalized representations of the state transition function of a Markov decision process (MDP) as well as of the underlying value function.

As opposed to selecting an action based on the best applicable schema, action selection then becomes a two-stage process in which all applicable schemas predict possible next outcomes and the schema is chosen for execution that predicts the maximally suitable outcome, that is, the outcome that is expected to yield the highest value according to the learned state value evaluation function. Thus, the inversion of the predictive capabilities takes place during action selection as well as by means of reward back-propagation mechanisms while learning the value function. More complex decisions or behaviors may be elicited if planning mechanisms are used to combine many schemas.

XACS has shown to be able to robustly learn compact representations of optimal behavioral policies. Policy learning was further sped-up by exploiting the knowledge of the predictive model using the DYNA-based update techniques discussed above (effectively speeding-up the adaptation of the value function). Thus, XACS is a schema system that combines a schema representation with a state value representation to learn a compactly represented optimal behavioral policy quickly, accurately, and reliably.

Recent gradient-based update mechanisms in XCS (Butz, 2006) can improve performance of XACS, so that XACS promises to serve as a robust learner in large, high dimensional MDP problems. With respect to behavioral plausibility, it was shown that ACS2 can be used to simulate the learning of behavioral patterns previously observed in rats (Butz and Hoffmann, 2002). Moreover, since XACS is a system that learns online and from scratch, the implementation of an enhanced XACS system is possible, which may comprise multiple, interacting reward learning modules that may be additionally controlled by motivational and emotional constraints. For example, dependent on the gained learning experiences, it is imaginable that the emotional patterns of the cognitive system may evolve differently resulting in, for example, a very "shy" or a very "bold" system.

5.2.3 Neural-Based Planning

Besides tabular and symbolic approaches, also a neural network-based schema approach (Baldassarre, 2001, 2003, 2002a,b) was implemented, which exploited the prediction and planning capabilities of the schema-based representation. The controller was tested on a simulated robot with a 1D surround camera that solves stochastic path-finding landmark navigation tasks (the robot moves in an arena with white walls and black pillar landmarks by selecting one of eight absolute-direction actions in each simulation time step). Unlike the DYNA-PI architectures, the controller can pursue arbitrary (novel) goals. In particular, the NN can plan with respect to the achievement of any externally or internally generated goal, thanks to the generation of internal rewards in association with them. Whereas DYNA learns to predict rewards assumed to be permanently associated to states, the NN planner is endowed with a "reward generator", which dynamically generates an internal reward when the system achieves its current goal.

The controller builds an efficient "partial policy" by focusing on possible start-goal paths and is capable of deciding to re-plan if "unexpected" states are encountered (Baldassarre, 2003). The simple "forward planner" version of the controller iteratively plans by the generation of chains of predictions from the position currently occupied by the robot. The more sophisticated "backward-forward planner" version of the controller iteratively generates chains of predictions from both the position currently occupied by the robot and the goal state. In both cases, the pseudo-experience so generated is used to train the reactive components of the system as in the DYNA systems. The forward models are composed of neural networks trained to predict the perceptual consequence of action executions. The "backward models" are composed of neural networks trained to "predict" the "origin state" from which the robot might have arrived to a certain state given the execution of a certain action.

Another version of the controller implements a simple form of neural abstract planning that enhances the exploration and evaluation updating capabilities of the controller (Baldassarre, 2001). Abstraction is implemented in terms of planning on the basis of macro-actions (actions composed of n actions of the same type, such as, north-north-north) and action execution at the primitive level.

A more sophisticated modular version of the controller (Baldassarre, 2002a) allows the system to store information about achieved goals and to recall such information so as to decrease the planning burden when the same goals are assigned more than one time. In this case, the goals are not only used to plan but also to satisfy a "motivation" signal that allows the reactive components of the system to recall the knowledge related to previously achieved goals.

An earlier, similar NN planner (Schmidhuber, 1990b,a, 1991c) learned a recurrent NN model and could show capabilities of reinforcement learning and planning in dynamic environments. He also investigated the capabilities of simulating curiosity and boredom with the architecture. Recently, parts of that NN planner were used in some of the modules of a modular and hierarchical control architecture (Gloye et al., 2004) that won the robot soccer world cup (FU-Fighters, Small Size, 2004).

5.2.4 Neural Network-Based Dynamic Programming

Finally, neural networks may be used to integrate schemas into highly flexible movement plans by neurally implementing dynamic programming. A recent computational model of motor learning and control, SURE_REACH, explains the high flexibility of human motor behavior (Butz et al., 2007a). This hierarchical architecture stores an associative model of state transitions as well as a redundant associative mapping of hand locations with arm postures. Population-encoded spatial representations enable the application of dynamic programming techniques. To move the hand to a desired location, the hand position is first translated into a representation of the redundant postures that coincide with the target hand position. This redundant intrinsically encoded goal representations and the encoded state transition model is then used to generate a movement plan by neurally implemented dynamic programming.

Without additional constraints, the minimum path in posture space is executed. However, if the task imposes additional constraints, alternative action sequences may be generated by simple neural inhibitions. Thus, SURE_REACH is able to reach hand targets while incorporating task-specific constraints, for example, adhering to kinematic constraints, anticipating the demands of subsequent movements, avoiding obstacles, or reducing the motion of impaired joints (Butz et al., 2007a; Herbort and Butz, 2007). The approach is generally similar to early self-supervised control approaches (Kuperstein, 1988; Mel, 1991), but extends them to the sensorimotor control of redundant bodies. Compared to previous neural network models of motor learning and control, SURE_REACH accounts for higher behavioral flexibility and adaptivity without the need for relearning.

5.3 Inverse Model Approaches

Schema approaches may be used to represent a model of the world in a very fragmented way and they require complex processes to turn a goal into an action. A different approach of modeling goal-directed behavior and the function of executive modules is put forward by the notion of the inverse model (Kawato, 1999). An inverse model is an internal representation that inverts the flow from action to effect. It thus generates actions that are useful to reach a desired state. To follow the example mentioned above, an inverse model might specify that thirst may be quenched by drinking from a cup held in the hand and that when in the kitchen without a cup, a cup should be grasped. Note that in this example, the model directly specifies which action to execute given start and goal. The model does not specify the actual consequences of actions, though. Rather, it merely suggests that the action in the given circumstances is usually advantageous for achieving the specified goal.

Thus, the inverse model approach is fundamentally different from the schema approach. Whereas in the schema approach, the fragments of information stored in the schema have to be processed to arrive at a decision or to generate an action, an inverse model aggregates such a process in a direct mapping from situations and goals to actions. An inverse model may thus be seen as the result of an aggregation of many executed schema processes that are combined and generalized into a simple mapping. A drawback of inverse model approaches is that the acquired mapping is highly inflexible because it generates a rigid mapping from goals to actions. Thus, if the environment changes or novel tasks have to be solved, alternative behaviors may be required to maintain effective behavior. Inverse models cannot provide alternatives so that expensive relearning would be necessary without schema knowledge. Of course, a direct mapping has the advantage that no potentially costly planning or other preparatory processes are necessary to determine actions. Thus, while inverse models appear well-suited for rigid, quick, automatized control, inverse models alone are rather inflexible and essentially may hinder the quick adaptation to novel situations or tasks.

5.3.1 Inverse Models in Motor Learning and Control

In computational neuroscience, inverse model approaches are implemented in feedback error learning (FEL) models of cerebral motor learning (Berthier et al., 1992, 1993; Barto et al., 1999; Haruno et al., 2001; Karniel and Inbar, 1997; Kawato et al., 1987; Kawato and Gomi, 1992; Schweighofer et al., 1998b,a; Wolpert and Kawato, 1998) In short, these model predicate that the cerebellum is an inverse model for goal-directed motor behavior. The cerebellum exerts control of goal-directed movements and adjusts its output according to an assumed cerebral linear feedback controller. During learning, the cerebellum thus learns a direct mapping from goals to motor outputs.

While FEL models rely on the accurate control of a simple controller, other inverse model paradigms learn their inverse models simply by the observation of randomly sampled actions or physical plant correlations. The most prominent class in these approaches are *direct inverse modeling* (DIM, Baraduc et al., 1999, 2001; Bullock et al., 1993; Kuperstein, 1988, 1991; Ognibene et al., 2006) and the related *resolved motion rate control* (RMRC) approaches (D'Souza et al., 2001; Jordan and Rumelhart, 1992; Whitney, 1969). Both techniques learn a situation-dependent mapping between goals and motor commands. For example, a non-redundant arm may learn its inverse kinematics by mapping a hand position goal to a corresponding arm posture, which may trigger suitable motor activity. RMRC is more robust in the face of redundant plants, storing that action for a particular goal and situation combination that was optimal during learning.

Redundant bodies or environments generally pose a problem to inverse modeling approaches, because one of many equivalent actions has to be associated with each particular goal. Thus, among all potential actions, those are stored in the inverse model that optimize additional criteria. These optimality criteria have to be defined to enable the acquisition of an inverse model (D'Souza et al., 2001; Engelbrecht, 2001; Todorov, 2004). An inverse model is thus only suited to optimize a single criterion, which was defined before learning. Changes in the criterion, for example, due to demands of novel tasks or changes in the environment, reduce the performance of an inverse model or may even render it completely incapable. In contrast, only the ability to adapt optimality criteria quickly from one movement to the next enables the flexibility of human behavior (Rosenbaum et al., 1995). Additionally, the need to adapt an inverse model to an optimality criterion seems to hinder unsupervised sensorimotor learning (Herbort and Butz, 2007). Thus, due to its inflexibility and limited learning capability, the inverse-model view of motor control has recently been challenged with the proposition of the SURE_REACH model (Butz et al., 2007a).

5.3.2 Inverse Models and Schema Approaches

Despite the principled difference that schema approaches store a general model of the world and inverse models encode preprocessed, task-specific goal-action links, both approaches are certainly strongly related.

First, inverse models and schema approaches may happen to represent identical sets of information, if a one-to-one mapping between goals and actions exists (of course, dependent on the situation). In such a context, each goal can only be pursued by a single action and executing this action is sufficient to reach that goal. Thus, inverse models and schema models inevitably represent the same information given that both models always yield the same action to execute. However, the general equivalence may only exist in rather abstract, artificial models, seeing that environments are usually continuously in flux.

Second, some schema approaches first prepare a behavioral policy, dependent on the goal and potential constraints, and then execute behavior accordingly (e.g., SURE_REACH). The resulting policy can be considered an ad-hoc inverse model, which has been generated solely and exclusively for the current goal and situational demands. In this sense, these approaches combine the advantages of inverse models with the flexibility of schema approaches.

Thus, it seems most plausible that efficient anticipatory behavioral control can only be accomplished with both representations present—schema approaches to know the environment and also to verify current action successes and inverse model approaches to effectively and progressively automatically control behavior.

5.4 Advanced Structures

The previous section has described a broad variety of approaches, which implement executive models and enable goal-directed behavior. In the following, we outline how the described approaches may be integrated and combined with predictive models to enhance performance and address more complex tasks.

It is evident from many lines of research in psychology, neuroscience, computer science, and engineering that efficient behavior is not only based on the quality of schemas or inverse models, but also on the quality of sensory data or the quality of the output processing. For example, sensory ambiguity may be reduce by integrating multiple sources of information or by predictive top-down connections. Likewise, motor control may be facilitated by being able to identify basic characteristics of plants or by dividing the generation of motor actions from high-level goals into computationally simpler sub-processes.

In this section, we want to highlight structures, in which schema approaches or inverse models may be combined or embedded to optimize behavioral control. First, we show that the combination of executive modules and predictive mechanisms can enhance behavioral performance. Second, predictive models and executive modules may be coupled to form higher order schemas, enabling effective behavior in varying contexts. Finally, hierarchical control structures may stabilize behavior and enable the solution of more complex types of problems.

5.4.1 Prediction and Action

The discussed inverse model approaches are capable of generating actions or behaviors to pursue certain goals. In this section, we discuss how these architectures may be integrated with predictive models, that is, forward models of schema approaches, to enhance control. Forward models enable the anticipation of changes of the environment or effects of one's own actions. In this section, we first introduce long short-term memory (LSTM) recurrent neural networks, which allow the prediction of a series of future events. Then, we give examples in which forward models provide internal feedback to stabilize and enhance control.

5.4.1.1 Recurrent Neural Network Approaches

Recurrent neural networks (RNNs) were proposed in Elman (1990) mainly as a language and grammar processing system. However, recent advances have applied RNNs to a variety of problems including time series analysis, speech processing, or robot navigation tasks. RNNs seem to have particularly strong potential for the formation of predictive and anticipatory structures. A good overview of a variety of RNNs can be found in Zappacosta et al. (2007). In the following, we focus on the LSTM system, which solves particularly hard grammatical problems as well as challenging time series analyses problems[1].

LSTM models are artificial RNN architectures that are endowed with neural gate-based structures (Hochreiter and Schmidhuber, 1997). Input gates and output gates guard input/output access to the internal states of neurons, enabling the algorithm to maintain memory over theoretically infinitely long periods of time. The networks effectively deal with the problem of vanishing gradients, which is usually a major

[1] Some of the other neural network approaches discussed below as well as the NN approaches discussed in relation to schema-based approaches also contain some form of recurrence but are usually not explicitly termed "recurrent" NNs.

limitation in other RNN structures especially in problems in which long-term dependencies need to be remembered. LSTM can remember and relate events distant in time. It is thus expected to be most suitable as a prediction tool for anticipatory systems that need to detect long-term dependencies (in memory) or that have to deal with partially observable MDPs (POMDPs).

LSTM RNNs use “memory units” that use the “constant error carousel” (CEC) to propagate error, theoretically, infinitely back in time. The memory units are protected by an input gate and an output gate that multiply incoming activation and outgoing activation to effectively “gate” the memory information so that it can apply only when necessary. In later papers, a “forget-unit” was added that makes the timing of the memory unit more precise and allows learning from continuous data streams (Gers et al., 2000). Kalman-filter enhanced learning (Pérez-Ortiz et al., 2003) increases learning speed by orders of magnitude but also increases the learning complexity of the system.

While LSTM RNNs thus showed to be highly useful in predicting challenging context-free grammars as well as real-valued data streams, it remains to be shown how the learned structures may be inverted to trigger effective online action selection and motor control.

5.4.1.2 Internal Feedback

After given an example of a predictive model, we now explain how such models may be used to improve behavior. In order to efficiently select effective actions, it is essential to know the current state of the world as exactly as possible. However, biological and artificial sensor systems are prone to noise, and information about the environment may only reach the executive modules through time-consuming processing pathways. In this case, a predictive model may compensate for these effects. For example, if sensor data is noisy or ambiguous, a prediction of the current state of the environment based on recent perceptions and recently executed actions may enable the formation of a concise representation of the environment. Thus, forward models may help to implement Kalman filtering approaches. Likewise, decisions or actions may be adjusted according to their expected impact on the world before feedback from the environment is available or even before the action is executed.

A processing control model, partially based on internal feedback, was proposed by Kawato et al. (1987), who applied it to arm movements. In the model, control is initially exerted by a linear feedback controller. The controller is not very efficient, though, because the delayed feedback results in a slow control process. A forward model is gradually learned and improves the control process by providing internal feedback, thus enabling a much faster control process. Finally, an inverse model replaces the linear controller to enable maximally effective movements.

In this setup, the forward model provides almost instant feedback about the expected consequences of the generated motor command. Thus, the motor system does not have to wait for external feedback to adjust its output but can adjust motor commands on their expected effects. As a result, the motor commands are much more accurate and movement dexterity increases. Later on, internal feedback can still level out small inaccuracies of the inverse model and compensate for forward model estimation errors.

Several recent studies have suggested that forward models, as body emulators, are essential for efficient body control (Kawato, 1999; Wolpert et al., 2001). Moreover, various studies suggest that internal forward model feedback is used to estimate spatial body location (Wolpert et al., 1995) as well as to improve behavioral control of fast reaching movement (Desmurget and Grafton, 2000) or of pole balancing (Mehta and Schaal, 2002). Grush (2004) relates such representation also to higher level cognitive processes. Thus, internal, anticipated sensory feedback appears to play an important role in behavioral control, state estimation, as well as higher cognition.

5.4.2 Coupled Forward-Inverse Models

In the review of schema approaches, the condition part of a schema mostly referred to a specific sensory state or situation, whereas the action part referred to a single action entity. However, it is also possible to form schemas from more complex notions of perception and action. Forward-inverse models directly couple forward and inverse model information essentially representing context in current forward model accuracy.

Adaptive agents should be able to operate in different contexts or environments and should quickly adjust to changes. Thus, it has been proposed that multiple executive modules, schema approaches or inverse models, may be represented for different contexts. However, this requires the quick identification of contexts to enable the selection of appropriate modules. Predictive models may play a key role to identify contexts and participate in the selection of the right executive module for the right task.

Several researchers have proposed such decentralized architectural schemes for the control of action that are based on coupling forward and inverse models, both, in a localized (Demiris and Khadhouri, 2005; Pezzulo and Calvi, 2006a; Tani and Nolfi, 1999; Wolpert and Kawato, 1998) and in a distributed fashion (Tani, 2003; Tani et al., 2004). The former approaches are based on the mixture of experts architecture (Jacobs et al., 1991) while the latter are based on the self-organization of the representational space in RNNs.

In these approaches, forward models are coupled to executive modules (that is, some form of inverse model), representing a higher level schema. Such a schema is applicable if the predictive model makes continuously accurate predictions. The condition part of a schema is a forward model, which enables the identification of the underlying properties of a situation, for example, how objects or bodies react to certain actions. These underlying properties may not be evident from regarding a single instance of the perceptual input. Likewise, the action part of a high level schema might not be a single motor command but an entire controller, which is especially suited to exert control in a specific context. Thus, the accuracy of the predictive model indicates the suitability of the executive module in each schema. As predictive model and executive module of a schema are trained exclusively in parallel and in an assigned context, the predictive model will only be able to make valid predictions in the context, for which also the executive module was trained.

These architectural schemes have been used for multiple purposes. For example, they were used to select actions appropriate to the context (Wolpert and Kawato, 1998), they were used for action observation and execution (Demiris and Khadhouri, 2005), and they were combined with a motivational system in which active drives influence action selection (Pezzulo and Calvi, 2006a) . When behaviors can be combined linearly, the models can also generalize behaviors. Algorithms for learning and combining contexts in non-linear dynamics have also been proposed (Vijayakumar et al., 2005).

This integration of predictive models and executive modules into a schema may help stabilize the selection of control strategies, even in noisy contexts. Furthermore, it is possible to deduce abstract properties of a situation, which may not be directly concluded from sensory input. A drawback can be that the difficulty between learning the inverse model and the forward model may differ significantly, so that the current accuracy of the forward model may not necessarily reflect the current suitability of the coupled inverse model.

5.4.3 Hierarchical Anticipatory Systems

Each of the so-far presented approaches is limited to comparatively small problem domain, such as the control of simple movements, planning a chain of actions, maze navigation, or the prediction of events of a certain kind. Whereas each approach may be well suited to solve the problems in its domain, it cannot be easily extended to the high variety of tasks that humans or autonomous artificial systems face, ranging from the need to determine long-term goals to the accurate control of basic actuators. This limitations can be overcome by integrating the described approaches into a hierarchical framework. Many neurological and psychological studies and models suggest that effective cognition and consequent behavior is based on hierarchically structured systems, for example, accounting for complex sensory processing, cognition, and behavior (Giese and Poggio, 2003; Koechlin and Summerfield, 2007; Loeb et al., 1999; Poggio and Bizzi, 2004; Powers, 1973; Riesenhuber and Poggio, 1999; Todorov, 2004; Wolpert et al., 2003).

By structuring a cognitive architecture, separate problems in tasks like sensory processing or motor control may be solved by different modules. For example, it was demonstrated that the spinal circuity is able to counteract some perturbations on its own, thus making motor control easier for the central nervous system (Loeb et al., 1999). Accordingly, the CNS provides a basic control strategy, for example, by setting reflex gains or muscle stiffness. This enables the CNS to flexibly adjust motor control to varying tasks without the necessity to react to small perturbation, whose compensation is left to the spinal system, which is well suited for this task due to its fast feedback control loops. Models of central motor control suggest that the cerebellum replaces and further optimizes cortical motor networks in the control of frequent, well-trained movements (Berthier et al., 1993; Barto et al., 1999; Kawato et al., 1987; Kawato and Gomi, 1992; Schweighofer et al., 1998b,a). Furthermore, also movement specifications may be based on incorporating movement selection biases on multiple levels (Cisek, 2006; Herbort et al., 2007). Also, current hierarchical models of vision (Riesenhuber and Poggio, 1999; Giese and Poggio, 2003) were suggested to be extended to motor control problems Poggio and Bizzi (2004).

In conclusion, hierarchical, modular systems can address specific computations in specific modules, consequently reducing the complexity of each computation and enhancing stability due to the partial autonomy of each module. However, in most approaches information flows in a single direction, for example, from visual sensors to abstract representations or from behavioral goals to movements. Systems in which layers bidirectionally influence each other seem to be more promising for the understanding of complex perception, cognition, and behavior (Hinton et al., 1995; Hinton, 2007; Rao and Ballard, 1997). Higher layers may try to model the behavior of the lower layers, correcting lower layer states when the lower layers do not have the knowledge of predicting their own state. In other words, higher layers may correct the state of lower layers by, for example, resolving ambiguity. Uncertainty measures of each module's state and also attentional influence may further modify the influence layers have on each other. In sum, combinations and integrations of effective sensory processing and motor control modules promise to yield highly flexible adaptive decision making and control structures that go far beyond the competency of a flat architecture.

In the following section, we now first evaluate the predictive and anticipatory capabilities of each considered system and then discuss correlations and contrast the distinct features of each system. In particular, we first list predictive and anticipatory capability criteria. Next, we discuss the various schema approaches and inverse model approaches with respect to these criteria. The subsequent discussion draws the attention to the currently missing system capabilities and proposes various future research options.

5.5 Evaluation of Predictive and Anticipatory Capabilities

We now evaluate the predictive and anticipatory capabilities of the introduced approaches on the basis of the taxonomies of predictive and anticipatory capabilities introduced in Chapter 2. To do so, we consider the capabilities of each system individually and finally discuss their correlations and differences as well as potential complementarities.

More concretely, we distinguish and discuss the following predictive qualities:

- *Symbolic vs. real valued predictions*: Does the system form symbolic, real-valued, or both types of predictions?
- *Discrete vs. continuous predictions*: Does the system form predictions about a discrete next time step or can it form continuous predictions over time?
- *Noise robustness*: Are learning and representation of predictions noise robust?

- *Sensory vs. payoff predictions*: Does the system form sensory predictions or payoff predictions?
- *Single vs. multiple predictions in representation spaces*: Does the system form one single prediction or multiple predictions (e.g. concrete and abstract)?
- *Full vs. partial predictions*: Does the system attempt to predict complete future states or is it able to focus on sub-states?
- *Exact vs. fuzzy, distributed predictions*: Does the system form on exact predictive representation or does it also estimate the confidence of its predictions?
- *Immediate vs. long term predictions*: Does the system predict only next states or is it able to form immediate long-term predictions (without chaining immediate predictions)?
- *Generalization capabilities in predictions*: Is the system able to generalize its predictions to similar states in the environment?
- *Single vs. multiple sources of information*: Does the system distinguish between multiple sources of information such as sensory, context, and motor activity information?
- *Markov-dependent vs. independent predictions (MDP vs. POMDP)*: Does the system rely on full observability to be able to reliably learn accurate predictions?
- *Distinction between self and other*: Does the system distinguish between predictive representations about own future states and other future states?

With respect to anticipatory capabilities, we distinguish the following qualities:

- *Direct vs. indirect inversion*: Are goals directly (inversely) mapped to actions, or is the mapping done indirectly?
- *Reward vs. plan-based inversion*: Is the inversion accomplished by means of back-propagated payoff representations or by means of explicit representations of future states?
- *Planning capabilities*: Is the system able to plan?
- *Full vs. partial planning*: Can the system also generate partial, abstract plans?
- *Online vs. offline representations*: Is the system bounded to generate future representations based on the current state or can is also generate anticipated representations offline?
- *Flexible goal-oriented behavior*: Can the system flexibly pursue novel goals?
- *Adjust to new task constraints*: Can the system immediately adjust behavior to novel task constraints?
- *Curious behavior*: Can the system act curiously, directedly exploring novel territory or environmental properties?
- *Epistemic actions*: Can the system act pro-actively in order to search for missing information?
- *Surprise mechanisms*: Can the system implement surprise mechanisms upon unexpected perceptions?
- *Motivational goals*: Can goals be chosen or pursued based on the system's current motivations?

With these distinctions of predictive and anticipatory system capabilities in mind, we now evaluate the considered systems.

5.5.1 Schema-Based Systems

5.5.1.1 Predictive Capabilities

Schema-based systems exist for symbolic and for real-valued representations as well as for discrete and continuous predictive representations. Generally, schema systems focus on sensory predictions and, dependent on the schema representation, this prediction can comprise multiple predictions in space as well as in time. Also, partial predictions are generally possible and the predictions often contain confidence estimates and thus represent fuzzy predictions. A tenet of schema approaches is that condition structures try to focus on those sources of information that are maximally suitable to generate representations of future states. Thus, schema systems are able to generalize, dependent on the representation and learning mechanisms employed. Multiple sources of information are usually considered—albeit not necessarily processed by different modules. Except for successful applications in deterministic POMDP problems (Holmes and Isbell, 2005; Landau et al., 2003; Métivier and Lattaud, 2003; Stolzmann, 2000), the successful development of internal states for the solution of stochastic POMDP problems still remains an open challenge. Finally, schema representations may also be projected, or mirrored, onto other entities in the environment, representing their sensory-motor correlations and internal states. To the best of our knowledge, currently no system exists that accomplishes such a task autonomously.

DYNA-PI More concretely, the tabular DYNA-PI model may be considered the most restricted system amongst the considered schema-based systems. It is able to work only on discrete, distinct, and symbolic state representations and generally cannot be consider noise robust. The prediction of the next state may be a distribution over states but originally was also restricted to exact next states. The original DYNA-PI approach does not contain any generalizations, either, which clearly poses a huge scalability problem. The table essentially grows linearly in the number of distinguishable states and actions. Provided diverse real-valued sensory input, the number of states may be infinite, which is an obvious limitation of tabular approaches. Thus, tabular DYNA-PI is only suitable to investigate internal planning and reinforcement propagation mechanisms rather than to actually apply the system to real-world adaptive system control tasks. More recent schema-based approaches tackle this limitation with various generalization mechanisms.

XACS The XACS system is a purely symbolic, discrete prediction learner. It combines model and RL learning to learn sensory predictions and payoff predictions. The sensory predictive module relies on determinism in the environment but can ignore fluctuations of irrelevant sensors (Butz, 2002a). The reward learning part is rather noise robust (Butz et al., 2004a, 2003a).

Besides concrete predictions, XACS also provides the certainty of its state predictions and the corresponding state value predictions. The predictions are generalized in that irrelevant attributes can be ignored. The predictive model is usually significantly smaller than a completely specified model. Currently, the architecture is flat without hierarchies. Irrelevant attributes are ignored and can be explicitly identified as irrelevant for accurate predictions or as unpredictable.

Experimental evaluations have shown that the system can ignore irrelevant attributes and, given irrelevant attributes, it beats learners that learn a tabular problem representation. The predictive models are always generalized and are usually much more compact than tabular approaches.

There are no hierarchies and predictions are currently on one time scale only. However, predictive chains can be generated so that long-term predictions are possible in a limited sense. The generalization mechanisms in XACS focuses on the attributes that are relevant for accurate predictions of the next sensory inputs and the consequent reward, respectively. Thus, regularities are detected and object clusters are expected to be identifiable by the mechanisms. More sophisticated actions such as hierarchical option-type actions (Barto and Mahadevan, 2003) or motor programs have not been investigated so far. Also contextual information, except for action codes, has not been treated separately from pure sensory inputs in any way.

Sequence learning capabilities or performance in POMDP problems were not investigated with XACS so far. Currently, the predictive learning capabilities are restricted to MDP problems, because no internal states are used. Since actions are directly included in the classifier structure, discrete action codes currently need to be used currently and encountered changes are related to one's own actions only.

Neural Based Planning Whereas the original tabular DYNA-PI architecture is a tabular lookup system, which operates on discrete symbolic inputs, the NN-based architecture is a continuous predictive system, which predicts feature-like sensory inputs or even continuous sensory changes. The NN-based architecture can be considered rather noise robust. It forms stochastic predictions. However, the forward models of the implemented architecture currently produce deterministic predictions of the full sensory input. It cannot ignore irrelevant or unpredictable input, nor can it focus on the prediction of partial input. Confidence values and fuzzy predictions may be employed by using an appropriate error functions and related weight update mechanism. As in DYNA-PI and XACS, the NN-based schema architecture represents concrete predictions in terms of sensory input and the architecture is flat enabling only immediate predictions (but see Baldassarre 2003, for a preliminary investigation of a NN planner whose predictions span further time steps ahead).

The NN-based system can generalize to similar events and similar sequences, but it does not develop object-oriented representations (no clustering) nor any type of grammar representation (no recurrence), except for potential emergent representations in the hidden layer of the NN due to back-propagation. Thus, only the back-propagation algorithm in the implemented NNs may be able to ignore sensory inputs.

The system distinguishes between two sources of information, sensory inputs, which are continuous, and action inputs, which are discrete and are used to select expert forward models dedicated to them. If continuous actions were included, then actions should be considered as an additional input or may be coded in a population code fashion. So far, there were no investigations that would include other context-based information or that would learn about environmental dynamics that occur independent of the system's own actions.

As there are no internal state representations, the neural-based planner relies on the Markov property. Moreover, the system is completely self-centered being currently unable to project one's own predictive knowledge onto other entities in the world.

SURE_REACH Devised as a model for movement generation rather than prediction, SURE_REACH does not explicitly implement any predictive mechanisms. Nevertheless, the neural networks that are used for control are also suitable to predict future states, the developed sensorimotor model is purely associative. The kinematic mapping may be used to predict hand locations from (predicted) postures and vice versa. Likewise, the sensorimotor model may be used to predict the trajectory that results from issuing a sequence of motor commands. Due to the population encoding of these internal models, also the uncertainty of predictions may be represented.

In sum, the system can be used to form real-valued predictions on a continuous time scale, predicting perceptions. It can be considered rather noise-robust and may even form, albeit limited, multiple predictive representations (arm postures and hand locations). These predictions are full but, due to the population encoding, inherently represent fuzzy distributions of states. Besides the option to chain predictions, as in all other schema systems, the predictions are immediate and restricted to one time scale. As the other approaches, it also integrates action and sensory information for prediction. SURE_REACH does not address partial observability nor the challenge to form distinct representations of self and other.

5.5.1.2 Anticipatory Capabilities

The schema-based systems have several anticipatory capabilities in common—most obvious is the fact that all rely on indirect inversions to trigger anticipatory behavior. That is, they need to use their predictive capabilities in some way to trigger goal-oriented behavior. This inversion mechanism is sometimes purely reward-based, sometimes plan-based and sometimes a combination of both. Moreover, all schema systems have the possibility to plan, albeit usually generating full plans only. Also the offline generation of mental representations is usually possible. Dependent on the involved inversion mechanisms, goal-oriented behavior can be more or less flexibly adjusted. Curious behavior is generally implementable, however, epistemic actions as well as surprise mechanisms require further additions. Motivational goals may also be included in each of the systems.

DYNA-PI DYNA-PI models are able to form explicit predictions online and offline. Thus, situation-dependent planning as well as offline reflections are possible and also implemented. DYNA-PI models are generally goal-oriented anticipatory system. However, DYNA-PI models do not have the flexibility to account for novel goals without any significant re-planning effort. That is, DYNA-PI models usually learn one (or a couple of alternative) behavioral policies. If none of the policies is currently suitable because, for example, the goal is relocated to a novel location or multiple goals are suddenly distributed over the environment, complete re-planning is often necessary. System behavior is not directly goal-initiated but goal-oriented, due to the learning of a behavioral policy. Thus, it is the reward inversion that results in the generation of goal-oriented behavior. Curious behavior has been implemented in some DYNA-based approaches, however, epistemic actions as well as surprise mechanisms remain to be further investigated. DYNA-PI may learn policies for several distinct goal representations so that it may choose between the pursuance of available goals based on the system's current motivational state.

XACS The anticipatory capabilities of XACS are generally similar to those of DYNA-PI with the advantage that the system is able to learn generalized policies in partially noisy environments that contain many additional, irrelevant sensory inputs. Moreover, XACS does not learn a pure policy representation and thus does not rely purely on reward inversion but also incorporates a plan-based inversion. That is, anticipated next states and their associated state values trigger action decisions. Partial planning is possible, but it is dependent on the chosen representation. However, the learned state value function cannot be changed without significant relearning effort. As the DYNA-PI model, the XACS model lacks the anticipatory flexibility to account for any possible goal, but it learns to adapt only to the goals, that is, the rewards, encountered during learning.

The behavioral policy can be improved to cause curious behavior (improving model and policy learning) as well as greedy behavioral patterns (improving and speeding-up policy learning, Butz, 2002b). The usage of a list of currently least accurate predictions combined in a priority list, similar to the work done by Moore and Atkeson (1993) in their prioritized sweeping mechanism, showed to improve behavioral learning even further (Butz and Goldberg, 2003). As in DYNA-PI, though, neither epistemic actions nor surprise mechanisms were further investigated so far. Currently only one RL module was implemented but different RL modules for different motivations are easily interpretable and combinable in the XACS framework. Thus, the system has strong potentials to study multiple motivational influences and emotional integrations. For example, opportunistic behavior may be triggered by combining current motivational utility measures with current predictions.

In sum, the XACS approach has shown that the combination of learning generalized representations of both a predictive model and a state-value representation is a highly suitable approach that yields effective anticipatory action decision making and control. However, as in the DYNA-PI approach, without additional mechanisms or representations, the system cannot adapt to novel task constraints or goal representations effectively.

Neural Based Planning Since the NN planner generates reward internally in correspondence to (externally or internally generated) goals by means of a specific module, the NN planner can self-generate reward. This reward is associated with any possible state that the system might happen to pursue as a goal. Thus, goals may be associated with motivations, triggering internal reward and also shaping behavior internally. During planning, the NN planner focuses on states that lie between the starting position and the goal, and those around them. When in planning mode, the actor is a model of itself acting in reality. In this sense, the system actually predicts its own choices in states potentially experienced in the future. This is similar to the XACS model, in which predictions are generated and compared to the state-value representations choosing that action that leads to the anticipated highest state value.

Since the NN planner does not distribute reward values, behavior may be more flexible but requires expensive, potentially exponential online planning upon the activation of a novel goal representation. Action decision is based on a plan-based inversion in that hypotheses are generated looking ahead, distributing reward, and assigning maximally effective actions to NN-based states. While the DYNA-PI system is a very simple planning system that only predicts (potentially stochastically distributed) concrete next states, the NN-based system plans based on predictable state aspects. Unpredictable inputs are ignored but partial predictions are currently not possible. Planning is quite robust in the sense that it involves all states that might likely be visited during action execution. It takes place precisely as an offline improvement of the control policy in relation to the assigned goal.

Curious behavior can be included easily by directing the behavior during exploration to NN regions in which the predictions have high uncertainty. Epistemic actions were not investigated so far and also surprise mechanisms have not been considered in further detail. However, goals can become motivations in the multi-goal versions of the planner so that motivational goals can be readily incorporated and the goal-based planning mechanism can also account for novel goals flexibly, even though this may be computationally very demanding.

With respect to action initiation, the system learns to associate goals with actions once a planning step has been applied successfully. In this case, goals trigger actions / motor (control) programs, especially in the multi-goal version of the system when same goals are assigned more than once. The planning process "compiles" goal-related information into the reactive components of the system. Thus, relying on schema representations, the neural-based planner actually forms a goal-dependent inverse model mapping.

In conclusion, NN-based DYNA-PI is a typical schema-based approach that is capable of online planning and of generating simulated experiences offline. The NN-based approach has the advantage of implicitly generalizing, filtering noisy inputs, and ignoring unpredictable inputs. Hierarchical implementations of the approach, possibly in combination with recurrent structures, await future research effort.

SURE_REACH The SURE_REACH model transforms spatial population-encoded goals (that is, hand locations) into intermediate, redundant population-encoded goals (that is, a corresponding subspace of arm postures). The posture representation is then used to generate motor commands by means of dynamic programming-based planning. Thus, also SURE_REACH is an indirect inversion model that uses its schema representations encoded in its inverse kinematics and inverse sensorimotor model to map goals to actions. The inversion is reward-based in the sense that goal activation, which represents reward, is inversely propagated through the population-encoded posture space. However, due to the population encoding, the operation of this mechanism is not exponential but polynomial and thus efficiently executable online upon goal activation. Thus, the system can yield highly flexible goal-oriented behavior being able to approach any reachable goal in space. Additional constraints can modify internal target representations or movement preparations to adjust behavior to situational demands. For example, new task constraints—be it disabled joints, preferred arm postures, or anticipated subsequent goals— can flexibly modify the unfolding anticipatory behavioral control structure.

Although not further investigated so-far, curious behavior patterns may be included as well—potentially causing movements to spatial regions that have not been sufficiently explored or had not been reached for an extended period in time. Epistemic actions, however, will require the incorporation of additional mechanisms. Also surprise mechanisms have not been further investigated, yet. Motivational goals, however, may be easily incorporated and a first goal-selection mechanism was coupled with the SURE_REACH model, showing avoidance behavior and the preference to reach rewarding locations in space (Herbort et al., 2007).

Thus, SURE_REACH may be considered the most flexible architecture of the schema systems considered. However, so-far SURE_REACH has only been applied to a rather restricted arm reaching tasks. Thus, the generality of the approach as well as the scalability of the representation to higher dimensional problems needs to be further investigated and developed (cf. Butz et al., 2007a).

5.5.2 Inverse Model Approaches

5.5.2.1 Predictive Capabilities

In the discussed inverse adaptive control approaches, the motor control capability is the most relevant and most investigated system part. Thus, the systems' predictive capabilities are of lesser importance and an inverse model may very well be combined with other predictive systems to yield stable online control in addition to the effective inverse control mechanisms, as done in the discussed forward-inverse model approaches.

Generally, DIM, RMRC, and FEL all work on continuous valued inputs and represent continuous changes. No discretization takes place. In general, also payoff representations are not included. All systems can be considered rather noise robust but usually have complete representations and have no mechanisms to focus on partial environmental inputs only. Thus, focusing capabilities are restricted to motor activity and goal dependence. Furthermore, inverse model approaches usually do not rely on exact representations and thus form rather fuzzy associations. The continuous changes are usually represented on one time scale so that more abstract concepts of change are not represented in space nor in time. Nonetheless, the goal representation and the implicit "belief" that the associated goal will be achieved may serve as an interesting representation of a desired future state.

All systems known to us process information in one task-context only and cannot switch between different contexts with different optimality criteria. Finally, the capability of forming internal predictions is restricted to the possibility of activating (desired) goal states, effectively predicting that an activated goal state will become actual. Triggering goal states based on current internal states, which might depend on current internal motivations and emotions, is another future research challenge. Also the distinction between self and other and the potential projection of the inverse model on observed behavioral patterns awaits future research investigations.

Anticipatory Capabilities The anticipatory capabilities focus on decision making and action initiation. As suggested by the name itself, direct inversions from goals to actions trigger behavior upon goal activation. Thus, the inversion is direct, neither reward- nor plan-mediated. The consequence is that the systems also do not have any explicit planning capabilities, and essentially also no re-planning capabilities. Behavior is either successful or fails. Upon failure, further learning is required. However, since most inverse modeling approaches are state dependent closed-loop control mechanisms, disturbances during behavioral control can usually be compensated for. In fact, all three approaches discussed learn inverse mappings from goal representations to motor activity that are conditioned on the current sensory input, which represents the current state of the body in the environment.

Environmental exploration may be biased dependent on the current system knowledge, which enables curious behavior. The possibility to impose goal-directed behavioral execution during learning may be further explored, potentially moving from very general, inaccurate representations to progressively finer-coded, more accurate control. Coupled forward-inverse models discussed above come into mind here, where the forward model accuracy may determine motor activity during learning, and thus bias the focus of inverse model learning. Also motivational constraints may be incorporated easily, resulting in the selection of a goal, which is then pursued by means of the inverse model. Thus, motivations may trigger goals in the reachable space, which can be approached without additional computational effort besides the invocation of the direct inverse mapping.

In sum, inverse adaptive control approaches may be considered as the tools that can realize anticipatory, goal-based action decision making and control. Due to their focus on this aspect of motor control and their general lack of predictive capabilities, it seems straight-forward to modularly combine inverse adaptive control systems with suitable forward predictive mechanisms, which may stabilize control in dynamic control problems. This has been done by the discussed forward-inverse model combinations.

5.5.2.2 Advanced Structures

The discussed advanced structures do either form only predictive representations of sensory inputs or couple some of the discussed system modules. Thus a separate discussion of the considered systems is not carried through.

It may be noted, however, that RNN-like mechanisms can be expected to be necessary to tackle POMDP problems, since internal state representations are necessary in this case. To combine multiple sources of information, these systems may need to be further modularized, which has been partially realized by the gating structure of neurons in the LSTM system. However, partial predictions, and predictions on multiple levels in time and space most likely require the incorporation of multiple forward and inverse modules and the successful knowledge exchange between these modules. Coupled forward-inverse model show one approach to successfully combine direct controllers with schema-based forward models. The discussed hierarchical system approaches may be used to form hierarchical controllers and generate predictions at multiple levels of abstraction in time and space. Moreover, additional challenges, such as epistemic actions, may be tackled with such system combinations. The current capabilities of all discussed systems are now further contrasted from a broader perspective, identifying current system shortcomings as well as arising challenges.

5.6 Discussion

The system classifications point towards several immediate and longer term challenges. In this discussion, we contrast the different systems with respect to their predictive and anticipatory capabilities and identify the most important challenges lying ahead. Hereby, the combination of several systems and system capabilities appears highly promising to generate more complex, autonomously learning, highly adaptive, flexibly behaving cognitive systems.

5.6.1 Contrasting Predictive System Capabilities

The system categorizations showed that there are a rather wide variety of predictive learning systems, each of which also have distinct anticipatory processing potentials. Although it is hard to contrast these potentials directly, Table 5.1 shows an overview of the predictive capabilities of the discussed learning architectures. All systems exhibit highly promising but in many cases differing predictive capabilities. The table may serve as an indicator of the most important challenges lying ahead for each investigated system and which aspects are the most immediate challenges that point towards successful system enhancements and improvements.

The table suggest that there is a current lack of system competencies in several seemingly highly relevant aspects of predictive capabilities: (1) the development of competent predictive system that are able to learn predictions on multiple levels of abstraction in time and space; (2) the development of systems that effectively incorporate context information in their predictions. These two points are discussed in the remainder of this section. Albeit also important challenges, the problem of handling environments with only partially available information (POMDP problems) as well as the problem of the self/other distinction is not further discussed due to the diversity of the problem and its strong dependency on representations and diversity in the approaches to this problem.

Although several of the predictive systems have the potential to predict multiple aspects and provide accuracy or confidence estimates of their predictions, it seems to be difficult to provide multiple predictions in parallel, such as the prediction of

Table 5.1 The contrasted predictive capabilities of the considered systems suggest further advancements as well as potential system combinations.

Aspect	DYNA-PI	XACS	NN-b.D	SURE_REACH	Inverse Mod.
Form	Symbolic	Symbolic	Real-Valued	Real-Valued	Real-Valued
TimeScale	Discrete	Discrete	Continuous	Continuous	Continuous
Noise Robust	No	Partially	Yes	Yes	Yes
Payoff/Sensory	Both	Both	Sensory	Sensory	Sensory
Multiple Space	Single	Single	Single	Single	Single
Full/Partial	Full	Partial	Partial	Full	Full
Det./Fuzzy	Deterministic	Partially Fuzzy	Potentially Fuzzy	Fuzzy	Potentially Fuzzy
Time: Imm./Longer Term	Immediate	Immediate	Immediate	Immediate	Immediate
Generalization	No	Yes	Yes	Yes	Yes
Info. Sources	Two	Two	Two	Two	Two
(PO)MDP	MDP	MDP	MDP	MDP	MDP
Self/Other	Self	Self	Self	Self	Self

next sensory input, plus the prediction of the position of an object in the input, or the prediction of other, often pre-processed environmental features. The hierarchical networks starting from Rao and Ballard (1997) might be an approach to realize such multiple abstract capabilities. The hierarchically combined layers, structured appropriately, may each have a different (emerging) type of abstract representation and thus also abstract predictions. It seems that the integration of other mechanisms, such as the clustering-for-prediction capabilities of the XACS system or the long-term dependency detection capability of the LSTM system, into these hierarchical network structures points towards a highly challenging but also highly rewarding future research direction.

Related but not identical to the capability of predicting at multiple levels of abstract representations lies the capability of predicting at multiple time scales. Again, hierarchical networks seem to have the most potential in this respect. However, even more important than representational abstraction is the question of how to abstract in time. To generate flexible longer time-bridging capabilities during learning, it needs to be clarified when predictive responsibility should be delegated to the next higher level. Early work in this direction suggests that learning at a higher level should be activated if the current level is well-predicting on average but currently encounters highly ill-predicted input (Schmidhuber, 1992a,b). Interestingly, it was recently shown that a very similar principle can serve for the effective detection and generation of options, that is, higher level motor programs, in reinforcement learning (Butz et al., 2004b; Simsek and Barto, 2004). In general, the information content received from the sensory inputs must be significantly higher and persistently high in order to delegate predictive responsibility to the higher prediction layer. Further research in this respect seems very important.

Another approach for multiple levels of abstraction in time is the consideration of delay in sensory feedback. Hierarchical control structures partially take these feedback constraints into account, such as the work of Kawato et al. (1987), in which a lowest-level PD controller serves as backup in case the higher level inverse model-based controllers and forward models are incorrect or inaccurate. The combination of these principles with more competent network structures, points towards another big future research challenge.

The incorporation of multiple sources of information for prediction, apart from the distinction of sensory inputs and action input, is also only partially realized in most of the predictive systems. Hereby, it can be expected that context information should not be simply included as an additional lower level input, but rather should be exploited as a different type of input that serves as a focusing and predisposition mechanism in the system. Thus, in the rule-based XACS system, context may pre-select currently relevant rules, or, in the LSTM system, context information may be used to open and close certain input, forget, and output gates in order to stream information flow in a context-dependent way. The usage of context information from Balkenius' context dependent attention-processing and reinforcement learning systems (cf. Chapter 4) may serve as an inspiration of how to incorporate such mechanisms in a more flexible way into predictive learning systems.

Besides these possible advances, it should be kept in mind that predictive system capabilities are only useful if they serve a purpose, that is, if they affect motor control favorably. To generate competent anticipatory cognitive systems, predictions need to be learned in order to improve learning and behavior. Thus, the general challenge is to develop more competent anticipatory decision making and control systems and possibly also bias the learning of predictions on the resulting anticipatory behavior capabilities. To achieve this endeavor, it will be necessary to combine several predictive systems and couple predictive and inverse systems for the problem structures at hand. Moreover, it will be necessary to exploit their respective competencies modularly to generate more effective anticipatory processing mechanisms. How this might be achieved is outlined in the following section.

5.6.2 Contrasting Anticipatory System Capabilities

Before contrasting the systems' anticipatory behavioral capabilities, we want to point out that the model learning components themselves are not as much influenced by their own predictive capabilities as might be advantageous. Although most considered systems use error-based learning principles, targeting learning resources towards task-specific, motivational goals poses an interesting additional challenge. That is, while "learning for control" may be the first principle, "learning for the achievement of ecological relevant goals" may be an even more focused principle that points out that learning should focus on those control aspects that are really relevant to the learning system.

Table 5.2 shows the current anticipatory capabilities of the discussed learning systems. Schema-based approaches all have similar properties, although the various implementations differ in certain respects depending on their generalization capabilities and utilized representations. Inverse modeling systems are additional mechanisms that may shape behavioral learning directly. In addition, RNN approaches are expected to be useful to learn the achievement of longer-term goals. Hierarchically structured top-down, bottom-up systems may serve well to form abstracted repre-

Table 5.2 The contrasted anticipatory capabilities also show several current drawbacks as well as potential system combinations and integrations.

Aspect	DYNA-PI	XACS	NN-b. DYNA	SURE_REACH	Inverse Mod.
inv.model	no	no	no	yes	yes
reward based	yes	yes	partially	yes	no
general planning	yes	yes	yes	yes	no
focused planning	no	yes	no	no	no
offline simulation	yes	yes	yes	yes	yes
flexibility	low	medium	medium	high	medium
flexible goals	no	no	limited	yes	limited
curious behavior	yes	yes	yes	yes	yes
epistemic actions	no	no	no	no	no
surprise mechanisms	no	no	no	no	no
motivational goals	no	limited	limited	yes	yes

sentations in space and time to be able to plan and act goal-directedly on a more conceptual level.

Besides the potential learning improvements by the means of anticipatory mechanisms, the table shows that several other capabilities require future research. First, faster behavioral adjustments due to unexpected sensory inputs have hardly been investigated. That is, surprise mechanisms could be exploited further for (1) fast self-stabilization mechanisms and (2) the activation of additional cognitive resources for more focused learning and adaptation. Kalman filtering-based updates and other error and information gain estimations may help to improve control and stabilization capabilities in this respect.

Second, task-dependent planning mechanisms may be investigated further. The combination of different predictive methods to enable prediction for action decisions on multiple levels of abstraction seems inevitable. It also remains an interesting question, how exact planning needs to be in order to be sufficiently effective. Davidsson (1997) showed that one-track predictions (those that predict only the usual behavior-dependent future and do not consider alternatives) are often sufficient to improve behavior by inducing preventive mechanisms if the usual behavior leads to undesired states.

Third, while curious behavior has been implemented in a few architectures, epistemic actions were not successfully shown in any of the considered architectures. Epistemic actions may, however, be realized in several systems. However, it remains unclear which predictive representations can most effectively trigger epistemic actions. It seems necessary that a system would need to generate hypotheses about the environment and trigger actions to verify uncertain but relevant hypothesis. For example, in a search task, a robot may look behind an obstacle to see if the ball might be there. Kiryazov et al. (2007) have generated a first realization of such a system on a real robot platform. The system is able to generate hypotheses based on analogy making, consequently triggering goal-directed verification activity.

Hierarchical NN-based system architectures may offer another solution for the realization of epistemic actions. Once higher levels are able to pre-activate lower level neurons, these pre-activations may not only lead to the faster detection of such inputs but also to the activation of suitable motor activity to search for the hypothesized inputs. In general, while systems might have a general curious action selection mechanism, for example, for improving predictive model learning, epistemic actions may be based on the same principle of predicted information gain, only that in this case, plasticity needs to be more dynamic in that the entropy of current important available information needs to be considered and selectively improved. Such mechanisms may lead to truly curious behavior and the automatic activation of epistemic actions.

Fourth, the coupling of motivational mechanisms and potentially even emotional mechanisms with the behavioral decision and control modules poses additional challenges. Context may be handled as a special input to the predictive and to the control system and it may reflect current system motivations. The activated contexts—activated, for example, by a neural activity pattern in the hierarchical neural architecture—should trigger matching motor programs and action decisions that usually lead to the activated context. As discussed, coupled forward-inverse models are a good candidate in this respect, selecting those coupled models that are maximally suitable given the current context.

5.6.3 Integration

The contrasted factors show that the challenges ahead in the design of competent, flexible, and highly adaptive cognitive system architectures comprise system improvements of predictive and anticipatory capabilities. However, possibly even more important, they require the effective combination and integration of various learning and representational mechanisms.

Research currently still focuses on the improvement of particular predictive system capabilities. In the future, though, we expect that successful combinations of different predictive system capabilities will become increasingly important. We expect the following enhancements to be particularly fruitful:

1. The development of predictive systems that process and combine different sources of information (such as context information, sensory information, and action information).
2. The implementation of predictive hierarchies that can generate predictions at different levels of abstraction in time and space.
3. The coupling of predictive representations with action control representations.

Especially the last point poses a great challenge but might be the key to the generation of actual cognitive systems. Perceptions need to be linked with appropriate action codes (causing affordances and bottom-up action predispositions). And, vice versa, action codes need to be linked to corresponding sensory effect codes that are expected to change after action execution. With such a cognitive structure at hand, many anticipatory capabilities might even emerge naturally from the system structure itself. However, even with a less sophisticated representation, several advanced anticipatory capabilities will need to be investigated:

1. Anticipatory representational shaping needs to be further developed, that is, the learning of representations directly for effective behavioral decision making and control.
2. The further development of curious behavior capabilities and epistemic action capabilities: to realize a cognitive system that automatically activates epistemic actions. It seems important that such behavior is triggered by the anticipated information gain that seems most relevant for the achievement of current goals.
3. Anticipatory top-down mechanisms need to be further developed, which influence bottom-up sensory processing. This includes attentional mechanisms (cf. Chapter 4) but also action decision making and control mechanisms since action decision making can be considered as yet another attentional process.
4. A motivational and potentially emotional module may be coupled with the predictive system in order to induce even better action decision making capabilities, enabling the execution of opportunistic actions and actions that are anticipated to satisfy expected motivations (such as taking food and water on a hike).

The anticipatory enhancements are certainly not stand-alone but are very interdependent and also highly dependent on the predictive representations used. Thus, the discussed enhancements of the predictive capabilities of the system should not (only) be pursued in isolation but rather should be designed from the beginning to serve the realization of effective anticipatory action decision and control mechanisms. It is expected that interactive, emergent, and unforeseen properties will be detected along the way of this research endeavor and will as well lead to novel insights in information processing, adaptive behavior, embodiment, and cognition as a whole.

5.7 Conclusions

This chapter has shown that there are various challenges ahead. In order to create competent, anticipatory, adaptive learning systems, the systems do not only need to be competent in learning accurate predictive models of their environment but also need to be able to effectively exploit the learned models for adaptive behavior. This process is expected to be interactive rather than iterative in that the developing predictive representations should immediately cause anticipatory mechanisms that, vice versa, immediately influence the further development of the predictive representations. The categorizations and contrasting discussions in this chapter may serve as guidelines for the development of such more effective anticipatory mechanisms and competent cognitive systems. It is hoped that this chapter does not only provide a useful overview of the discussed systems but that the chapter also encourages further assessments of learning systems with respect to their predictive and anticipatory capabilities and the creation of combinations of these systems to tackle the challenges ahead.

Chapter 6
Anticipation and Believability

Carlos Martinho and Ana Paiva

6.1 Introduction

In this chapter, we discuss the relation between anticipation and believability. We start by introducing the concept of believability and the importance of both emotion and anticipation in the generation of believable behaviour. Then, we present some related work examplifying how emotions and anticipation have been used in the field of synthetic characters, with a particular focus on anticipation. Afterwards, we present how the relation between anticipation and emotion has been researched by the authors to create the *emotivector*, an anticipatory mechanism aimed at assisting the generation of autonomous believable behaviour for synthetic characters. As a fusion between the fields of affective computing and anticipatory computing, the emotivector generates affective signals from the mismatch between predicted and sensed values, and is inspired by the psychology of emotion and attention. We present two applications of the emotivector, demonstrating its adequacy: one in the realtime control of a situated, embodied agent inhabiting a virtual world, and another, in the real-time control of the affective expressions of a robotic affective chess 'buddy'. Finally, we describe a successful integration of the low-level emotivector mechanism in a high-level cognitive agent architecture and the benefits of such integration.

6.1.1 Animation and Believability

The word animation derives from the Latin *animare*, meaning "to breathe life". However, to create the illusion that a synthetic character has a life of its own, motion alone is not sufficient. Every movement in an animated scene must have a reason for being. This argument lead John Lasseter, chief creative officer at Pixar and Walt Disney Animation Studios (WDS) to state that technology could never bring life into a character (Lasseter, 1987). Now, technology is taking up the challenge.

Artificial Intelligence (AI) researchers have long sought to create autonomous creatures. The thought of these entities brings special delight when they are imag-

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 115–133, 2008.
© Springer-Verlag Berlin Heidelberg 2008

ined to project a sense of being "really there", appearing to have a life of their own. In their quest towards creating the Illusion of Life, a term coined by Ollie Johnson and Frank Thomas from WDS (Thomas and Johnston, 1994), AI researchers have opened the door for traditional character animator techniques to become part of the synthetic character inheritance.

One of the first concepts transferred from traditional animation to the field of synthetic characters was the concept of *believable* character, a character that "provides the illusion of life and thus permits the audience suspension of disbelief" (Bates, 1994). Although the concept falls prey to its subjective nature, it has being explored in seversal media (e.g. literature, theatre), and WDS, in the late 1920s, developed a set of practices that became the fundamental principles of traditional animation to achieve believability (Thomas and Johnston, 1994). This reference is relevant for AI researchers as it promotes the mind of the character as a driving force of the action. When all the character's movements are understood as the result of its thought processes that connect them, then the character becomes more important than the techniques that went into its animation. The audience forgets that it is an animation and is actually entertained by a synthetic character, that in the process has gained a life of its own.

6.1.2 Emotion and Exaggeration

The pioneering work by Joseph Bates' group on the role of emotions in believable agents brought two important concepts from WDS guidelines to the field of synthetic characters: *emotion* and *exaggeration*. Based on the work of professional animators, Bates argued that the clear portrayal of emotions is a central requirement for believable characters as the consistency of the expression relates to the perception of personality and as such constitutes the affective base of believability (Petta and Trappl, 1997). Before using the term *believable* characters (Bates, 1994), Bates had referred to such characters as *emotional* characters (Bates, 1992). The importance of clearly expressing emotions in the creation of believable synthetic characters presented by Bates has been confirmed by subsequent work, such as Elliott et al. (1999), Martinho et al. (2000), and Dias (2005). As a result, synthetic characters generally have an underlying affective model.

Exaggeration appears as a means to convey more clearly the modelled emotional states. As human are experts in recognizing human behaviour, a highly realistic human in a synthetic character places unconsciously increased demands on the character motion. Given the difficulty in producing completely realistic animations, the success of the animation industry, both traditional and computer animated characters, has been the result of the creation of stylized and caricatured animated characters. As characters are not real, exaggeration can be used to convey more effectively a certain emotion that otherwise could go unnoticed and break the suspension of disbelief.

6.1.3 Anticipation

This chapter focuses on a complementary line of research that is inspired by another principle from WDS guidelines: *anticipation*. For an animation to be clearly understood, the audience must know at all time: what is going to happen, what is happening and what had just happened (Lasseter, 1987). The first part of the action is the preparation for action, also referred as *anticipation*, and guides the attention of the viewer to make sure that the motion is not missed, and meaning lost. Although anticipation is such a central concept in the language of traditional animation, it has had but a secondary role in the field of synthetic characters.

The quest for believability has sent researcher on two parallel paths: a pragmatic approach inspired by arts such as drama and character animation, and another that strives for higher levels of autonomy by providing the synthetic character with biologically plausible (Blumberg, 2003) or psychologically sound behaviour (Marsella and Gratch, 2003). Both paths emphasize the concept of believability as a dimension of synthetic performance closely related to the adequate expression of emotion, and different computational models of emotions have been proposed to aid achieving believability — e.g. the several implementations of the cognitive theory of emotions by Ortony et al. (1988) in Loyall (1997), Martinho et al. (2000), and Dias (2005). However, few computational models explicitly integrate the concept of anticipation in the creation of life-like behaviour. Anticipation is usually found 'diluted' in the planning mechanisms of the synthetic character (Marsella and Gratch, 2003) or disguised as an emotion by itself, such as in Plutchik's theory of emotions (Plutchik, 1991). This chapter highlights anticipation as an essential part in the creation of believable behaviour and the relation of anticipation with affect.

6.1.4 Anticipation, Emotion, and Believability

Anticipation and emotions are closely related. One of the principal functions ascribed to emotions is precisely that of anticipating events, especially when such events are relevant to the central concerns of the organism (Strongman, 1996). *If I am walking in the woods and, suddenly, 'something' ahead on the path lets out a loud roar, my heart races, my muscles tense, I 'feel' afraid and ready to run away.* Using emotions, I was able to decide among a great range of possible actions, by eliminating most of the consequences of each from consideration a-priori, that is without any time being wasted on their consideration (Guttenplan, 1994). Emotions framed the process by rendering salient only a tiny proportion of the available alternatives and conceivably relevant facts, anticipating which parameters should be taken into account in the decision process and readying resources anticipated to be relevant for the outcome.

As emotions may elicit anticipatory behaviour, the anticipation of an event may also elicit an emotion. Continuing our example... *To avoid the 'frightening creature', I decide to take an alternative albeit longer path to return to the village where I am headed. Unfortunately, near the end of the path, I realize that the bridge over the uncrossable chasm has been destroyed — probably the storm, last night...* I experience disappointment, but also discouragement, as I am unable to do anything about the situation. Anger appears as a result of the unfairness of the situation: *I was to deliver an important antidote to the village.* Emotions elicited by anticipation are often related with expectations, commitment towards important goals, and the validation or invalidation of both expectation and goals. As such, a same outcome can lead to a wide range of emotional experiences, based on different types of expectations. To be prepared for what is to come is a crucial factor in survival. The emotional valence of an expected event provides the ground on which to develop strategies that enable us to adapt quickly and efficiently to changes. Even when the valence of an expected event is unknown, there is evidence that our brain process this information taking a pessimistic approach (Herwig et al., 2007). From an evolutionary perspective, we cope better with a potential threat in the environment by anticipating the worst case.

Mood can also influence and be influenced by anticipation. Continuing our example... *I'm in a bad mood. I start anticipating all the consequences of not being able to bring the antidote to the village soon. My mood worsens as negative thoughts invade my 'mind'.* In a bad mood, not only perception will favour negative events, but also the anticipation of events will tend to be biased by the mood valence (e.g. positive or negative) (Miceli and Castelfranchi, 2002).

As it is possible to go from anticipation to an affective state and back, anticipated emotions may appear in the loop. Continuing our example... *As I approach the chasm, I perceive a feeble cry for help from below. I climb down using the rope from the broken bridge, and find a small boy I had never seen before. He has clearly been poisoned, those marks on his arm... are unmistakable, I have seen them too many times. I have only enough for one application of the antidote, though... If I walk away, the child will not stand a chance, and his face will haunt me for the rest of my life. I know it, but I cannot bear to come empty handed back to the village: the lives of several villagers depend on this antidote being analysed by the village shaman. I carefully place the child on my shoulder and start moving back to the first trail, determined to face the 'frightening beast' if necessary. I just can't leave him here to die, I would not be able to live with it.* When making decisions, people often anticipate the emotions they might experience as a result of the outcomes of their choices. In the process, they simulate what life would be like with one outcome or another (Meller, 2001). The anticipation of post-behavioural feelings can influence people's behaviour, as behavioural choices have been found to be based upon anticipated emotional reactions following a particular behaviour (Richard et al., 1996). While anticipating my behaviour and its context, I 'pre-felt' the emotion I expected to feel in the anticipated situation, to some degree of intensity, and this had a relevant impact in my decision process. Anticipating salient emotions is another vital function of an organism. The expectation of an affective event can trigger regulatory processes that prepare the organism to cope with a possible threat, for instance.

The use of both anticipation and emotion provide consistent and more explicit meaning to motion, assisting in the creation of more believable scenes. Consider the following story: *upon returning to the village with the antidote, I decided to take another path to avoid possible danger, but the bridge at the end of this path was broken due to the storm of the previous night. I heard a cry for help. Using the rope from the bridge, I found a child that was poisoned, and carried him on my shoulder* *back to the first trail, decided to face the danger ahead.* Although this story contains the same amount of detail as the previously described, it lacks believability. Using anticipation and emotion, everything becomes more consistent, continuous, more “meaningful”, more believable. Anticipation is much more than a design trick from traditional animation but also an important factor to be taken into account in the design of affective systems, when their use is to provide meaning to motion and create believable behaviour.

6.2 Related Work

This section briefly describes relevant works in the field of synthetic characters that address the issue of creating believable behaviour using emotions and anticipation, with a strong emphasis on the anticipatory approaches.

6.2.1 Oz Project

In 1992, Bates, in collaboration with Loyall and Reilly, founded the Oz project which aimed at creating virtual interactive theatres, populated by synthetic characters developed following three important principles from traditional animation (Thomas and Johnston, 1994): (1) the viewer should, at any time, be able to attribute a clear emotional state to a character; (2) as the actions of a character depict its inner thoughts and such thoughts are influenced by emotions, the viewer should be able to view emotions in the actions of a character; (3) the actions should be exaggerated to ensure that such emotions are clearly understood by the viewer.

To achieve such goals, Reilly and Bates (1992) developed an agent architecture (Tok) composed of two subsystems: the EM system, based on the theory of emotions from Ortony et al. (1988), which generates emotions based on perception of the virtual environment and the internal state of the agent; the HAP system, which performs action selection according to current goals and emotions of the character. Using the Tok architecture, one of the created worlds, the “Edge of Intention”, presented three synthetic characters (the Woggles) the user could direct to improvise stories in real time. The user could change the character’s mood, select a location in the virtual world towards which the character should move, or select certain behaviour that would make the Woggle do or say something. Each character immediately obeyed such directions by improvising an appropriate course of behaviour, colouring its improvisation with life-like qualities: normall variability, idiosyncrasies, mood-related modulations of behaviour, event-based changes in mood, and adherence to social conventions.

6.2.2 EMA

Gratch and Marsella (2004) lay out a computational framework, EMA (acronym for EMotion and Adaptation) where appraisal and coping act as core reasoning components for human-like autonomous agents. This approach is grounded in Lazarus (1991) cognitive-motivational-emotive system, and focuses on process rather than surface behaviour, namely: appraisal, which characterizes the person's relationship with the their environment, and coping, which suggests strategies for altering or maintaining this relationship. Cognition informs both of these processes.

EMA implements the following algorithm: (1) construct and maintain a causal interpretation of ongoing world events in terms of beliefs, desires, plans and intentions; (2) generate multiple appraisal frames that characterize features of the causal interpretation in terms of appraisal variables; (3) map individual appraisal frames into individual instances of emotion; (4) aggregate emotion instances into current emotional state and overall mood; (5) adopt a coping strategy in response to the current emotional state.

EMA contributed to the design of a virtual reality training environment that teaches decision-making skills in high stakes social situations: the Mission Rehearsal Exercise (MRE) training system. In MRE, intelligent agents control virtual humans, playing the role of locals, friendly and hostile forces, and other mission team members. In the evaluation of MRE, Gratch and Marsella emphasize behavioural *consistency* as a key challenge facing the design of interactive life-like agents, i.e. the coordination of all functions of the agent (e.g. perception, planning, natural language processing) into a single coherent individual, over time. Although the term anticipation is never used, anticipation can be found diluted in the affective planning mechanism.

6.2.3 Duncan the Highland Terrier

In the attempt to build increasingly sophisticated autonomous interactive synthetic creatures, Blumberg's Synthetic Character's Group at MIT Media Lab developed a synthetic character to experiment with different kinds of *expectations* in graphically embodied creatures: Duncan, the Highland Terrier. Duncan lives in a graphical environment which he perceives through a synthetic perception system which includes simulated audition and point-of-view rendering.

Anticipation is an important part of Duncan's behaviour, generally using the term "expectation formation" for anticipation. Two aspects related with the use of anticipation will be discussed here: the work by Isla and Blumberg (2002) on object persistence, and the work by Burke and Blumberg (2002) on apparent temporal causality.

6.2.3.1 Object Persistence

Isla and Blumberg (2002) argue that the absence of expectations based on spatial structure significantly impairs any pretension to life-likeness — "if a ball rolls behind a wall, it would appear either broken or colossally stupid for the creature to then not know where to look at it" — and provided Duncan with a sense of *object persistence*, following the definition by Piaget (1954): "persistence of mental images of objects after they have stopped being perceived, the ability of making deductions about where the objects could be, and to act on these deductions".

Isla and Blumberg defend that a Gaussian distribution, although more compact, is not an adequate representation in virtual environments, and propose a "probabilistic occupancy map", which is an hexagonal grid overlaid on the environment, where each node contains the probability of a target object to be contained at that location. At each step, each node passes some fraction of its own activation to its neighbours, and an additional modifier that favours diffusion in the direction of movement of the mobile object. The verification of the expectation is performed by the observation of the environment. When an object is perceived, the probability of the node representing the location is set to 1 and all others to 0.

Isla and Blumberg defined *salience* as "the degree to which an observation violates expectation". Following the work by Kline (1999), they point out two types of violations: unexpected observations (which they designate *surprise*) and negated expectations (which they designate *confusion*). At any time, the most salient object is selected as the focus of attention.

6.2.3.2 Apparent Temporal Causality

Burke and Blumberg (2002) argue that because the representation of time and the rate at which a creature experience relevant stimuli are fundamental for living systems, such representation can be used to enhance the life-like qualities of a synthetic creature. Inspired by Ethology, they developed an approach on learning and action-selection based on Gallistel and Gibbon (2000): Scalar Expectancy Theory (SET) and Rate Estimation Theory (RET). These theories require an animal to represent the length of the interval between stimuli and the rate of reinforcement associated with various stimuli.

According to SET, when a creature perceives a salient stimulus, the creature starts an internal timer that records the (subjective) interval between that stimulus and another salient stimulus. Later, when the first stimulus is perceived again, the animal starts another timer, and decides when to respond by comparing the ratio of the elapsing interval to the remembered interval to a certain threshold. In RET, the decision whether or not a stimulus merits a response is based on an animal's growing certainty that a stimulus has a substantial effect on the rate of reinforcement. By including SET and RET, Burke and Blumberg went beyond "traditional architectures that integrate an analysis of the past with the ability to react to the present, and include a representation of the future". As such, the approach is of an anticipatory system and the use of both models endow a synthetic creature with the ability of predicting and planning for future events by discovering causal relationships in the world.

6.3 Emotivector

6.3.1 Architecture

The emotivector is an anticipatory mechanism extending the agent architecture of Russell and Norvig (1995) and is located on the pathway from the sensors to the

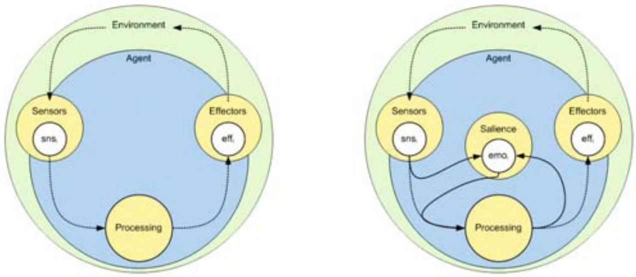


Fig. 6.1 On the left, Russel and Norvig's agent architecture, composed by a set of sensors sns_i sending encoded signals to a processing module that controls a set of effectors eff_i acting on the environment. On the right, the same architecture enhanced with a module composed of several *emotivectors*.

processing module. Figure 6.3.1 shows the agent architecture and the emotivector enhancement.

The emotivector observes the signals flowing from the sensors to the processing module[1] and, from this observation, anticipates the next expected value. The mismatch between the expected and sensed values generates an *affective* signal that is sent along with the sensed value.

In other words, the emotivector is a simple anticipatory mechanism coupled with a sensor that: (1) monitors the value of the sensor and predicts its next state; (2) determines the affective state that arises from the mismatch between the prediction and the sensor input value; (3) sends this information along with the sensor value. When a value from the sensor reaches the processing module of the agent, the tag provides a recommendation such as 'this signal value is much worse than expected: you should look at it carefully', or 'nothing new here: it is slightly becoming brighter, as expected'. The processing module of the agent can then take these recommendations into account for its further processing.

We will first present the emotivector's anticipation model then describe each one of the two main components of the affective signal: *salience* and *sensations*.

6.3.2 Anticipation Model

The computation of the emotivector relies on its capacity to predict the next sensor value. Each time a new signal reaches the sensor, the emotivector computes the sensor next expected value. The predictor implements an hybrid algorithm based on the Kalman filter (Kalman, 1960) and the generalized recirculation algorithm (Hinton and McClelland, 1988), whose learning rate is mediated by the current affective state (i.e. experiences associated with intense affective states will have a greater

[1] For the sake of simplicity, all signals flowing from the sensors to the processing module are one-dimensional signals within the interval between 0 and 1.

influence on the prediction). This algorithm provides the emotivector with a simple and efficient generic predictor which does not require any initial fine-tuning to work. More details on the implementation of the prediction algorithm can be found in (Martinho and Paiva, 2006b).

6.3.3 Salience Model

When a new value reaches the emotivector, it is confronted with its previously computed expected value. The emotivector then estimates the signal a-priori salience by computing two components inspired by Posner's two-system model of attention (Posner, 1980): the exogenous component and the endogenous component. The *exogenous* component is inspired in bottom-up, automatic reflex control of attention, and emphasizes unexpected values of a signal: the greater the difference between the expected value and the input value, the greater the exogenous salience. The *endogenous* component is inspired in top-down, voluntary control of attention, and emphasizes the closeness of a signal value to any actively searched values. In other words, if the agent is looking for something, similar things will likely attract its attention: the closer the value of the signal gets to a "desired" value, the greater the endogenous component will be. The interaction between both components define the attention grabbing potential of the signal, by adding or subtracting a component from the other (Müller and Rabbit, 1989).

6.3.4 Sensation Model

The existence of a desired value for a sensor allows the emotivector to associate a certain 'quality' to the signal in the form of a basic affective state: a *sensation*, in the sense given by Harlow and Stagner (1933). To generate such a sensation, a model inspired in early behavioural theories of emotions is used. The sensations are defined across two dimensions, as in (Young, 1961): valence and change. The model works as follows. The emotivector anticipates a reward (if the prediction is closer to the desired value than the current value) or a punishment (in the opposite situation). When the 'real' reward or punishment reaches the emotivector, it is confronted with the reward or punishment expectation. As a result, and following an approach inspired in the behavioural synthesis of Hammond (Hammond, 1970), one of the four following basic sensations is triggered: S+ or positive increase, if reward is stronger than expected; $+ or positive reduction, if reward is weaker than expected; S- or negative increase, if punishment is stronger than expected; and $- or negative reduction, if punishment is weaker than expected. Figure 6.2 shows the four sensation model and their eliciting conditions.

6.3.5 Selection Model

When all signals from the sensors finally reach the processing module of the agent, each one with a tag from the associated emotivector, how does the agent select which are relevant?

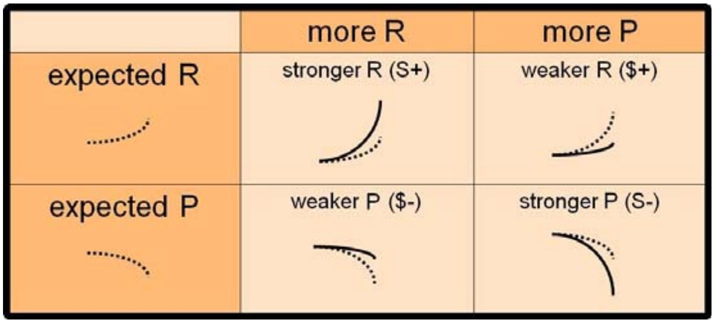


Fig. 6.2 Four-sensation model: S+ (positive increase) if the reward is stronger than the anticipated reward, $+ (positive reduction) if the reward is weaker than the anticipated reward, S- (negative increase) if the punishment is stronger than the anticipated punishment and, $- (negative reduction) if the punishment is weaker than the anticipated punishment.

Different strategies can be used. However, and to avoid the limitations of the winner-takes-all strategy, the low efficiency of salience order processing, and the inelegance of setting salience thresholds, *error prediction* was added to the emotivector (for a discussion of the selection strategies, please refer to (Martinho, 2007)). A second (identical) predictor was added to the emotivector which receives the error prediction each time a new value enters the emotivector and is compared with the expected value. Based on the history of error predictions, and using the same algorithm as the first predictor, this second predictor estimates the next prediction error. When the 'real' prediction error is greater than the estimated prediction error, the emotivector marks the signal as *relevant*.

6.3.6 Uncertainty

Error prediction provides an error margin for the estimation of the next sensor value. Each time the first predictor estimates the next value of the sensor, the second predictor estimates how thrustworthy this prediction is, by estimating the prediction error. As such, the emotivector is able to model, in a certain sense, the *uncertainty* associated with the prediction.

The introduction of uncertainty had an interesting side-effect on the sensation model: it allowed to extend the four-sensation model to a nine-sensation model. In the four-sensation model, when the emotivector is expecting a reward (i.e. is expecting the value of the signal to move closer to a desired value), the outcome can only be two-fold: the reward can be better (a S+ sensation) or worse (a $+ sensation) than expected. With the uncertainty associated with a prediction, the outcome can now be three-fold: 'significantly better than expected', 'significantly worse than expected' or 'better, as expected'. This approach more than double the sensations automatically generated by the emotivector. Figure 6.3 represents the nine-sensation model.

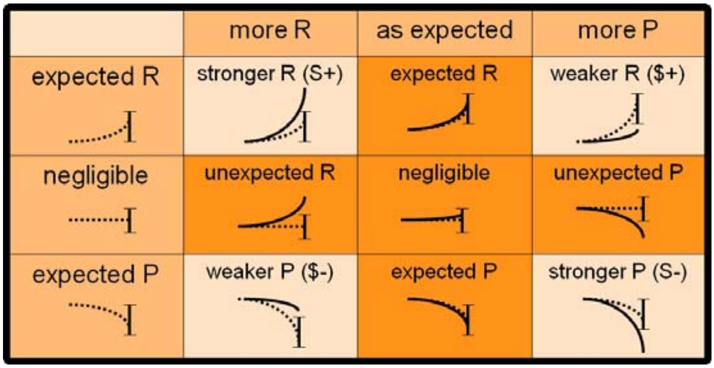


Fig. 6.3 Nine-sensation model. In the figure, *R* stands for reward and *P* for punishment. The first line of the figure shows the three possible outcomes when the emotivector is expecting a reward: 'significantly better than expected', 'better, as expected', and 'significantly worse than expected'. The use of uncertainty allowed for the introduction of five new sensations (represented by darker cells). From top to bottom, and left to right, they are: 'reward is as good as expected', 'unexpected reward', 'no significant reward nor punishment, as expected', 'unexpected punishment', and 'punishment is as bad as expected'.

6.4 Aini, the Synthetic Flower

Aini is a synthetic flower "living" in a stretch of shallow water, a synthetic character created to help us evaluate the adequacy of the emotivector mechanism in the creation of believable behaviour. Aini is a situated, embodied virtual agent which dynamics are controlled by a physics engine. A detailed description of the implementation of Aini's virtual body can be found in (Martinho and Paiva, 2006a).

An interactive task was designed to test the emotivector model: a word puzzle game in which Aini helps the user to uncover a four-letter word by reacting consistently to the user's actions in the virtual environment. To complete the word, the user has to place a set of letter cubes onto wooden platforms representing the word letters and their relative position in the word. To allow for a more immersive interaction and account for situatedness and embodiement, the simulation is fully tridimensional and physics controlled. Figure 6.4 shows Aini in her virtual environment. For a fully detailed description of the experimental settings and the challenges involved in such a game, please refer to (Martinho, 2007).

6.4.1 Emotivectors in Action

For this experiment, a grid of 25 emotivectors was used to build Aini's synthetic vision system (see Figure 6.5). This approach allowed behaviours such as 'casually look around' and 'quickly look at something new' to emerge from the interaction with the user. For more details, please refer to (Martinho, 2007).

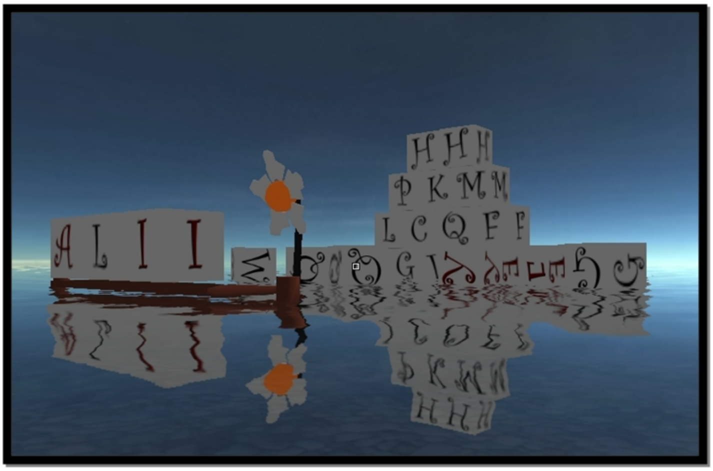


Fig. 6.4 Aini and the Word Puzzle Game. The three last letters of the word ('I', 'L' and 'A') are uncovered (they are read from Aini's perspective). Aini is currently expressing an 'unexpected reward' sensation towards the last introduced letter ('L'). The user had been playing around with the cubes for so much time (instead of performing the task) that Aini was not expecting any progress to happen so soon.

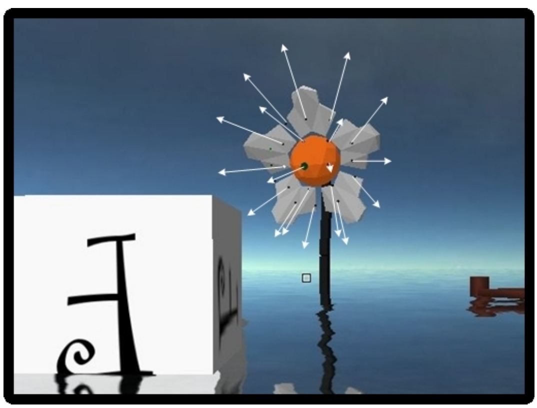


Fig. 6.5 Aini synthetic vision system. Aini's synthetic vision system is implemented as a 5 x 5 sensor grid. Each sensor is associated with an emotivector and measures the distance to the nearest object in a predefined direction within Aini's field of view. At any time, the focus of attention is decided based on the salience of the grid emotivectors, while the intensity controls the quickness of the animation.

Additionally, a single emotivector was associated with the progress of the task[2] and generated the task related sensations, each one mapped onto particular animation parameters. An interesting fact is that, for the user, these sensations seemed to account for the history of the interaction. This is best understood through an example.

Consider the case of a user trying to uncover the word 'ZEAL'. Aini starts in a neutral state. The user starts playing around in the environment rather than to concentrate on the task. Because no change occurs in task progress, the prediction error drops, representing the fact that Aini is pretty convinced that nothing is going to change relatively to the task.

When the user finally places her first letter ('L') in the wrong place, Aini expresses an 'unexpected punishment' sensation. Aini stops moving and lowers its 'head' (a similar expression is represented on the right of Figure 6.6). Aini now expects more 'punishment' to come although she is very uncertain about it.

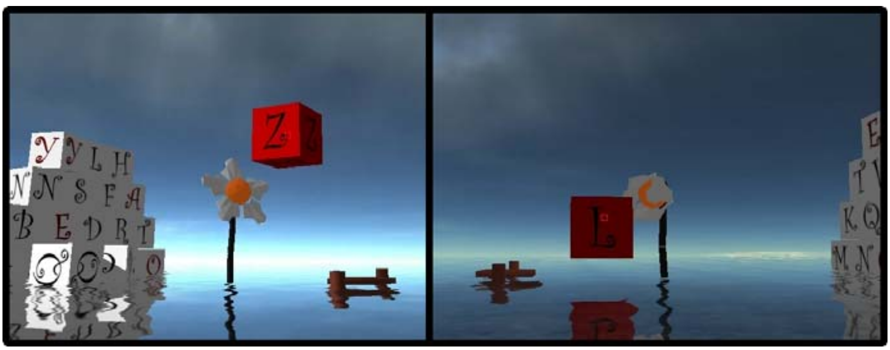


Fig. 6.6 Aini's affective expression. For the purpose of the experiment each one of the nine sensations generated by the emotivector was directly mapped into animation parameters. The picture on the left shows a positive sensation being expressed while the picture on the right depicts the expression of a negative affective state.

Influenced by Aini's negative expression, the user removes the letter 'L' from the word. As the task progress increases by a significant amount, Aini expresses a 'weaker punishment' sensation: it rises its 'head' and waves encouragingly (a similar expression is represented on the left of Figure 6.6). It is important to note the following: if the user decides to place the letter 'L' again in the same (wrong) placeholder, Aini will react differently from the first time, as the emotivector inner state has changed in the meantime.

The user now places the 'L' letter cube in the correct position. As there is still a great uncertainty associated with the emotivector prediction, an 'expected reward' sensation is triggered and Aini expresses her confidence regarding the user progress in the task. Afterward, the user places the 'A' letter in the right position. As the

[2] The progress of the task is measured from '0' ('none of the letters belong to the word') to '1' ('all the letters are in the correct position') using an approach inspired in the game 'mastermind' (M. Meirovitz, 1971). The initial value is '0.5' ('no letters on the wooden platforms').

margin of error has now dimished, the value entering the emotivector is outside the error prediction margin, and a ‘stronger reward’ sensation is triggered, making Aini express total bewilderment to the user.

This simple example shows how rich an interaction can become by simply mapping the emotivector basic sensations onto a set of affective expressions. The history of the interaction and the timing of the different actions will make the emotivector trigger different basic sensations in response to a same user action.

6.4.2 Evaluation

The experimental assessment of believability is a non-trivial task, mainly due to the subjective nature of the concept of believability itself. In the synthetic character field, believability is usually evaluated by asking the subjects to answer a questionnaire evaluating their satisfaction regarding the interaction with the synthetic character. To help evaluating believability, the experiment was designed in such a way that for the user to finish the task, the behaviour of the synthetic character had to be emotionally consistent (i.e. how the synthetic character express its affective state should be consistent with the user actions) and demonstrate intentionality (i.e. how the character expresses its affective state has to be consistent with its perceived intentions).

Subjects were asked to perform a series of word puzzles with four synthetic characters sharing the same graphical appearance but behaving differently. Two acted as control characters and the two others evaluated the emotivector approach, comparing it to another approach using in the current generation of computer games. More than 280 word puzzles were played by more than 60 male and female subjects from 5 to 79 years old, and with different computer skills.

The results confirmed the adequacy of the emotivector approach: no subject was able to finish the game with the control characters, suggesting that it is impossible to finish the game by brute force alone, and while all subjects succeeded with the emotivector-based synthetic character, only 20.6% finished the game with the game-based approach.

The experiment revealed three interesting results. First, that emotivector based synthetic vision provides with a natural form of interaction (e.g. by waving an object to draw the character’s attention to it). Second, that even a single emotivector (in this case, the emotivector connected to the task progress) can create rich and non-repetitive behavior that seems to account for the history of past interactions in a meaningful manner. Third, that the emotivector based behavior may, in certain situations, outperform significantly the approach used in the current generation of computer games. A detailed description of the results can be found in (Martinho, 2007).

6.5 iCat, the Affective Game Buddy

Social robots are robots especially designed to interact with people, helping them to perform tasks in a plethora of diversified environments. One particular domain of application of such robotic characters is as pedagogical agents that work as assistants to a learning task and provide immediate feedback regarding the learning experience, with the purpose to improve the learner's performance. Furthermore, the introduction of embodiment in learning environments positively affects the perception of the learning experience, especially when such synthetic characters convey emotional responses to the tutoring situation (Breazeal, 2003). This section focuses on a specific type of learning environment: turn-based educational games.

An experiment was designed based on the hypothesis that if the synthetic character acts as a tutor (or game companion) to the user with whom it is interacting, and its emotional behaviour reflects what is happening in the game, then the user will be able to better perceive the game state and her performance will increase. To generate such emotional states, and control the consequent emotional expression of the robotic character, the emotivector mechanism was used. If successful, such an approach would suggest that any game for which it is possible to provide an evaluation of the game state at a certain time, from the point of view of one player, could be coupled with a robotic character, that could automatically display believable behaviour, strengthening what is occuring in the game.

To evaluate such an approach, a chess scenario was developed where users could play a game of chess against a iCat robot from Philips using an electronic chessboard from DGT Projects (DGT, 2007). The iCat is an available plug-and-play robot capable of mechanically rendering facial expressions and was designed to simulate human-robot interaction under the perspective of social robotics (Breemen, 2004). Figure 6.7 shows the system setup and portays typical interactions occuring during a chess game.

6.5.1 Emotivectors in Action

The system is composed of two main parts: the chess subsystem and the emotion subsystem. The *chess* subsystem contains the interface with the electronic board and a chess engine (Kerrigan, 2007) used to evaluate the board state and compute the iCat's next move(s). The *emotion* subsystem is responsible for managing the character's emotional state. It receives information from the chess subsystem, and sends animation commands that blends the prescripted animations and behaviours in the iCat platform. In other words, the emotion module receives the evaluation of the game state performed by the chess engine, and generates the affective state of the robotic agent, which is composed of two components: *instant reactions* and *mood*, inspired by Scherer's work (Scherer, 2000).

Instant reactions refer to the relatively brief episodes of an external or internal event as being of major significance. They have a short duration but are quite explicit. Instant reactions can be associated with previous expectations, particularly in a turn-based game, where we inevitably build an idea of the opponent's performance, and tend to anticipate her performance during the game. In the chess scenario, instant reactions are generated by an emotivector mechanism, which input is the evaluation of the board by the chess engine after the user has played. Each one

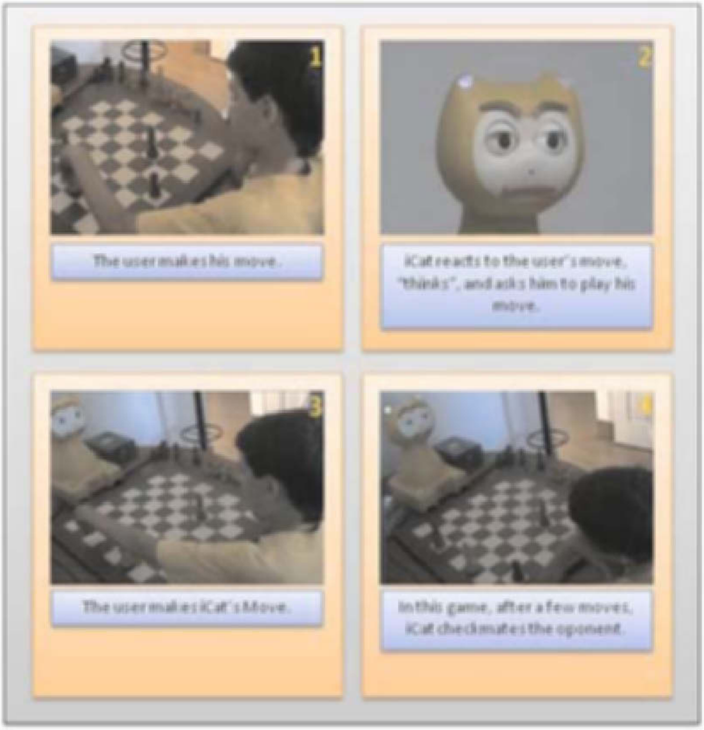


Fig. 6.7 Interaction during a chess game

of the nine sensations is mapped into predefined animations, which are blended in the iCat robot.

Mood is a relatively lasting affective state, less specific, often less intense, and thus less likely to be triggered by a particular stimulus or event. Mood is an unidimensional variable depicting intensity and valence and is computed as a function of the current game state. In other words, the better the evaluation of the game from the iCat perspective, the more positive the mood will be, and conversely for the negative mood. The mood value is used to control some paramenter used by the idle animation that are played in between the instant reactions triggered by the user's plays.

6.5.2 Evaluation

To test the effect of the iCat's emotional behaviour based on the emotivector, an experiment with 9 participants with ages ranging between 7 and 31 years old was conducted. The main focus of the evaluation was to find out if the iCat's emotional behaviour had any impact on the user's perception of what was happening on the game.

To measure the success of the user perception of the game, what the user "thinks" about the game at a certain moment is compared with the value obtained from the chess engine's evaluation function. As such, at a certain board position, if these two variables match with each other (e.g., if the user thinks that iCat is loosing and the chess evaluation function also indicates that iCat is in disadvantage), the user is considered to have successfully perceived the game state.

The experiment was conducted with three different control conditions regarding to the iCat's emotional behaviour:

1. The behaviour generated was in agreement with the emotivector model;
2. The emotional behaviour was "incoherent" and random. In this case, the emotional expressions to the user's move are randomly chosen between all but "coherent" responses;
3. Without expressing the emotional state, that is, a neutral/idle, behaviour.

These three conditions were used with all the users (in different orders) in three different exercises (of different difficulties). At the end of each exercise, the experimenter asked the user what was the state of the game and that perception was recorded. As such, the tests provided a sample of 27 values for each one of the three control conditions.

The results were as follows: in the "neutral" condition, 20 out of 27 were successful at perceiving the state of the game; in the "random" condition, 16 out of 27 were successful; and in the "emotivector" condition, 23 out of 27 were successful. Thus, the success measure varied among the three different control conditions, and, the results are better when the iCat exhibited the affective behaviour described (emotivector based). These results suggest that such behaviour helps the users to better understand the game.

A Spearman correlation test with a two-tailed test of significance for the samples of each one of the three control conditions was performed with three different variables to correlate: (1) the user's perception of the game based on the iCat's expression; (2) the user's perception of the game based on her overall analysis and (3) the "actual" game state, obtained from the chess evaluation function. The main result found was that the correlation between the "user's own analysis of the game" and the "actual game state" variables is higher when the iCat is controlled by the emotivector mechanism, with correlation of 0.930 ($p < 0.001$). With the random emotional behaviour samples, the correlation decreases to 0.485 ($p = 0.010$) and in the games that iCat did not express any affective state the value is 0.680 ($p < 0.001$). This result suggest that the user's perception of the game increases when the iCat's emotional behaviour is in agreement with the actual state of the game.

Regarding to the correlation between the perception of the game based on the iCat's expression and the user's own analysis of the game, correlations between these two variables were found in two of the three control conditions. When using the values from the emotivector system, such correlation is really strong (0.958 for $p < 0.001$), whereas in the neutral emotional behaviour the value decreases to 0.580 ($p = 0.002$). Despite the decrease in the value, the variables still remain correlated with the neutral behaviour. This may result from the fact that users may interpret the iCat's neutral behaviour taking into account their opinion in what is happening in the game. However, as far as the random condition is concerned the variables are negatively correlated (-0.116), although with a poor significance ($p = 0.564$).

6.6 Emotivector Integration in Agent Architectures

Although designed to assist in the automatic generation of believable behaviour, the emotivector is a low level anticipatory mechanism working at the sensory interface level, and as such can be integrated in higher level agent architectures. An example is provided by Piunti et al. (2007d), where the notion of subjective expected utility (SEU) guiding the action-selection process was substituted by an affective expected utility (AEU) which incorporated the affective information provided by the emotivector. Basically, the AEU added a modulating term that would reinforce a certain stimulus in the case of a positive feeling and diminish it in the case of a negative feeling.

The design, encoding and testing of the integrated system showed that, although the AEU would not affect performance significantly in the case of static trends in stimuli, when seasons and sinusoidal progress of stimuli were used, the AEU agents outperformed the SEU agents. Piunti et al. (2007d) point out that the advantages of using the emotivector were two-fold. First, the emotivector assessed a better prediction model for signals that evolve continuously over time. Second, the emotivector contributed with a relevant affective bias, as the sensations generated by the emotivector would provide the agent with a modulated motivation to decide which is the most hopeful area to explore. In opposition to the SEU agents, the AEU agents could distinguish the case in which a great likelihood was coupled with a low utility to the case where a high utility is coupled to a scarce likelihood. Additionally, the AEU agents could use the additional affective component indicating how good is the feeling toward a certain choice, thus anticipating the potential affective consequences of alternatives.

6.7 Conclusions

In this chapter, we discussed the relation between anticipation and believability, and how emotions are an inherent part of the discussion. We saw how the pioneering efforts addressing the creation of believable behaviour emphasized the importance of traditional animation, namely the importance of emotions. We also discussed how anticipation, although an important concept in traditional animation techniques, has had but a secondary role in the field of synthetic characters, and presented some of the few recent works that consider anticipation as a critical feature in the generation of believable behaviour.

We discussed how the relation between anticipation and believability was researched by the authors to create the emotivector, a simple anticipatory mechanism aimed at assisting the generation of believable behaviour for synthetic characters implemented as software agents. An emotivector, when coupled with a sensor: (1)

monitors the value of the sensor and predicts its next state; (2) generates an affective state that arises from the mismatch between the prediction and the sensor input value; and (3) sends this information along with the sensor value. When a value from the sensor reaches the processing module of the agent, the emotivector affective tag provides a recommendation such as 'this signal value is much worse than expected: you should look at it carefully', or 'nothing new here: it is slightly becoming brighter, as expected'. The processing module of the agent can then take these recommendations into account in the selection of action, namely in the control of behaviour related with emotional expression.

We presented two scenarios depicting possible applications of the emotivector mechanism in the generation of behaviour perceived as believable by the user, and their evaluation. The first scenario showed how the emotivector was successfuly used to control a realtime situated embodied character inhabiting a virtual world, Aini, that interacted with a user to solve a word puzzle game. The second scenario showed how the emotivector was successful in controlling a robotic character acting as a game companion, the iCat robot, in a pervasive game of chess. The results of the evaluation of both scenarios suggest that, in some applications, the emotivector mechanism is able to generate affective states that can be used to control the affective expression of a synthetic character, and that such behaviour is believable and understandable by the user. Furthermore, because it is a low level anticipatory mechanism working at the sensory interface level, the emotivector can also be integrated in higher level cognitive architectures. We presented a successful example of such an integration and the impact on this particular system.

The context-free nature of the emotivector allows it to be portable to different contexts, in both virtual and real world environments. For instance, the same system used for the chess game could a-priori be applied to any game in which it is possible to obtain the (eventually heuristic) evaluation of the state of the game for one of the players. Using such an evaluation, the game companion would exhibit automated believable affective behaviour, without needing to know the peculiarities of the game it is playing.

It is important to remember that, at the core of the emotivector mechanism, is *anticipation*: the affective state generated by the emotivector results from the mismatch between sensed and *expected* values, and the mismatch between sensed and *expected* errors in prediction. As such anticipation plays an important role in the automatic generation of *believable* behaviour by the emotivector.

Chapter 7
Anticipation and Emotions for Goal Directed Agents

Emiliano Lorini, Michele Piunti, Cristiano Castelfranchi, Rino Falcone, and Maria Miceli

7.1 Introduction

Breakthrough challenges in the cognitive modeling are in providing artificial agents with abilities to operate in *open systems* where dynamism, partial knowledge and non-determinism exact agents not only to to quickly react to events but also to *anticipate* decisions, facing with uncertainty and unpredictability of future events. One of the key issues is then to enhance cognitive reasoning with emotions and affective abilities. The rationale behind the introduction of emotions into artificial agents is manifold. Since the grew up of cognitive sciences and applied psychology during 60s, the negative bias against emotions, and more generally against 'irrational' or 'not reason-based' responses, has been practically reversed. The functional value of emotions has now been widely acknowledged and their evolutionary contributes have been especially emphasized (Simon, 1967; Fridjda, 1986; Parrott and Schulkin, 1993; Lazarus, 1991). From the perspective of biological evolution, emotions can be considered psychological mechanisms that evolved to solve adaptive problems (Toby and Cosmides, 1990) (i.e. escaping threats or predators, finding food, shelter and protection, finding mates) and thus surviving and delivering one's genes to one's own offspring. According to this view, emotions mediate behaviors for organisms in order to enhance long term adaptation and to answer their recurrent ecological demands. Otherwise, emotions provide evolutionary solutions to many of the critical problems implied by agents' situated interactions with their environments, for instance by enhancing proactiveness, by favoring the adaptive allocation of bounded computational resources, by providing anticipatory mechanisms for adaptation to mutable contexts, etc.

Traditionally AI emphasized individual problem solving in closed domains, where agents face with short term interactions and are assumed to deal with a bounded rationality and a narrowed number of goals. Taking the perspective of *Affective Computing*[1] has forced designers to deal with different domains, where cog-

[1] The term is introduced by Rosalind Picard (1997).

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 135–160, 2008.
© Springer-Verlag Berlin Heidelberg 2008

nitive agents are assumed to juggle deliberation between multiple goals, facing with heterogeneous influences, uncertainty and multiple problem solving styles. Most recognized role played by emotions for agent cognition is their informative function exploited for agents adaptation. The idea is that, to be adaptive in open environments and dynamic contexts, artificial agents strongly require a proactive adjustment of their internal model over time, in order to adaptively assess resources, exploit informational feedbacks, learn from experiences and become better at achieving goals. To this end, emotions serve as a control mechanism for cognition may accomplish the functions to help the cognitive system in arbitrating different goals in uncertain conditions, by informing the underlying reasoning processes when some particular event is requiring servicing (Simon, 1967). In general terms, emotions can play the pivotal role of connecting the various layers of a cognitive architecture, allowing suitable functional shortcuts for reactivating components (or information) which were not available, or under attention, at a given moment. Anyhow, one of the neglected functions ascribed to emotions in computational models of cognitive systems is precisely in *enabling anticipation*. As we will explain in this chapter, emotions play many pivotal roles for anticipation, i.e. in building expectations, in pro-actively responding to events, in allocating computational resources, in assessing knowledge and regulating purposes, especially when they are relevant to the concerns and well-being of the organism. A substantial part of this chapter is devoted to define goal directed architectures enabling anticipation through emotion and expectation processing. An ontology of anticipatory mental states (e.g. predictions, expectations, etc.) is provided to clarify what cognitive anticipation exactly is. This part of the chapter is fundamental as a conceptual basis for the implementation of expectations and emotions in a cognitive system. Focusing on goal directed systems, we will mainly deal with cognitive *expectations* and their related emotions such as surprise, disappointment and relief. In particular, we will focus on mismatch based emotions (in cognitive reasoning, discussing them as mechanisms playing a pivotal role in learning, attention, belief revision and action execution. In order to functionally assess the dealings between emotion and anticipation we will refer on their manifold causal relationships. Following the theoretical assessment provided by Castelfranchi and Miceli (ress) and strongly focusing on the various aspects of cognitive expectations, we will consider the following themes:

- The mediating roles of emotions between the behavioral stimulus-response attitudes elicit forms of anticipation producing *emotion-based expectations*. On the one side preparatory emotions may trigger anticipation which is not based on explicit predictions of future states and events. On the other side, premonitory emotions may accomplish the function of signalling and activating latent mental states and expectations which are physically represented within agent's internal states.
- Besides, the esteem of expectations and predictions of future states upon which an agent may be concerned with, can trigger emotions[2]. This kind of *expectation-*

[2] Notice that, differently from cognitive appraisal on ongoing events, here the appraisal is anticipated and rely on predictions of future events.

based emotions can be generated by the invalidation of an expectation, thus a mismatch between what was expected and what is perceived in a given instant of time.

- A third class of interaction between emotions and anticipatory abilities assesses the *anticipation of future emotions*. Rather than on events and future states to face with, here the anticipatory ability handles representations directly referring on expected emotions, which may strongly affect agent's decisional processes and strategies.

By adopting a functional approach based on expectations processing, we will discuss a set of important behavioral and mental changes as byproducts of expectation invalidation, from which surprise originates. Besides, we will claim that there is not just one single form of surprise, but several forms of surprise should be distinguished on the basis of cognitive appraisal involving current mental states. Furthermore we will survey computational models converging with recent studies that are pointing out the enhancement of adaptive abilities in introducing expectations and emotions in reasoning of artificial, goal directed agents.

The reminder of this chapter is organized as follows. Section 7.2 presents recent advances in the field of affective computing, discussing the peculiarities of the different approaches. In section 7.3 we discuss the typology of expectations and predictions proposed by Lorini and Castelfranchi (2007). A related typology of surprise is given in section 7.3.2. Section 7.3.3 is devoted to clarify why surprise is so important in cognition. Section 7.3.3.1 focus on a specific role of surprise in cognition by presenting the computational model of surprise developed by Lorini and Piunti (2007). In particular, we describe surprise-based filter mechanism that is responsible: 1) for signaling the inconsistency between beliefs and an incoming input which is relevant with respect to the current task to be solved; 2) for the revision of beliefs and expectations on the basis of the incoming relevant information. Section 7.4 describes a computational model for goal directed agents, exploiting mismatch based emotions like surprise for enhancing their anticipatory abilities. The manifold relationships between anticipation and emotions are described along with the discussion of the architecture. Finally, in section 7.5, we provide discussion and concluding remarks.

7.2 Related Works in Affective Computing

Even if many of the instrumental relations between complex emotions and their function within the mind are far to be explicitly formalized and appreciated, numerous theoretical models have been proposed during the last decades dealing with affective reasoning. Many of these models of emotions have been coherently formulated in order to be tightly implemented with a computational system, and rapidly imposed their contributes to the research community in the field of cognitive modeling and Artificial Intelligence. Recent computational systems for deliberative agents are enucleating the cognitive processes underlying emotions and their functional roles in driving intelligent agents. Unfortunately, a series of critical issues arise when emotions were applied in practice. In turn, standard goal directed architectures provide arbitrary strategies and typically are implemented using highly domain-specific patterns, thus resulting slightly scalable, monolithic, not flexible enough to embed *anytime* and domain independent processes. The wide part of these approaches are limited to the implementation of domain specific or rule based systems, aimed at enhancing adaptive capabilities in dynamic environments. More than on the key aspect of anticipation, the actual computational models mainly focus on a causal description of basic emotional processes and attempt to implement this description more or less directly in agents. Simple affective states are generally placed in terms of their effects on agent's reasoning. Various influences on behavioral and attentive activities are described, while emotions are generally coupled with agent control mechanism in order to tight different processes and computational modules underlying reasoning.

Different solutions span from reactive methods of control, similar to those employed in primitive biological organisms and artificial life (Scheutz and Sloman, 2001), to the affective control of computational resources inspired by Simon (1967) and the decision making processes affected by emotions (Doyle, 1992; Gmytrasiewicz and Lisetti, 2000; Lowenstein and Lerner, 2003; Busemeyer et al., 2003).

The computational model given by Macedo and Cardoso (2001, 2004) proposed a solution for exploration of unknown environments with motivational agents where surprise govern the intentions and 'action-goal' processes, thus eliciting action-selection through evaluation of utility functions.

The remarkable work by Scheutz (2002, 2004a) studied artificial agents implementing emotional control mechanisms and incrementally define a framework for embodied agents, where basic emotions are argued to enhance their adaptiveness in ecologically inspired tasks. Scheutz works further defined the general, domain independent principles for the evaluation of the effectiveness and the utility of emotions through extensive computer simulations (Scheutz, 2004b).

A particular class of emotion driven architectures are those based on *appraisal theory* (Arnold, 1960; Frijda, 1987; Lazarus, 1991). They provide emotions to coordinate the different computational and physical components required to effectively interact in complex environment. Emotional signals are generally used as a causal precursor of the mechanisms to detect, classify, and adaptively respond to significant changes of environment. Early approaches used domain dependent schemes and rules to derive and support to appraise events and govern action selection and planning strategies (Elliot, 1992; Moffat and Frijda, 1995).

In their Extended Mental State Framework (EMSF), Correa and Coelho proposed a goal directed architecture where goals are defined with different attributes (i.e. 'importance', 'intensity', 'insistence' and 'urgency' with different degrees) and included in higher level Mental States, defined as informational structures informing agents behavior and allowing to relate situations, in the world, to actions (Correa and Coelho, 2004). In more detail, Mental states are placed through programmable structures of rules and constraints, supporting complex forms of social organization and enhancing agent interactions with the environment.

Gratch and Marsella (2006, 2004) presented a domain independent computational model stressing the many different causal relations between emotions and cognition. Their computational model EMA placed cognitive processes and emotions in mutual relations: emotions arise from an evolving subjective interpretation of agent relation whit his environment either affecting either being affected by cognitive states and behavior. To process emotions, EMA defines a domain independent taxonomy of appraisal variables and produce emotions by processing the causal attribution of chains of events. By comparing ongoing beliefs, desires and intentions with external events and circumstances, emotions are elicited in terms of appraised variables, namely a superset of rules and criteria placed in terms of *desirability* (goal importance) and *likelihood* (probability of a given event). Once appraised, emotions are then responsible for two kind of coping strategies: 1) *problem-focused coping strategies* modifying agents behavior in terms of action selection, planning and allocation of resources; 2) *emotion focused coping strategies* acting on internal states and used for causal reinterpretation, shift of motivations, belief revision, goal reconsideration etc.

An alternative approach is posed by evolutionary models based on connectionist architectures which are introducing emotions as particular signals to be handled within sub-symbolic processes. Busemeyer and Johnson (2004) presented a neural network system for decision making, where the accumulation of affective evaluations produced by actions execution are exploited in a wide variety of cognitive tasks such as perception, categorization and memory processing. As reported in chapter 2, emotions like surprise and curiosity have been variously adopted in goal oriented systems to govern action execution: Schmidhuber (1991c, 2007) considered curiosity as sources of 'internal motivation' either to explore the environment and improve relevant information either to enhance the repertoire of actions aimed at achieving goals. Besides, recent tendencies in machine learning are exploiting *anticipatory* and *intrinsic* rewards (Singh et al., 2005), in order to allow the system to adapt to circumstances and particular contexts, regulate the trade-off between exploration and exploitation, proactively explore and learn more efficiently. Ahn and Picard (2006) proposed a cognitive system with affective and anticipatory abilities, modeled by positive and negative appraisal of expectations and where extrinsic rewards (rising from external goal, or costs) are integrated with intrinsic rewards (rising from internal circuits, emotions and motivations).

7.3 Expectations and Surprise

7.3.1 A Typology of Expectations and Predictions

The expectation system of a cognitive agent is an amazingly complex system whose operating characteristics vary enormously across time and specificity, at both conscious and unconscious, and learned and innate levels. Indeed, the crucial feature of cognitive agents is their being pro-active not only reactive due to their anticipatory representations. Cognitive agents have the ability to deal with the future by mental representations or specific forms of learning. For guiding and orienting the action a representation of the future and more precisely a representation of future effects and of intermediate results of the action is needed. To have a mind means to have anticipatory representations, i.e., predictions and goals; not just perception, beliefs, memory.[3]

Lorini and Castelfranchi (2007) have identified a rich typology of expectations. The typology is the one given in figure 7.3.1. The most general distinction is between *low-level* and *high-level* expectations and predictions. *Low-level expectations* and predictions correspond to sensory motor expectations and predictions based on some form of statistical learning on frequency and regular sequences, on judgment of normality in direct perceptual experience, on the strength of associative links and on the probability of activation (Kahneman and Miller, 1986). Low-level expectations and predictions play a prominent role in automatic behavior (Norman and Shallice, 1986) where primitive forms of anticipation are involved (Stolzmann, 2000). On the other hand, *high-level expectations* and predictions have many different sources: from analogy ("The first time he was very elegant, I think that he will be well dressed") and, in general, inferences and reasoning ("He is Italian thus he will love pasta"), to natural laws, and - in social domain - to norms, roles, conventions, habits, scripts ("He will not do so; here it is prohibited"), or to Theory of Mind ("He likes Mary, so he will invite her for a dinner; He decided to go in vacation, so he will not be here on Monday"). These forms of high-level expectations and predictions are involved in intentional behavior and deliberative activity.

The category including high-level expectations and predictions is then refined by the identification of specific sub-species of anticipatory mental states. *Scrutinized* expectations and predictions are distinguished from expectations and predictions *in background* (passive expectations) which operate at an unconscious and automatic level (Kahneman and Tversky, 1982). On the one hand, scrutinized expectations and predictions occupy consciousness and draw on the limited capacity of attention. They are coupled with the current intentions of the agent. For example, while trying to find a cheap flight from Rome to London in the Ryanair website and having the intention to do this, an agent consciously expects to find a cheap flight from Rome to London. Scrutinized expectations and predictions are endogenous anticipatory explicit representation of the next input which have to be matched with the incoming data. On the other hand, expectations and predictions in background are available at a mere automatic and effortless level. Either they can be the product of priming[4] or they are part of a *presupposed mental framework* supporting the agent's scrutinized expectations and predictions. There are expectations and predictions constituting the *presupposed mental framework* of the agent which have the form of *conditional*

[3] As previous works have already clarified (Miceli and Castelfranchi, 2002; Castelfranchi et al., 2003) expectations have to be distinguished from mere predictions (i.e. beliefs about the future). Expectations are the functional coupling of a belief on the future with a goal concerning the contents of that belief. The presence/absence of this motivational component marks the distinction between mere predictions and proper expectations.

[4] Several empirical evidences exist showing that, in being active and available at an automatic and effortless level, background (passive) expectations can affect subject's performances and judgments and can conflict with conscious (scrutinized) expectations (see for example (Sommer et al., 1998)).

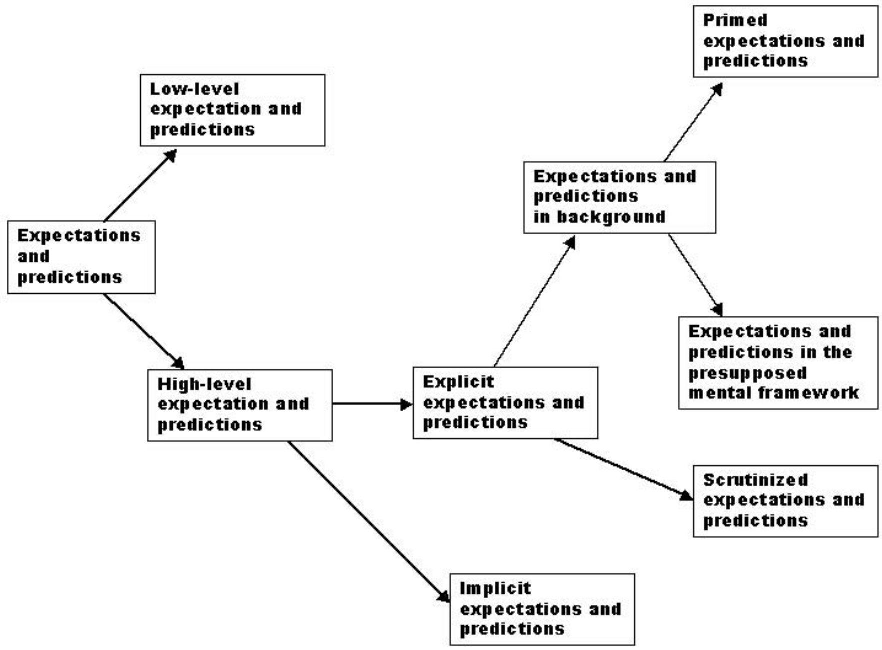


Fig. 7.1 Typology of expectations and predictions

expectations and predictions.[5] They encode a causal knowledge used by the agent for interpreting the context where his action and perception are situated. For example, while trying to find a cheap flight from Rome to London in the Ryanair website and having the intention to do this, an agent consciously expects to find some cheap flight from Rome to London (i.e. the agent has a scrutinized expectation to find some cheap flight from Rome to London). Such a scrutinized expectation is supported by a background belief of the form "I believe that I have entered into the Ryanair website" and by a conditional expectation in background of the form "I expect that if someone enters into the Ryanair website then he will find some cheap flight from Rome to London".[6]

Scrutinized expectations and predictions, and expectations and predictions in background are included in the more general category of *explicit* expectations and predictions, that is, all those expectations and predictions which are available to the agent either at a conscious level or at an automatic and effortless level. Explicit expectations and predictions are members of the agent's explicit knowledge

[5] The notions of conditional expectation and conditional belief have been extensively analyzed in analytical philosophy and AI (Boutilier, 1996; Rott, 1989; Stalnaker, 1981).

[6] These conditional expectations and predictions are generated on the basis of the agent's knowledge encoded in scripts and frames (Schank and Abelson, 1977). For example, the agent's conditional expectation "I expect that if someone enters into the Ryanair website then he will find some cheap flight from Rome to London" is generated from the agent's script of the Ryanair website.

base including also explicit beliefs about the present, explicit beliefs about the past, explicit assumptions, etc... Finally, explicit expectations and predictions are distinguished from *implicit* expectations and predictions, that is all those potential expectations and predictions that can be inferred from the agent's explicit knowledge base (Ortony and Partridge, 1987; Levesque, 1984).

7.3.2 From the Typology of Expectations to the Typology of Surprise

Surprise is perhaps the most primitive form of emotion and is tightly related with expectations and predictions. In fact, surprise is the emotional response which is associated with the invalidation of predictions and expectations. Surprise is a felt signal (Reisenzein, 2000) which provokes an immediate reaction/response of alert and arousal due to an inconsistency (discrepancy, mismatch, non-assimilation, lack of integration) between an incoming input and our predictions and expectations. It invokes and mobilizes resources at disposal of an activity for a better epistemic processing of this strange information (attention, search, belief revision, etc...), but also for coping with the potential threat (Lazarus, 1991). Surprise is aimed at solving the inconsistency and at preventing possible dangers (the reason for the alarm) due to a lack of predictability or to a wrong anticipation.

In agreement with the model of cognitive surprise we developed in (Lorini and Castelfranchi, 2007, 2006), we conceive surprise as an expectation-based cognitive phenomenon[7] playing a fundamental role in expectation dynamics. On the basis of the typology of expectations and predictions sketched in the previous section 7.3.1, a related typology of surprise can be given. A specific type of surprise can be associated to each type of expectation and prediction and different levels of surprise can be defined. Our typology of surprise is the one sketched in figure 7.3.2.

Low-level surprise is conceived here as the most peripheral form of surprise, due to the perceptual mismatch between what the agent sees and his low level (i.e. sensory-motor) expectations. This forms of surprise is distinguished from forms of *high-level surprise* due to symbolic representations of expected events, and to the process of information integration with previous long-term knowledge and to the explanation of the perceived data (Meyer et al., 1997). Hence, several sub-species of *high-level surprise* are defined. *Mismatch-based surprise* is due to a recognized inconsistency between a perceived fact and a scrutinized expectation. An agent can have an anticipatory conscious representation of the next input (i.e. a scrutinized expectation) and can try to match the incoming data against it. If there is a mismatch between the two representations, then surprise will arise. For example, while trying to find a cheap flight from Rome to London in the Ryanair website and consciously expecting to find some cheap flight from Rome to London, an agent might discover that there are not cheap flight from Rome to London. From the mismatch between the agent's expectation and the incoming input surprise will arise.

[7] See also (Casati and Pasquinelli, 2006) for an expectation-based model of surprise.

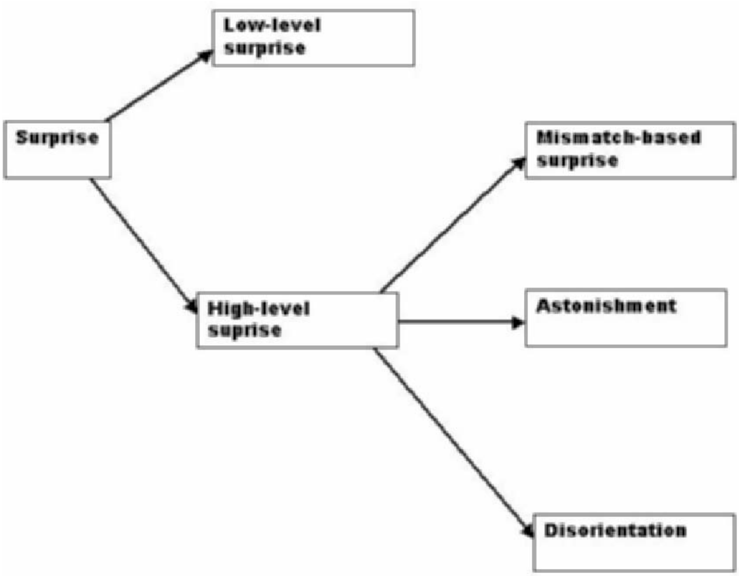


Fig. 7.2 Typology of surprise

Besides, we call *astonishments* those forms of surprise due to the implausibility and unexpectedness of the new input. When an agent is astonished about something, he cannot believe what he sees and this presupposes that he is trying to believe, he is trying to find an explanation for what he sees, but he is suspended. Astonishment is due to a difficulty, to a delay due to this process of integration, of accounting for, which in this case is not automatic and fast, not immediately successful. We cannot in fact believe something just putting it in our belief base; we must check about consistency (especially if there are reasons for suspecting some inconsistency). If the actual input generates an intense astonishment then it means that the input is unexpected and rather unpredictable from the agent's expectations and predictions in background or from the agent implicit expectations and predictions. For example, after having typed the Ryanair website's address "www.ryanair.com" in his browser in order to connect to the Ryanair website, a message could appear saying "Ryanair website has been definitively closed since the company went bankrupt". This will induce an intense astonishment in the agent since either, from his implicit knowledge, the agent is able to infer that the fact "Ryanair went bankrupt" is quite implausible or he presupposes that "Ryanair did not go bankrupt and the Ryanair website is still accessible" (given that he intends to enter into the Ryanair website).

Finally, we call *disorientation* the form of surprise due to the the collapsing of the presupposed mental framework of the agent. More precisely, *disorientation* is due to the invalidation of conditional expectations and predictions constituting the presupposed mental framework of the agent. As noted in the previous section 7.3.1, such conditional expectations and predictions are used by the agent for categorizing the whole situation or context in which his action and perception are situated. When one of those conditional expectations and predictions is invalidated, the agent has to reinterpret the entire situation and context. He is not just surprised but disoriented. For example, when entering in a restaurant a person assumes to be entered in a restaurant, and assumes to be entered in a normal restaurant with no elephants, no lions or other strange things. Indeed, a normal person has a conditional expectation of the form "I expect that if I enter into a restaurant then I will not see a lion". If the person enters into a restaurant and sees that there is a lion, then he will be completely disoriented because of the invalidation of his conditional expectation which will lead him to reconsider the mental framework used for interpreting the whole situation.

7.3.3 Roles of Surprise in Cognitive Processing

As discussed in section 7.3.1, a cognitive agent has sensory-motor expectations (low-level expectations) which play a prominent role in automatic and routinized behavior. Given a certain stimulus *S*, the agent selects a certain response *A* with the sensory-motor expectation of the reward he will obtain by doing A. An unexpected (positive or negative) reward of the performed action A can be responsible for generating *low-level surprise* (as defined in section 7.3.2) and for adjusting reward predictions. This has been confirmed by empirical researches. In (Schultz, 1998) the neural correlates of such surprise mechanisms are studied. The advanced hypothesis is that dopamine could function as a signal of unpredictability of actions and then be required for learning novel actions. These neural mechanisms on which surprise is based signal error in reward prediction and can be used for learning the value, positive or negative, of an agent's actions. Surprise also plays a crucial role in neuro-computational models of action control. According to these models, after a system has *selected* some motor programs for achieving some desired state *S*, he *anticipates* the effects of the selected motor programs (motor prediction), that is, he forms some sensory-motor expectations of the next input. If the some sensory-motor expectations is invalidated by the perceived input, surprise can intervene during performance and be responsible for a correction of the ongoing action. For instance, suppose that an agent is trying to grasp a moving object in front of him. Appropriate motor programs are selected (e.g. moving the hand with a certain angle, direction, velocity; bending the elbow in certain way; etc.) and the effects of selected motor programs are anticipated in such a way that the body can be adjusted in order to succeed in grasping the object. For example, while trying to grasp the object the agent has sensory-motor expectations of the spatial position $Pos_{Expected}$ of the object in the next future. If the perceived position $Pos_{Perceived}$ of the object turns out to be different from the expected position $Pos_{Expected}$ of the object, surprise will arise. Such a surprise is responsible for a correction of the ongoing action of grasping the object (e.g. after being surprised the agent changes the direction of the movement).

A further functional role ascribed to surprise is its function for a shift of attention (Botelho and Coelho, 1997; Baldi, 2004; Itti and Baldi, 2006). This is part of the immediate response and short-term function of surprise. The felt feedback of surprise (the feeling of surprise) is responsible for redirecting attention towards unexpected stimuli, and for concentrating cognitive resources on them and for interpreting them. In this sense surprise is crucial for *learning* in cognitive systems (Berlyne, 1960). In the reinforcement learning community in the last few years, several authors have been devoted to encompass traditional reinforcement learning models in order to endow cognitive systems with *intrinsic motivations* such as curiosity and surprise (Singh et al., 2005; Schmidhuber, 1991c). Curious agents have been designed in such a way that they shift their attention towards salient stimuli and initiate an exploration those parts of the environment which turn out to be novel. This line of research in the RL community takes inspiration from some recent studies in neuroscience (Dayan and Balleine, 2002; Kakade and Dayan, 2002) where it has been shown that dopamine also plays a critical role in the intrinsic motivational control of behaviors associated with novelty and exploration. For instance, salient, surprising sensory stimuli inspire the same sort of phasic activity of dopamine cells as unpredicted rewards.

Related to the previous roles of surprise, is the function of surprise for a shift from an automatic (reactive) level of performance to a deliberate level (Ortony et al., 2005). In fact, when an agent is engaged in an automatic and routinized activity, the violation of his sensory-motor expectations of the effects of his actions or the violations of his reward predictions can be responsible for a shift to the deliberate level consisting of a new planning process. For example, while trying to grasp an object an agent has a sensory-motor expectation $Temp_{Expected}$ of the surface temperature of the object. In the normal case the agent expects that the object will not be red-hot. If the agent's sensory-motor expectation is violated by the input $Temp_{Perceived}$ (e.g. the agent touches a red-hot object), low-level surprise arises and the agent's behavior might be interrupted. After this the agent might be deliberately replan and choose a different course of action (e.g. it might decide to grasp the red-hot object with tongs).

The crucial role of surprise in higher forms of learning has been stressed in cognitive and experimental psychology. With higher forms of learning we mean those forms of learning which consist in a revision, update, change of high-level predictions, expectations, beliefs of a cognitive agent (as defined in section 7.3.1). As some psychologists have stressed (Meyer et al., 1997) surprise often culminates in a process of belief and expectation change. More precisely, surprise is often a signal and causal precursor of a process of belief reconsideration. In fact, when agent's high-level expectation is violated, *high-level surprise* (as defined in section 7.3.2) arises and the agent becomes aware of the fact that he has wrong knowledge of the environment. This is the preliminary step towards a revision of the agent's preexistent high-level predictions, expectations and beliefs.

In what follows we refer to autonomous and proactive software agents, in particular we refer to a strong notion of agency, dealing with a special kind of Goal-Directed entities using an explicit internal representation for their goals (purposes) and their beliefs (knowledge). Such a kind of cognitive systems is assumed to create, update and manipulate a symbolic representation of the world through mental states. Differently from adaptive and merely *goal-oriented* agents who try to adjust their epistemic representations to the world in order to make them as accurate as possible, a true *goal-directed* agent modify the external world, through its activities and according to his endogenous representation, trying to make it as close as possible to the desired states. Whereas the assessment of the belief base determines agent's internal states, the internal goal representation makes it possible to determine the 'end state' of a given activity, as far as to deliberate between concurrent goals and decide the one to adopt. Typically, goal directed agents use a reasoning process, including a deliberation engine to resolve and select the goal to adopt among the applicable ones.

The next two section are devoted to discuss a computational model of surprise we developed by Lorini and Piunti (2007) and Piunti et al. (2007b,d). In such models we devise out some of the functional roles that surprise and expectation invalidation emotions play with respect to anticipatory abilities.

7.3.3.1 Surprise-Based Belief Update: A Computational Model

In Lorini and Piunti (2007) the role of high-level surprise in belief change is investigated and a computational model of surprise-based belief change is developed. In such a computational model - which takes inspiration from the general theory of surprise presented in Lorini and Castelfranchi (2007) and briefly described in the previous section 7.3.2 -, the notion of mismatch-based surprised is operationalized. Mismatch-based surprised is conceived as a filter mechanism which is responsible: 1) for signaling the inconsistency between beliefs and an incoming input which is relevant with respect to the current task; 2) for the revision of beliefs and expectations on the basis of the incoming relevant information. The computational model consists in the operationalization of two general hypothesis. On one hand, it is supposed that at each moment an agent is focused and allocates his attention on a particular task that he is trying to solve. That is, the agent has a certain number of scrutinized expectations and, as noted in section 7.3.1, each scrutinized expectation is coupled with an intention representing the pragmatic solution that the agent has selected in order to accomplish the task (Bratman, 1987)[8]. Hence, the agent ignores all incoming input which are not relevant with respect to the current task on which he is focused and only considers those information which are relevant. On the other hand, it is supposed that if a relevant input mismatches with a scrutinized expectation of the agent, surprise arises. The surprise reaction is a causal precursor of a belief update process. In fact, a surprise with a certain intensity relative to the incoming relevant input "signals" to the agent that things are not going as expected and that beliefs must be reconsidered. The surprise-based mechanism of belief update is implemented in a $belief-desire-intention$ (*BDI*) architecture (Wooldridge, 2002; Rao and Georgeff, 1992) and the performances of a standard *BDI* agent and of a *BDI* agent endowed with a surprise-based filter of belief change (called *BDIS* agent) are compared. The control loop of the standard *BDI* agent is described in the

[8] A bayesian network is also used to encode the agent's causal knowledge of external environment and to define those conditional expectations and predictions that, as argued in section 7.3.1, constitutes the presupposed mental framework of the agent.

$\mathcal{BDIS}$ **agent control loop**	$\mathcal{BDI}$ **agent control loop**
1. $B := B_0$;	1. $B := B_0$;
2. $I := I_0$;	2. $I := I_0$;
3. while (true) do	3. while (true) do
4. get new percept Γ;	4. get new percept Γ;
5. if $\mathbf{S}(I,\Gamma,B) > \Delta$ then	5. $B := bu(\Gamma,B)$;
6. $B := bu^*(\Gamma,B,I)$;	6. $D := options(B,I)$;
7. end-if	7. $I := filter(B,D,I)$;
8. $D := options(B,I)$;	8. $\pi := plan(B,I)$;
9. $I := filter(B,D,I)$;	9. $execute(\pi)$;
10. $\pi := plan(B,I)$;	10. end-while
11. $execute(\pi)$;	
12. end-while	

Table 7.1 The two typologies of agents

right column of Table 1, whilst the control loop of the *BDIS* agent is described in the left column of Table 1. The formal description of the control loop of the standard *BDI* agent is similar to the one given in Wooldridge (2002); Rao and Georgeff (1992). In lines 1-2 the beliefs (beliefs about the present, as well as predictions) and intentions of the agent are initialized. The main control loop is in lines 3-10. In lines 4-5 the agent perceives some new facts Γ and updates his beliefs according to a function *bu*. In line 6 the agent generates new desires by exploiting his means-end rules. In line 7 he deliberates over the new generated desires and his current intentions according to the function $filter$.[9] Finally, in lines 8-9 the agent generates a plan for achieving his intentions by exploiting his planning rules and he executes an action of the current plan. The main difference between the standard *BDI* agent and the *BDIS* agent is the belief update part in the control loop. It is supposed that a process of belief update defined by the function bu^* is triggered in the *BDIS* agent only if the degree of mismatch (i.e. $\mathbf{S}(I,\Gamma,B)$) between the incoming input Γ and the scrutinized expectation of the agent associated with an intention in I is higher than a threshold Δ (line 5 in the control loop of the *BDIS* agent). In this sense, the *BDIS* is endowed with a cognitive mechanism of surprise-based belief change. In fact, this mechanism filters out all perceived facts that are irrelevant with respect to the current intentions and with respect to the current scrutinized expectations. Thus, the *BDIS* agent only updates his beliefs by inputs which are surprising and relevant with respect to his current intentions and scrutinized expectations. Differently, at each round the standard *BDI* agent updates his beliefs indiscriminately: for any fact he perceives, he updates his beliefs whether the perceived fact is relevant or not.

[9] Space restrictions prevent a formal description of the function $filter$ here (see Wooldridge (2002) for a detailed analysis). Only note that this function is responsible for updating the agent's intentions with his previous intentions and current beliefs and desires (i.e. $filter : BEL \times 2^{INT} \times 2^{DES} \mapsto 2^{INT}$).

7.4 Expectations and Emotions for Goal-Directed Agents

The most recognized approach in design of computational systems for affective agents is inspired by *Appraisal Theory* of emotions (see Arnold (1960); Frijda (1987); Lazarus (1991), among others). According to this model, appraisal processes are performed by agents to assess the relationship with their environment and external events (Elliot, 1992; Moffat and Frijda, 1995; Gratch and Marsella, 2006). In particular, appraisal allow agents to characterize the significance of events from a subjective perspective and can be used to respond (*coping*) to events by affecting agent internal states (i.e. knowledge, desires, intentions). In so doing agents mediate, with affective states, their interaction with the environment[10]. Whereas appraisal processes are generally implemented by assessing specific appraisal variables and then identifying causal chains that, from past events, lead to the current state, we here propose a different approach, embedding cognitive appraisal directly within agent reasoning. As in Piunti et al. (2007b,d), we identify two *integrated* levels of reasoning, involving cognitive, slow deliberative processes as well as fast automatic and associative ones. Both levels integrate various mechanisms required to manage expectations, used to assess alternatives and choices, and to direct cognitive resources towards anticipated events. In more detail, we include *high level, active and scrutinized expectations* and *background, passive expectations* (See Fig. 7.1). At the higher level we deal with expectations modulating decisions and thus goal deliberation: we include in the reasoning process a quantitative influence on the terms given by the expected utilities used for arbitrating between alternative courses of actions. These influences can be adjusted on the basis of past practices and enable agent to learn from experiences. Besides, in order to enhance agent's adaptiveness, we model situated, low level expectations to elicit interrupts of deliberative processes in order to control unexpected events requiring services and thus reconsider the course of agent activities. Typically, these particular kind of reasoning is not part of the specification of an agent in his purposive behavior, rather can be let to *emerge* on the basis of context information as a result of the interactions in the environment. This approach is intended at: *1)* Exerting a top-down modulation of emotional reasoning as a result of deliberative process and adaptive responses to relevant events and *2)* Integrating adaptiveness in decision making along with expectations and their causal relation with the subjective appraisal/evaluation of events.

7.4.1 Expectations and Decision Making

Goal directed systems refer to a strong notion of agency, where internal goal representation makes it possible to deliberate between concurrent goals and decide the one to adopt (Castelfranchi, 1998). In so doing, cognitive systems are able to create, update, manipulate a symbolic representation the world through mental states

[10] Lazarus indicated this particular process as a reflexive re-assessment of the internal state under context awareness rather than a explicitly deliberated process.

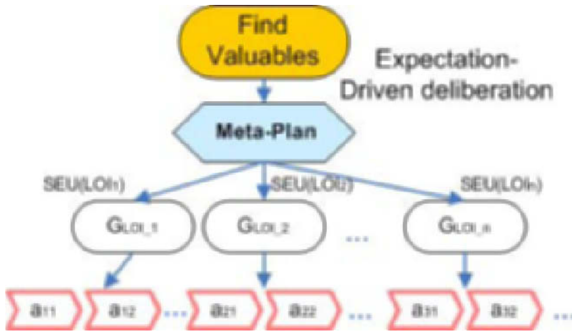


Fig. 7.3 Given a terminal goal, *Expectation Driven Deliberation* compares Subjective Expected Utilities to choose the most promising course of actions

(i.e. Beliefs, Goals). This approach is reflected by the wide adopted Belief-Desire-Intentions (*BDI*) model of agency (Rao and Georgeff, 1995; Georgeff et al., 1998), according to which an action is performed when the agent has an intention to achieve a given goal, and some beliefs indicating that the action helps in achieving that goal. Our model builds, on top of a *BDI* engine, an *expectation-driven* decision making, thus combining deliberative, logical aspects of a *BDI* model with more quantitative, numerical aspects of decision theory. To allow agents to take decisions based on scrutinized predictions and expectations we model a long term memory entertaining endogenous anticipatory representations. Each agent's (sub)goal is given along with the representation of its activation formulae (typically first-order belief formulae (Thangarajah et al., 2002)) and a network of inhibition links (indicating if a given goal has the priority on another goal and under which conditions this priority is applicable). Filtering can be managed through a dynamic arbitration network, providing disambiguation between the precondition rules and the resolution of the relative dependencies (inhibition links) between the concurrent goals. The *BDI* deliberation engine react to changes in the belief base (i.e. internal events thrown by a belief update) and use the current internal state to filter out enabling conditions for arbitrating the goal adoption (see Braubach et al. (2004) for more details). Imagine an agent being engaged in a *foraging task*: in normal conditions, the terminal goal is to look for valuables moving to a series of rooms towards some Location of Interest (*LOI*). Expectation-driven deliberation lets the agent to decide on which *LOI* to look for, considering how the various alternatives are 'promising' (Fig. 7.3).

As for the decision theoretic paradigm of 'rationality' (Savage, 1954; Doyle, 1992), an artificial agent may behave in order to maximize the expected utility, given multiplying utilities (desirability) and probabilities (likelihood). In our model this strategy is delivered at a meta-level reasoning, typically when the agent has to select between alternatives to achieve mutually exclusive sub-goals. Expectations are built upon two independent quantitative dimensions: *Belief strength*, as a degree of subjective certainty placed in terms of likelihood (the agent is more or less certain about their content) and *Goal value*, a subjective importance strictly dependent on desirability of the goal state and the related motivating forces, but also on context conditions and mental attitudes (Miceli and Castelfranchi, 2002; Castelfranchi et al., 2003; Castelfranchi, 2005).

Desirability is assessed in terms of utilities, and is coupled to rewards obtained upon goal completion: it is calculated according to the extent to which an intention (i.e. a given sequence of actions) has fulfilled a certain goal. This makes it possible to endow expectations with their *valence*: expectations can be considered *positive* (or *negative*) according to their contribution (or determent) to the ongoing intentions and mental states (e.g. Goals, Beliefs).

Likelihood are subjectively assessed as predictions. They are assigned through a forward model, able to quantitatively assess the subjective likelihood of a certain future state in domain of probability[11].

7.4.1.1 Emotions Modulating High Level Expectations

The fact that emotions influence decisions is widely acknowledged. A paradigmatic example is offered by neuropsychological studies on the role of emotions in decision making (Bechara et al., 1997). The model proposed by Damasio (1994) suggested the introduction of explicit *somatic markers* creating a sort of *veto* in the branching tree of alternatives, thus reducing the fan out of a possible decision. Otherwise, an emotion can inform cognition by affecting the desirability and the likelihood of an outcome. Early models indicated for emotions an additional motivation regulating utilities. Lowenstein and Lerner (2003) observed important distances between classical decision theory and emotional decision making. They explicitly introduced an anticipatory effects of emotion in regulating decisions. A similar solution has been formally proposed by Gmytrasiewicz and Lisetti (2000), while Busemeyer et al. (2003) formalized how needs change over time under the pressure of external 'stimulation' and internal 'deprivation'.

As far as our computational model concerned, we based emotion affecting decision through feedback signals of appraised mismatches upon purposive action completion. Given a scrutinized expectation upon a possible reward, our agents can appraise their experiences matching the expected utility and the effective achieved reward. While *Active perception* (see section 7.3.3.1) is used to update the knowledge model used for determining probabilities of future events and predictions, feedback of mismatches between expectations and experienced outcomes are used to experience emotions and adjust either utilities and predictions. As described in section 7.3.2, recognized inconsistencies between a perceived fact and a scrutinized expectation are at the basis of *mismatch-based surprise*. On the basis of a monitored signal (i.e. obtained reward) and a given expectation (i.e. expected reward), we recognized the following six cases of appraisal :

1. *Positive increase (S+)*: the agent is expecting a punishment but receives a reward. The achieved reward is stronger than the one expected. Can be related to to **excitement** or **positive surprise**.

[11] In the actual implementation, we are testing different mechanisms for unsupervised learning to determine conditional probabilities of future events given a sufficiently wide open knowledge base (i.e. EM algorithms for Bayesian networks (Dempster et al., 1977)).

2. *Negative increase (S−)*: the agent is expecting a reward but receives a punishment. The achieved punishment is stronger than the one expected. Can be related to **distress** or **negative surprise**.
3. *Positive reduction* (\$+): the agent achieve less reward than the one expected. Can be related to **disappointment**.
4. *Negative reduction* (\$−): the agent achieve less punishment than expected. Can be related to **relief**.
5. *No Surprise (NS)*: the reward matches the expectation and is exactly the one expected.
6. *Surprise due to ignorance (IS)*: the reward is not deducible from prior knowledge due to lack of experiences, not mismatch-based surprise

Once appraised, agent can use these feelings to update his expectations and thus to give more or less preference to the related alternatives, i.e. deciding towards which course of action to be committed in the next future. The experienced mismatch (surprise) enhances the importance of a certain goal G_i, hence the agent is biased to believe to fulfill more value from G_i achievement. We here have the expectation of a certain event (positive or negative, depending on its accordance with ongoing goals) that is eliciting an affective response. Therefore there is an expectation inducing an affective bias by which the agent introduces in decision making an *Expectation-based emotion* providing an intrinsic anticipatory effect[12]. Thus the decision, and the related goal adoption, is ruled on the basis of anticipated states (high level expectations) relying on the likelihood of the outcome (prediction) and on its desirability (goal importance). We defined *Affective Expected Utilities* (AEU) in terms of:

$$AEU(G_i) = \sum_{a_j \varepsilon Plan(G_i)} [A_b \times U(O_{G_i})] \times P(O_{a_j}|a_j) \tag{7.1}$$

where G_i is the i^{th} goal to adopt between candidates, O_{G_i} is its related outcome, $U(O_{G_i})$ is the subjective utility of that outcome, a_j the j^{th} action of the plan triggered by G_i and $P(O_{a_j}|a_j)$ is the probability of that outcome, given that the j^{th} action of the plan will have the proper O_{a_j} outcome.

Respect to the `subjective expected utility` originally proposed given in Savage (1954), we introduce A_b on the basis of a qualitative and a quantitative appraisal of the experienced mismatch. Differently from Damasio (1994) somatic marker hypotheses, A_b gives an additional, quantitative reinforcement into the deliberation process and further modulates the expected utility in affective terms. The branching factor of a decision is then re-modulated, whereas the alternatives acquire different weights due to appraised expectations. In particular, the positive increase ($S+$) and the negative reduction (\$−) of the monitored signal provide a positive indication about the progression of the goal value. Hence, when associated with a specific decision alternative, they present a *positive feeling* towards the related outcome. On the contrary, the negative increase ($S-$) and the positive reduction (\$+)

[12] Notice that differently from cognitive appraisal where a possibly updated belief induces an emotion (Lazarus, 1991), here the appraisal is projected to future events (Castelfranchi and Miceli, ress).

cause the agent to have a *negative feeling* towards that choice, thus inhibiting its value. This is implemented by reinforcing the utility of a choice with the additional factor, in case of a positive feeling, and diminishing it in the case of a negative feeling. A_b is positive for positive feelings and negative for negative ones:

$$A_b(G_i) = \begin{cases} 0.0 & \text{if } E_s(G_i) \text{ is in } \{NS, IS\} \\ (\gamma_+) * E_r(G_i) & \text{if } E_s(G_i) \text{ is a pos. feeling } \in \{S+, \$-\} \\ (\gamma_-) * E_r(G_i) & \text{if } E_s(G_i) \text{ is a neg. feeling } \in \{S-, \$+\} \end{cases}$$

where $E_s(G_i)$ comes from the last appraised mismatch on G_i's reward, $E_r(G_i)$ is the distance between expected reward and sensed reward, γ_+ and γ_- are discount factors (with $\gamma_+ \ll \gamma_-$).

A number of relevant remarks are worth making in this expectation driven deliberative process. First of all, the high level expectation and its invalidation (mismatch) play a pivotal role for affective states (either for the negative ones: disappointment, negative surprise either for the positive ones: relief, positive surprise). In fact, these emotions can't be elicited without anticipatory states. Mere goal fulfillment or frustration, if devoid of any specific mental state can elicit some emotion (i.e. sadness, joy). But no one can feel real cognitive relief (and excitement) unless a given prediction and concern with some uncertain goal, threaten to not come true. The same for disappointment and frustration, that can proper arise only if the goal importance is accompanied by a more or less certain prediction about its fulfillment, and this prediction has been invalidated. Cognitive Expectations emerges from the contemporary presence of predictions (beliefs on the future) and Goals (ongoing desires). Finally we assumed the intensity of the A_b as a function of its components: *a)* the more (subjectively) certain the prediction, the more intense the A_b and *b)* the more (subjectively) important (desirable) the goal, the more intense the A_b.

It is worth remarking further relations between decisions and emotions. For instance, considering emotions arising from counterfactual analysis of lost chances and alternative courses of actions leads to the further modulatory contribute of 'regret' (Gilovich and Medvec, 1994; Coricelli et al., 2007). As a further step one may involve in decisions not only the expectations about a given outcome, but also the associated *pre-felt* emotion. By arising from evaluating future consequences, expected emotions have been supposed to affect the terms of a decision (Lowenstein and Lerner, 2003; Castelfranchi and Miceli, ress). An expected emotion may induce a further change in agent goal ranking, whereas positive or negative expected emotions may induce a goal or reinforce its value (i.e. reinforcing importance for outcomes holding to positive emotions or inhibiting outcomes referable to negative ones). When the expected states also coincide with an expected emotion the agent is leaning to *as-if* reasoning. This would be possible for an agent which, by anticipating some future course of action, is also likely to anticipate that he would feel a given emotion.

7.4.2 Situated Agents and Affective States

A central claim of appraisal theory (Lazarus, 1991) is that emotions are associated with patterns of subjective judgment that characterize the personal significance of external events (e.g., was the event expected in terms of prior beliefs? is the event congruent with adopted goals?; is there the power to alter the consequences of this event?). For instance, coping strategies elicited by *surprise* can be modeled as a *momentary interruption* of deliberative and practical reasoning processes, e.g. diverting attention to past episodes or focusing sensors and effectors to a restricted area. Therefore, appraised events and elicited emotions can be used to activate background expectations dormant in agent's mind. The agent is activating emotions based expectations: affective states (emotions) have the role to activate mental states (expectations). Besides, events can be compared with agent goals and endogenously valued as *positive* (indicating that some event establishes the preconditions for achieving goals or create a new opportunity) or *negative* (some event represent a threat or thwart agent current goals). Here we are in the domain of *premonitory* and *preparatory* emotions (Castelfranchi and Miceli, ress), whose function is to provide some insight for inferring cognitive expectations signalling the need of behavioral changes.

Perceptive activities for situated agents require, at any instant of time, agent to sense context information and so have an up to date knowledge of what is happening in environment. As placed in Piunti et al. (2007c), perception and filtering are the epistemic processes responsible to store surprising events adding items to a Situated Associative Memory (SAM). As described in section 7.3.2, in this case surprise arises when the agent relieves a perceptual mismatch and then is used to activate background, passive expectations. The idea is to store, for each of these surprising events, informational reports in an associative memory. These reports contain descriptions of a defined set of situated properties: they have a symbolic representation including time-stamp, positive or negative valence of the originating event, location where the event has been detected and other specialized fields[13]:

```
evItem { valence: enum value="pos/neg"
         time-stamp: class="Time"
         location: class="Location"
         helps: class="Goal"
         thwarts: class="Goal"
       }
```

Once events are translated to their symbolic representation and stored in the SAM, they can be manipulated as percepts[14]. Henceforth, they can be exploited as a 'fast' source of information to adapt the behavior in the near future and anticipate world

[13] In the case of the scenario described in Piunti et al. (2007c), we distinguished negative events as harmful entities, fire threats, and positive events as food objects, valuables and *LOI* discovering.

[14] Perceptual items have a propositional content but a different nature respect to the beliefs. They are vague beliefs, only describing a situated event but not still adequate to define a causal envi-

changes. The intuition behind the mechanisms is provided by the well known principle of spatial and temporal locality, according to which one may assess that recently cached items of a certain class are likely to be retrieved in the near future. The amount of item locally present in the `SAM` can be used as an indication to infer latent, background expectations about the local context. It is worth pointing out that, once appreciated, these situated emotions have a mediating role between the stimuli and the reactions. Emotions here directly interfere and inform higher agent's reasoning (i.e. deliberation, intention reconsideration), in particular they force a bottom up interference with rationality [15]. The agent is using affective states as cues to his 'unconscious' assessment of the situation (Elster, 1996).

7.4.2.1 Functional Description

As in Schank and Abelson (1977), background expectations are generated on the basis of the agent's states and encoded in scripts and frames. This give to passive expectation a weak form of representation and an emergent dynamics, ruled by the assessment of ongoing affective states. Library of coping strategies, action alternatives and resource allocation strategies have been clustered within a discrete set of frames exploited as *control states*. Effects of coping are then modeled in different temporal scale, from immediate and short term reactions, to most persistent and

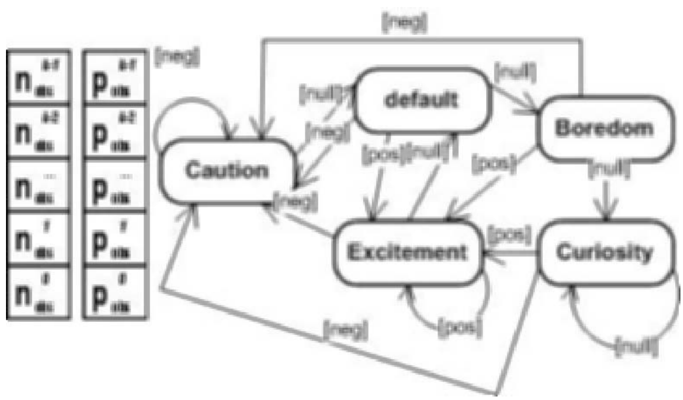


Fig. 7.4 Controller for Affective States: appraised positive (p) and negative (n) events are fed to a transition function in order to shift from different affective states.

ronmental relationship. Notice that situated percepts may hold to deceitful appearances, including false positive or negative items (Pollock, 1997).

[15] The mediating role of emotions can be explained in evolutionary terms. For instance, emotions have a *modulatory* and *energizing* effect on behavior, that may result more or less vigorous, more or less persistent (Dickinson and Balleine, 2002). Moreover, one may assess various stimuli to elicit the same affective state that in turns, will trigger the behavioral response. Another important function of emotions is for learning: they allow reinforcement since they can be pleasant or painful and thus remain associated to a given stimulus and enable faster reactions based on past experiences (Miceli and Castelfranchi, 2000).

long term effects. Given in functional terms, coping strategies includes emotional responses to overturn (in the case of negative emotions) or trigger (in the case of positive ones) control signals to be signalled to the reasoning process.

Emotions are thus assumed to monitor and signal goal pursuit, achievement and failure. Once an emotion has signalled the failure of a certain goal a behavioral response is elicited, which in turn may imply the adoption of some alternative goal. Besides, Affective States (ASs henceforth) are suitable control mechanisms for intention reconsideration. Traditional reconsideration strategies indicate an agent to abandon an intention when a related goal is achieved, when a goal become infeasible or when the agent relieve some inconsistencies between the world state and the external conditions necessary for goal achievement. Our model allows basic emotions to elicit an *interruption* on normal cognitive processes when unexpected events require servicing. Once based on expectations of future states, intention reconsideration becomes anticipatory and can be used to coordinate behavior with prediction of future states.

Furthermore ASs embed a particular kind of goal activation, bypassing the underlying deliberation processes normally used for practical reasoning. For example, on the short term a AS may attempt to resign the agent to a threat by signalling to the deliberative engine to abandon a goal (thus a related intention) that is becoming inconsistent with the actual belief base or the actual environment state. On the contrary, positive events may elicit goal activation to exploit new opportunities.

Each AS adopt a *context dependent* configuration of resources (i.e. vision, speed, perception rate, belief update). Becoming aware of his context, the agent can dynamically adapt his *control frame* in order to reduce performance payoffs and avoid wasting resources for useless activities. Control frames are characterized by the following tuple of dynamic values: $Cf = \langle En, r, Sr, s, G_a s \rangle$, En indicating the current amount of energy, r the range of vision where sensors can retrieve data, Sr the situated perception filtering rate, s the instant speed and $G_a s$ the goal to be activated in order to pro-actively respond to the situated events to cope[16].

Each frame defines the roles that the related ASs play for situated adaptation to contexts and environment dynamism. Imagine, in the foraging task described above, that the environment presents some threats for agent activities (i.e. the fires, adversary agents etc.). Once the agent has deliberated the best expected location to explore, through the evaluation of the sub-goal's related AEUs, it may happen the agent registers a close series of harmful (unexpected) events, i.e. fire collisions (Fig. 7.5.A). This may elicit the negative expectation that the agent is approaching to a dangerous area, and thus induce him to pass to a **Cautious state** (Fig. 7.5.B). This negative, passive expectation causes the agent to adopt a new goal, re-allocating his resources to cope harmful circumstances. Cautiousness causes changes both in the long and the short term: firstly it induces arousing by modulating attentive resources (i.e. enhancing Sr, looking ahead and augmenting r and reducing s, see Tab.7.2). A risk avoidance goal $G_a s$ interrupts the ongoing practical action to escape from threats and accordingly the agents arranges activities to better check the situation.

[16] We assume that agents spend energy and resources according to a combination of the previous costs (e.g. the higher the speed and perception-rate, the higher the consumed energy).

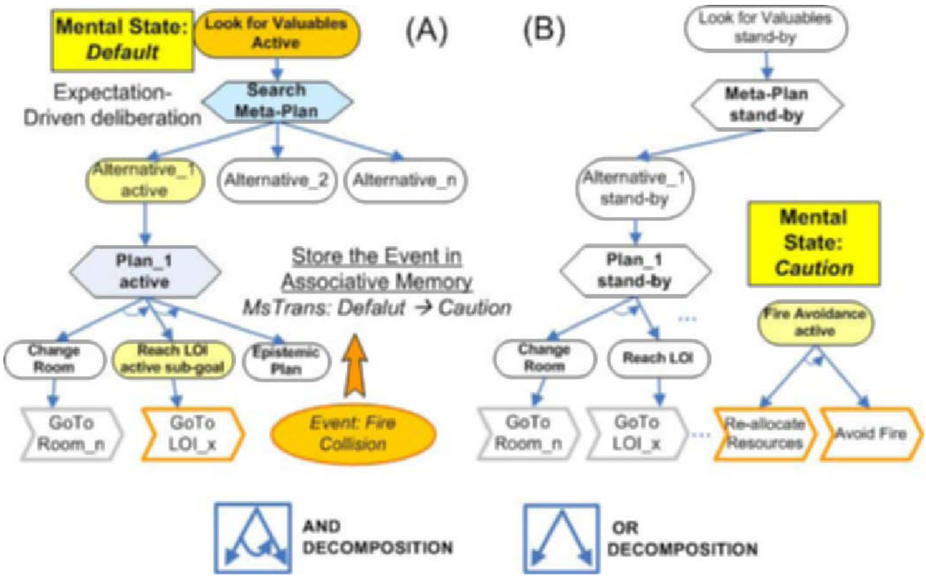


Fig. 7.5 Intention Reconsideration upon the activation of the Cautiousness Mental State

Table 7.2 Mental States elicit the adoption of control frames for attitudes, confidences and resource allocation strategies

AS	Mood	γ_{AS}	Resources		
			r	S_r	S
Default	Exploitation	1.0	.33	.33	.33
Excitement	Reinforcement	1.3	.275	.275	.45
Caution	Prudence	0.5	.45	.45	.10
Boredom	Exploitation	1.0	.33	.33	.33
Curiosity	Exploration	1.0	.45	.10	.45

On the long term, cautiousness brings to a watchful mood, by reducing the self confidence on beliefs (γ_{AS}), augmenting the control (e.g. enhancing perceptive iterations Sr) and/or performing the action in a less risky way (e.g. using safest alternatives in repertoire). Prevalence of positive surprising events induces the agent to shift to **Excitation**, that on the short term is used to arouse the agent, to augment epistemic activities and to search for those 'good' events. A positive surprise (i.e. valuables discovering) may induce the agent to abandon a previous intention and to reformulate his behavior to exploit the new opportunity triggering a new goal G_as. On the long term, excited agent adopt an 'optimistic' mood increasing the confidence (γ_{AS}) of those unexpected, positive events[17].

The lack of surprise progressively empties the SAM and reduces situated perceptive activities. In the long run, it produces a special frame: **Boredom**. Boredom

[17] Notice that, differently from Excitement emerging on appraisal of an achieved goal, Excitation has been related to a situated positive surprise.

indicates that the environment is almost stationary (no unexpected events are happening) and that the agent can fully exploit his purposive behavior governed by the deliberation driven reasoning. This enhances the subjective confidence in beliefs and in building predictions. Further persistence of boredom leads to **Curiosity**, a control state used to automatically arbitrate from exploitation to exploration activities. The exploration attitude is goal driven: once the agent does not recognize relevant events in his SAM[18], he may infer the low-level expectation that the environment is becoming more static, hence biases his activities towards actions that shows promise to perform a better field coverage and to maintain an updated knowledge. Bypassing the deliberation of practical reasoning, the curious agent pro-actively activates the epistemic $G_a s$ of exploring new rooms, searching for new facts and events. This has a twofold effect: on the one side it enhances territorial exploration augmenting the chances to discover new LOIs, on the other side it improves knowledge and maintains updated beliefs[19].

Transition Function and Information Fusion. `SAM`'s content is constantly monitored by an appraisal process in order balance the presence of items and thus decide which is the AS to adopt. Passing from one state another depends on how the events are relieved and appraised in execution time. This process can be described through a push down automaton (Piunti et al., 2007c; Gmytrasiewicz and Lisetti, 2000). Generally the agent supervises the buffers (through a background process) by balancing their registered contents: prevalence of negative items leads to passive expectation of undesirable states (i.e. contingencies, risks), hence to cautious attitudes, while positive events lead to positive expectations (i.e. opportunities) and excitement. For instance, the presence of negative items registering a close sequence of threats and obstacles may induce the agent to infer a low expectation of further threats and risks, thus to pass to a cautious mental state (Fig. 7.4). In more details, the current state is inferred by the previous state and the perceived input by a transition function $AsTrans : AS \times IN^{\star} \longrightarrow AS$, where AS is the set of definite affective states and $IN^{\star}$ the input events stored in the (possibly empty) `SAM`. $AsTrans$ realize an *information fusion* within the symbolic items. Notice that the presence of items of different nature may elicit inconsistencies to be resolved (i.e. presence of elements of different meaning as, for instance, interleaved sequences of positive and negative events). To address this problem $AsTrans$ uses a set of rules for combining and aggregating the items of the same type and circumvent the inconsistencies on the basis of the temporal sequencing given by the time stamps. As suggested by Cholvy and Hunter (1997), the rules used to govern the fusion can be composed of meta-level and domain specific information. For instance, a simple rule of *balancing* may assert to aggregate the items of a given typology, in order to circumvent the set of lower cardinality and to take into account only the information related to

[18] Heuristic thresholds define the k-length time window used for passing from Boredom to Curiosity.

[19] The benefits of interleaving epistemic and practical activities are generally accepted in situated cognition (Kirsh and Maglio, 1994). Different policies can be retrieved in literature to manage exploitation Vs. exploration. Among others, Ahn and Picard (2006) proposed an affective signal to abandon exploitation and trigger the process of exploration.

the bigger aggregate. By balancing the presence of items of a given type, the appraisal process suitably distinguishes between positive and negative expectations. A different approach was used in Piunti et al. (2007c), where two distinct buffers are handled to store positive and negative events and the current sate is let to emerge on the basis of the comparison of buffers sizes (Fig. 7.4). To prevent the agent to switch to an inconsistent state, the transition function is built to take into account a certain grade of *inertia*, thus providing more robustness against occasional events, false positive or negative items (i.e. due to noise or sensor faults etc.).

7.4.3 Confidence of Predictions and Modulation of the Probability Function

An additional mechanism is provided to handle the effects of ASs in order to reconcile the deliberative level, ruling intentions, goal adoption and action information, and the situated level, providing context information. The intuition behind this integration relies on the fact that each AS embodies a certain grade of self-confidence (due to the ongoing mood) that can be related to the belief base. Once we detailed beliefs with a certain strength due to uncertainty on the environment state, one may introduce the self-confidence as a *discount factor* to affect the likelihood of the predictions. In so doing, agents dynamically adopt a 'more or less confident' capability to build their predictions. For instance, positive moods as excitation can induce the agent to optimistically over-estimate the probability of a certain outcome. On the contrary, negative moods like cautiousness may introduce pessimistic under-estimations. On these basis, respect to the one given in (7.1), the affective expected utility results:

$$AEU'(G_i) = \sum_{a_j \varepsilon Plan(G_i)} [A_b \times U(O_{G_i})] \times [\gamma_{AS} \times P(O_{a_j}|a_j)] \tag{7.2}$$

where γ_{AS} is associated to the ongoing mental state (Tab.7.2). By associating a given confidence to the subjective capability to make predictions, γ_{AS} introduces a further affective modulation on agent rationality.

7.4.4 Discussion

As showed in the previous sections, the presented architecture for gaol directed agents introduces affective reasoning and enables different kind of anticipation based on surprise and mismatch based emotions. We first distinguished between long-term practical reasoning and situated reasoning, providing mechanisms for predictions on different time scale and based either on action information either on context information (see chapter 2). We discussed how the disambiguation of slow, decisional processes from situated ones elicits a clear methodological separation of concerns and may greatly assist the modeler by breaking down the work into two separate and independent activities: while the former is defined referring to explicitly represented mental states (related to beliefs and goal) and clearly involves decisional processes, deliberation and goal arbitration, the latter is defined by weakest representations (percepts) and can be defined through control frames, clustering domain dependent strategies, aggregates of heuristics and functional even affective responses used to anticipatorily react to local events. In a second phase, we reintegrated the two processes by taking into account the correlations and the relative interactions, enlightening how low situated reasoning can be used to inform higher reasoning and decisional processes. To this end the contribute of Affective States is twofold: from the one side they can relieve the deliberative and the attentive processes from the burdens to process weakly relevant information in decision processes, excluding action alternatives that are likely to be less promising or have vanishing likelihood to be achieved. Besides, ASs provide ready to use action selection and resource allocation policies that may relieve agent's need for resource-demanding and meta decision processes. The emergent nature of expectations and affective states enables agent to adopt ASs as control frames, while both expectations and emotions are conveyed to inform reasoning, for redirecting resources and adopt long term strategies once a surprising event is detected.

7.5 Conclusion

In this chapter we described some of the manifold relations between emotions and anticipation, enlightening some of the benefits which their computational processing can bring in a cognitive system. We mainly focused on the role of expectations, and their related emotions (i.e. surprise, caution), identifying several functional roles both at the level of automatic behavior (e.g. reactive behavior) where sensory-motor predictions are involved, and at the level of deliberated behavior, where high-level expectations, predictions and beliefs play a pivotal roles for goal oriented behavior, knowledge inference and decision making.

Accordingly, we surveyed some of the traditional systems dealing with affective computing. It is worth remarking that, whereas traditional models are aimed at enhancing agent adaptiveness to indeterministic and dynamic environments, we here adopt a different approach, enlightening the different processes by which on the one side emotions may produce anticipation (either *effect* anticipation either *start* anticipation) and on the other side anticipation may produce emotions. Besides, we introduced the motives and functional advantages in embedding emotions in reasoning process. In particular we envisaged those anticipatory abilities enabling agents to experience an enlarged set of emotions, and provided them with additional cognitive abilities with regards to decision making, attention and resource allocation. That is, *some kind of emotions cannot be experienced by systems devoid of anticipatory abilities*: this is the case, for instance, of relief, disappointment, surprise, caution etc. Moreover, the capacity for anticipatory representations enable agent to process expected emotions, hence allows forms of reasoning influenced by the anticipation of emotions.

In order to specify our anticipation based approach to computational emotions, we described two different goal directed systems implementing different genres of affective reasoning. The former architecture propose a computational model for surprise where the epistemic activities are governed by surprise as causal precursor of belief update. This allow intelligent strategy for resources allocation By balancing Epistemic Vs. Pragmatic activity Exploitable for in information-rich environments (i.e. information seeking agents).

Chapter 8
A Reinforcement-Learning Model of Top-Down Attention Based on a *Potential-Action Map*

Dimitri Ognibene, Christian Balkenius, and Gianluca Baldassarre

8.1 Introduction

How can visual selective attention guide eye movements so as to collect information and identify targets potentially relevant for action? Many models have been proposed that use the statistical properties of images to create a dynamic bottom-up *saliency map* used to guide saccades to potentially relevant locations. Since the concept of saliency map was introduced, it has been incorporated in a large number of models and theories (Rao and Ballard, 1995; Itti and Koch, 2001a; Rao et al., 2002; de Brecht and Saiki, 2006; Hoffman et al., 2006; Singh et al., 2006; Walther and Koch, 2006; Chen and Kaneko, 2007; Shi and Yang, 2007; Siagian and Itti, 2007). Saliency maps have shown to be useful both as models of human attention and for technical applications (Balkenius et al., 2004).

These bottom-up mechanisms have been enhanced with top-down processes in models that learn to move the eye in search of the target on the basis of foveated objects. In many of these systems, top-down attention is guided by task-related information that is acquired through automatic learning procedures (Dayan et al., 2000b). For example, Schmidhuber and Huber (1991) built an artificial fovea controlled by an adaptive neural controller. Q-learning was used in the model of Goncalves et al. (1999) to control attention based on multimodal input and reinforcement signals. Another model that uses reinforcement learning to control visual attention is described by Minut and Mahadevan (2001). In this model a first component learns by reinforcement learning to direct the gaze to relevant points in space, whereas a second component performs a "within fixation" processing directed to analyse the foveated space and identify targets. Reinforcement learning was also used by Shibata et al. (1995) to control the movement of a visual sensor over an image. The goal of the system was to find the optimal fixation point for object recognition. In this model, the same neural network was used both for object recognition and to produce the sensory motion output. Balkenius (2000) presented a model that uses instrumental conditioning as a basis for learned saccade movements. This model was later extended to support contextual cueing where several visual stimuli to-

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 161–184, 2008.
© Springer-Verlag Berlin Heidelberg 2008

gether suggest the location of a target (Balkenius, 2003). However, this model could only keep one potential target location active at each time.

Here we propose a novel model that improves on this type of top-down mechanisms by using an eye-centred *potential-action map* (PAM). The PAM keeps track of all the potential locations of targets based on the information contained in a *sequence* of fixations (cf. Chen and Kaneko, 2007). In this respect, the PAM works as a short term memory for potential target locations. Each fixation suggests potential locations for targets or other relevant cues and the evidence for each possible location is accumulated in the PAM. The location of the potential target locations are based on both the identity of the currently fixated object and its spatial location (Deco and Rolls, 2005). A shift mechanism triggered by eye movements allows the potential target locations activated in the PAM to be always updated with respect to the location of the *current* fixation (similar mechanisms might be used by real brains, cf. Gnadt and Andersen, 1988; Dominey and Arbib, 1992; Pouget et al., 2000; Di Ferdinando et al., 2004; Shadmehr and Wise, 2005). Overall, the PAM makes up an efficient mechanism for accumulating evidence for potential target locations in a action-oriented compact format readily usable for controlling eye movements. As we shall see, the results reported here indicate that, thanks to the PAM, the model suitably integrates bottom-up and top-down attention mechanisms and outperforms simpler models that only search for targets based on a single, currently foveated object.

In contrasts to the majority of models tackling the object-localisation tasks, the system proposed here was designed not only to find the target, but also to stay on the target once found. This is accomplished with multiple saccades that keep the eye's fixation point on the target. This combines the features of the cue-target based systems describe above and systems that are more directed toward tracking (e.g. Shibata and Schaal, 2001; Balkenius and Johansson, 2007). The idea underlying this functionality is that *vision serves action*, in particular that attentional selection is a precursor of action and it is intimately related to it (Allport, 1990; Ballard, 1991; Balkenius and Hulth, 1999; Castiello, 1999; Casarotti et al., 2003; Di Ferdinando et al., 2004). In this respect, the system presented here was designed to be used within a future architecture, which will guide a robotic arm engaged in reaching rewarded targets in space. As previous models (Ognibene et al., 2006; Herbort et al., 2007), within this architecture the targets of the arm's reaching movements will be selected on the basis of a neural competitions fuelled by the information flow coming from perception, in a way similar to what happens in the primate brain (cf. Cisek and Kalaska, 2005). With respect to this mechanism of action selection, the capacity of the attentional system to keep the fixation point on the target will allow the model to bias the competition between alternative goals of the arm's movements in favour of objects relevant to the system.

The rest of the paper is organised as follows. Section 8.2 will first illustrate in detail the architecture of the architecture proposed here and the detailed functioning and learning processes of its components, and then it will illustrate the tasks used to train and test the system. Section 8.3 will analyse in detail the function of the architecture's components, in particular how the potential action map can keep a memory of the information returned by cues and can integrate information on the target returned by several cues. Finally, section 8.4 will illustrate the strengths of the architecture and the limitations of it which will be tackled in future work.

8.2 Methods

This section first presents an overview of the system and its underlying assumptions, then explains the details of its different components and their functioning. The overall architecture of the system is shown in Fig. 8.1. It consists of the following parts which are all implemented as neural networks:

- An RGB input image, which is the sensory input of the system.
- A saliency map that selects targets for eye's movements by integrating excitatory signals from the periphery map (bottom-up attention map), excitatory signals from the potential action map (top-down attention map) and inhibitory signals from the inhibition-of-return map.
- A fovea, covering the central part of the input image, which is used for recognising objects.
- A periphery map, which plays a bottom-up attention function.
- An inhibition-of-return map, which prevents the eye from looking back to already explored locations.
- A potential action map, which is a top-down attention map that accumulates evidence for different locations where the target might potentially be found.
- A reinforcement-learning actor-critic model Sutton and Barto (1998), which allows the system to store knowledge on the possible (deterministic or probabilistic) relative spatial relations existing between different foveated objects and the target.

These components allow the system to explore new images on the basis of the bottom-up attention components. This attracts the eye to high-contrast areas while the inhibition-of-return components promote the exploration of areas with progressively lower contrast. With experience, the actor-critic components learn the spatial relations existing between the cues and the rewarded targets (Posner, 1980; Balkenius, 2000). While the system explores various targets, this allows the system to accumulate evidence for different potential target locations in the potential action map. This map plays the role of a working memory in which the identity of different objects explored over time can contribute with evidence for potential target locations relative to the currently fixated positions. This can be viewed as a form of what-where associations. Moreover, in the case of a rewarding target, the top-down attention process can also learn to override the inhibition of return mechanism to stay at a target one it has been localised. We now describe the functions of each of the components in detail.

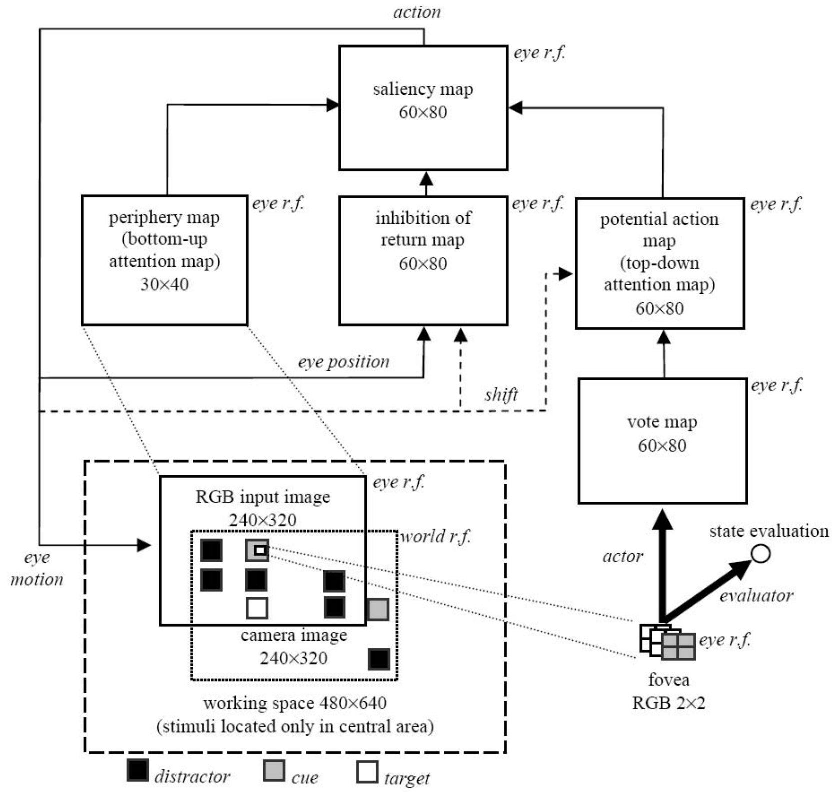


Fig. 8.1 The architecture of the system. The dashed box represents the work space, that is, the portion of environment that the eye can explore (the stimuli are presented only within the dotted-box sub-part of the working space). The plain boxes within the working space represent the periphery and the fovea input to the system. All other plain boxes represent different two-dimensional neural maps. The names and size (definition) of the maps are described in the boxes. Arrows represent information flows. Plain arrows represent one-to-one connections with unitary weight. Dashed arrows represent information that triggers a (hardwired) shift of the visual information in a direction opposite to the saccadic movement. Thick arrows represent all-to-all connections trained through reinforcement learning. The circle represents the output unit of the evaluator of the critic in the reinforcement learning component. “r.f” stands for “reference frame”.

8.2.1 RGB Camera Input

The camera input might be produced by a motorised pan-and-tilt camera simulating a moving eye. Here however, we use a stationary camera image and only simulate the eye movements. The work space that the eye can explore is an area formed by 480×640 RGB pixels. The objects relevant for the current tasks can appear in a sub-region of this space consisting of 240×320 pixels. Each object is a 20×20 pixel square uniformly coloured in either red, green or blue. The actual input to the system consists of a 240×320 pixel simulated camera with which the system explores the working space. The implemented system can also operate on a real camera, but this has not been used in the tests reported here.

All the different components of the system represent information in a eye-centred reference frame. This important assumption is based on the idea that the brain uses an eye-centered representation close to the sensory organs and puts off the computationally heavy remapping to motor coordinates until it is needed for motor control. This is an idea which is gaining increasing support within the neuroscientific literature on visuo-motor transformations taking place in parietal cortex (Shadmehr and Wise, 2005). The computational advantage of such deferred processing exploits that representations close to the sensory organs tend to contain much information, whereas later stages closer to the actuators use more abstract representations.

8.2.2 Saliency Map and Action Selection

A saliency map combines a number of visual feature maps into a combined map that assigns a saliency to every location in the visual field (Itti et al., 1998). Each feature map is typically the result of applying some simple visual operator to the input image. For example, a feature map could consist of an activity pattern that indicates all the vertical edges in an image. Other types of feature maps may code for intensity, color, motion or some other visual feature. The result of summing the different feature maps is that the saliency map will code for locations in the image with many features. For example, a region with many edges and bright colors will be more salient than a locations without any such features.

The idea of saliency map can be extended to include not only "bottom-up" information from feature maps, but also information from other sources. Here, the saliency map selects saccade targets by summing the topological input signals coming from three different sources (neural maps). The first source of input the periphery map which detects colored objects in the scene and issues bottom-up signals to the saliency map via topological one-to-one connections with equal fixed positive weights denoted by denoted by β. The second source is the potential action map which implements top-down attention and activates the saliency map through topological one-to-one connections with equal fixed positive weights denoted by τ. The final source of input is the inhibition-of-return map which encodes the last locations visited by the eye and activates the saliency map via topological one-to-one connections with equal fixed negative weights denoted by denoted by ι.

The choice of the saccade target is performed by selecting the position corresponding to the unit with maximum activation (the "winning unit"). During training, the units of the map were added random values before computing the wining unit. These noisy values were randomly drawn from a uniform distribution having a range $[-n, n]$ decreasing with time t:

$$n = \nu \cdot \left(1 - (t/T)^2\right) \tag{8.1}$$

where T is the duration of the training phase.

This process is a computational abstraction of the competition taking place in real brains between potential target stimuli (Koch and Ullman, 1985; Desimone and Duncan, 1995). Such a mechanism might be neurally implemented with a map of units having reciprocal long-range inhibitory connections and short-range excitatory connections (Amari, 1977; Erlhagen and Schöner, 2002). Since the more biologically realistic mechanism did not show qualitatively better results we opted for using a simple max function since it is much more computationally efficient. Note, however, that the neural version of the saliency map might have various advantages over the simple max function in tasks more complex that those considered here, as, for example, it allows the system to select targets that do not lay on the vertexes of the grid of neurons that form it and it tends to select targets located close to the barycentre of objects.

After the winning unit is selected, the map's units a_{ij} are activated on the basis of a Gaussian function depending on their distance from winning unit itself. In particular, the units have a higher activation the closer they are to the winning unit:

$$a_{ij} = exp\left[-d_{ij}^2/\sigma^2\right] \tag{8.2}$$

where d_{ij} is the distance between the unit ij and the winning unit, and σ^2 is the standard deviation of the Gaussian. As we shall see in section 8.2.7, this Gaussian activation is used to train the top-down attention component of the system (actor-critic model).

8.2.3 Fovea

The input image was sub-sampled to produce three RGB maps of 2×2 pixels each representing the component of the system capable of distinguishing between different objects. For simplicity, in this work uniformly coloured squares of 20×20 pixels were used to identify different objects, but the architecture can be used with more sophisticated object-recognition methods (e.g. as those proposed by Rao and Ballard, 1995; Riesenhuber and Poggio, 1999). The pixels of these maps were activated with $\{0,1\}$ on the basis of whether the corresponding RGB pixels of the input image were on or off.

8.2.4 Periphery Map

The periphery map, which is incapable of discriminating between different objects, is in charge of guiding the eye to high-contrast regions of the work space. Given the goals of this work, the presence of a colour was sufficient as a contrast indicator, but more sophisticated bottom-up saliency maps may replace this mechanism to process more complex scenes (e.g., cf. Koch and Ullman, 1985; Itti et al., 1998; Itti and Koch, 2001a). Here, the input image is used to activate a 30×40 B/W low-resolution periphery map. For this purpose, the activation of each element of the map is obtained by averaging the RGB colour values in the range $[0,255]$ of a group of topologically corresponding 8×8 pixels of the input image so as to obtain a gray scale value in the range $[0, 1]$.

8.2.5 Inhibition-of-Return Map

The inhibition-of-return map works in a way analogous to what happens in real organisms (Tipper et al., 1991; Klein, 2000), and produces an efficient exploration of the different potential targets in a scene. This map is activated by the last visited locations (cf. Koch and Ullman, 1985; Klein, 2000) on the basis of the fovea position (the "eye's position in the orbit"). After each saccade, the previous position of the eye activates a cluster of units of the map on the basis of a Gaussian function. In particular, each map's unit c_{ij} is activated as follows:

$$c_{ij} = min[c_{ijt-1}(1 - \varepsilon) + exp[-d_{ij}^2/\sigma^2], \phi] \tag{8.3}$$

where ε is a decay coefficient, d_{ij} is the distance between the unit ij and the unit corresponding to the foveated point, σ^2 is the standard deviation of the Guassian, ϕ is the maximum activation of the map's units. Note that the bound imposed on the maximum activation of units avoids that excessive inhibition accumulates in correspondence to places that are foveated multiple times, as in the case of targets (see section 8.3.4).

During each saccade a hard-wired mechanism shifts the pattern of activation of the map in the direction opposite to the eye's motion in order to maintain it activation in an eye-centred reference frame.

Section 8.2.9 shows that the tasks used to test the models are organised in blocks each composed of eight presentations of the same image. When the image changes from one block to another, the activation of all units of the inhibition-of-return map is set to zero. This implies that the system empties the memory of the previously visited positions which so are no longer inhibited. This hardwired reset mechanism is used here to avoid interference between blocks and might have a correspondent in real brains where the inhibition-of-return process seems to be actually reset when the scenes abruptly change (Klein, 2000).

8.2.6 Potential Action Map

The potential action map ("PAM") implements the top-down attention processes of the system. At each time step, the PAM is updated so as to accumulate the evidence collected by the system while exploring different cues on the different potential positions of the target. At each step, the information on the possible localisation of the target rendered by the currently foveated object is expressed by the "vote map". This is the output layer of units of the actor component of the actor-critic model described in section 8.2.7. As we shall see, the units of this map learn to be more active for positions where the target might be with respect to the currently foveated object.

The PAM accumulates the step-by-step activation of the vote map and is subject to a decay. More precisely, the PAM is formed by leaky neurons p_{ijt} which receive a topological activation y_{ijt} from the vote map and have the following activation:

$$p_{ijt} = (1-\delta)p_{ijt-1} + y_{ijt} \tag{8.4}$$

where δ is a decay coefficient.

Note that in the tests reported in section 8.3 two versions of the model were tested, one with the PAM map storing a memory on the information returned by the previously explored objects, and one without such memory. The model with the memory was obtained by setting the parameter $\delta > 0$, whereas the memoryless version of the model was obtained by setting $\delta = 0$. For ease of reference the two models with and without memory will be henceforth called "BASE model" and "PAM model" respectively.

As in the case of the inhibition-of-return map, during each saccade a hard-wired mechanism shifts the activation of this map in the direction opposite to the eye's motion in order to maintain it in a eye-centred reference frame. Moreover, a hard-wired mechanism sets to zero the activation of the PAM's units each time the scanned image changes between different blocks of the tests.

8.2.7 Actor-Critic Model

The actor-critic model allows the system to store knowledge about the possible (deterministic or probabilistic) relative spatial relations existing between different foveated objects and the target. The actor-critic model consists of two main components, the actor and the critic (Sutton and Barto, 1998). The actor is a two-layer neural network which has the fovea units' activation x_{kl} as input, all-to-all connections w_{ijkl}, and a map of units as output, called the "vote map", whose activation y_{ij} is computed as follows:

$$y_{ij} = \sum_{kl} \left[w_{ijkl} \cdot x_{kl} \right] \tag{8.5}$$

The critic is mainly formed by a two-layer neural network, here called "evaluator", which is in charge of learning to assign an evaluation to the foveated objects. The evaluator has the fovea units x_{kl} as input, all-to-all connection weights w_{kl}, and a linear unit v as output:

$$v = \sum_{kl} \left[w_{kl} \cdot x_{kl} \right] \tag{8.6}$$

The evaluator's weights are updated on the basis of the TD learning rule (Sutton and Barto, 1998):

$$\Delta w_{kl} = \eta \cdot s_t \cdot x_{klt-1} \tag{8.7}$$

where η is a learning coefficient and s_t is the "surprise" computed as follows:

$$s_t = (r_t + v_t) - v_{t-1} \tag{8.8}$$

where r_t is the reward signal that the system receives when it foveates the target.

The actor's weights are updated on the basis of the surprise signal s_t and a modified Δ-rule (initially proposed in Ognibene et al., 2006) that takes into consideration the fact that the system represents actions with "population codes" (Pouget et al., 2000) or "neural fields" (Erlhagen and Schöner, 2002):

$$\Delta w_{ijkl} = \alpha \left(\left(y_{ijt-1} + a_{ijt-1} \cdot s_t \right) - y_{ijt-1} \right) \left(y_{ijt-1} \left(1 - y_{ijt-1} \right) \right) x_{klt-1} \tag{8.9}$$

where α is a learning rate, $\left(y_{ijt-1} + a_{ijt-1} \cdot s_t \right)$ plays the role of desired output and $\left(y_{ijt-1} \left(1 - y_{ijt-1} \right) \right)$ is the derivative of the sigmoid function. In this formula, the desired output is such that it tends to increase y_{ijt-1} when the surprise s_t is positive, and to decrease it when the surprise is negative (and to do so only for units with $a_{ijt-1} > 0$: these play the same role of the unit encoding the "winning action" in discrete-action reinforcement learning). Note that the formula can be easily rewritten to show that at its core there is a Hebb rule involving the units x_{klt-1} and a_{ijt-1} of the input and output maps:

$$\Delta w_{ijkl} = \alpha \cdot s_t \left(y_{ijt-1} \left(1 - y_{ijt-1} \right) \right) \left(a_{ijt-1} \cdot x_{klt-1} \right) \tag{8.10}$$

8.2.8 Parameter Settings

In the experiments reported in section 8.2 the parameters were set as follows: standard deviation of the Gaussian functions $\sigma^2 = 1.6$ (where 1 is the distance between two units); parameter of noise added to the saliency map for action selection: $\nu = 0.08$; connection weights of bottom-up attention map $\beta = 0.15$; connection weights of top-down attention map $\tau = 0.1$; connection weights of inhibition-of-return $\iota = 0.16$; decay coefficient of inhibition of return $\varepsilon = 0.5$; maximum activation of the units of the inhibition-of-return map $\phi = 0.5$; reinforcement-learning critic's discount factor $\gamma = 0.1$; critic's learning rate $\eta = 0.001$; actor's learning rate $\alpha = 0.001$; training phase: $T = 160,000$ images (equivalent to 20,000 "image blocks", see section 8.2.9).

8.2.9 The Tasks

The BASE and the PAM models were tested with two tasks, and some variants of them, having an increasing level of difficulty. As the solution of the tasks requires the same knowledge, the models were first trained on the basis of the simplest version of the first task and then tested in all other conditions (see section 8.3). The tasks are now explained in detail.

1-cue/x-dis task. Fig. 8.2 shows three example images used in this task. The images are randomly created by positioning a green cue (object) in a random vertex of a 5×5 grid, and the red target in a randomly selected vertex having either the same column or row of the cue. Note that this setting, which varied in different "blocks" (see below) of the tests, creates a *stochastic regularity* in the relative positioning of the cue and the target: the target had an equal chance of being on a vertex of the grid positioned on a "cross-shaped" area centered on the cue. Variants of the task were obtained by positioning a certain number of blue distractors on the remaining vertexes of the grid. In the tasks the number of distractors varied from one to ten: "x-dis" in the name of the task stands for the specific number of distractors used in the various tests (see section 8.3).

In all tasks each randomly image generated was presented for a sequence of eight simulation time steps: henceforth each sequence will be called a "block". The task with no distractors was used to test if the models were capable of integrating bottom-up and top-down attention. To do so the models had to learn, via the top down-attention components, in which area they could find the target with respect to the cue, and then they had to find the target within such area via bottom-up attention. The version of the task with distractors was used to test if the models were capable of keeping the memory of the target area suggested by the cue in the case one or more distractors in such area were foveated before the target was found (this was supposed to be a capacity possessed by the PAM model but not by the BASE model).

Fig. 8.2 Examples of images used in the "1-cue/x-dis" task (here x is equal to zero, three and five distractors, respectively in the three images). The black, dark grey and light grey squares in the images represent respectively the green cues, the red targets and the blue distractors.

2-cue/x-dis task. Fig. 8.3 shows an example of a sequence of four images used in this task. If the images are overlapped, they produce a whole image composed of: (1) two cues that are set on two vertexes of a 5×5 grid that are selected at random but have different columns and rows; (2) a target set on one of the two possible vertexes (selected at random) corresponding to the column of one cue and the row of the other cue; (3) a certain number of distractors set at random on the remaining vertexes (with the exception of the position, out of the two potential positions of the target, left empty). Note that here the images used in the task were divided into four images presented in sequence respectively one, one, one and five times during each "block" (Fig. 8.13). This was done to avoid a local minimum that prevented us to test the integration capabilities of models due to the fact that we had only three colours and this prevented us from having two different cues. In fact, we initially tried to directly use the whole images described above (i.e., the images including the target, two cues and the distractors) but once the models foveated the first cue they immediately tried to search the target on the row and column of the first cue instead of searching the second cue (this strategy was actually more efficient as in the setting described above the second cue was set outside the area indicated by the first cue: it was not possible to put the second cue inside the area indicated by the first cue – as suggested by the example of the car in the street suggested below – since, as the cues had to be the same, it was not possible to create a hierarchy between them). In future work, the use of a more complex object recognition component will allow us to remove this simplification (see section 8.4), introduced here to allow running simple clear tests of the of the basic functionalities implemented by the model. This task required not only to integrate bottom-up and top-down attention and to retain information in memory, as in the previous task, but also to integrate information coming from two cues, for example by searching the target within an area corresponding to the intersection of the two areas suggested by the two cues. This function can be implemented only by the PAM model as it requires the memory of the action suggested by the first one of the two foveated cues. To give an idea of the use of this capability, think about an eye that has to foveate a person in a city. A strategy to solve this task might be finding a street, scanning its surface in search of a car, and looking the through the car's windows to see if there are people inside. In this example, a first cue (the car) is used to search a second cue (the car) and then the target by integrating the information given by the two cues.

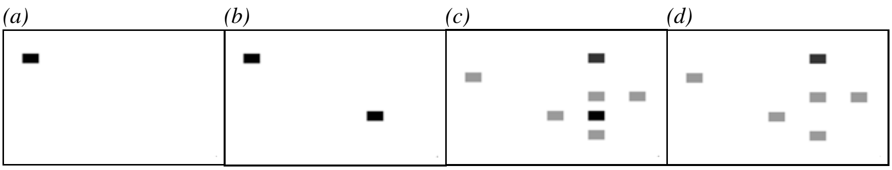


Fig. 8.3 Example of images used in a block of 2-cue/5-dis task. (a) The first image of each block contains only the first cue, used as the first input to the models. (b) The second image reports both cues to allow the test program to evaluate if the models select the first cue, and to allow the models to see the second cue. (c) The third image reports the second cue, to allow the test program to evaluate if the models select the second cue, and the target plus the distractors. (d) The fourth image, repeated five times in each block, contains only the cue and the distractors.

In all tasks the cue(s) have maximum luminosity (i.e., their RGB colour values are set to 255) whereas the target and distractors have a lower luminosity (colour values set to 230). This simple technique was used to bias the system to first foveate cues and then other objects so as to ease the statistical analysis of the models (see section8.3). Both training and tests consisted in the presentation of a certain number of blocks of images to the systems plus a reward signal of one each time the models foveated the target.

The performance was always computed as the percent of times in which the systems' eye was on the target. These settings implied that the maximum theoretical performance in the tasks, without considering the negative effects of distractors, was as follows:

- *1-cue/x-dis task*, 7/8=.8750: the optimal model would first foveate the cue, then the target and then stay on it.
- *2-cue/x-dis task*, 6/8=.7500: the optimal model would first foveate the first cue, then the second cue, then the target and then stay on it.

Note that the tests use simplified images (e.g., objects identified by colours, and positions of objects on a grid) to ease the analysis of the models reported below. In fact they allow computing optimal performance, analyse the sequences of behaviour, explain the functioning of the systems' components, etc. However, as mentioned in section 8.4, this is not a limit of the architecture as it might be endowed with more sophisticated components, such as a more sophisticated object-recognition component, in order to let it tackle more complex tasks involving real-world images.

8.3 Results

This section illustrates the performance and functioning of the BASE and the PAM models when tested with the tasks illustrated in section 8.2.9. In particular, the models are first trained on the basis of the simplest version of the first task as the solution of the two tasks requires the same knowledge (section 8.3.1). Then the models are tested with various versions of the tasks to analyse the functioning of their various components: the bottom-up attention map and the inhibition-of-return map (section 8.3.2), the vote map (section 8.3.3), and the potential action map (sections 8.3.4, 8.3.5, 8.3.6).

8.3.1 Learning and Performance of the Models

Fig. 8.4 illustrates the learning curves of the BASE and the PAM models trained with the first task for 20,000 image blocks each having a random number of distractors (from one to five). The figure also reports two simulations where the parameter ν of the exploration noise was set to zero in the two models. These experiments were carried out with two goals in mind: (a) testing if the bottom-up mechanisms driving the systems is capable of generating the necessary explorations needed by the functioning of reinforcement learning algorithms without the addition of noise: this is an interesting function that this mechanisms might play if reinforcement learning algorithms are used to learn eye motion; (b) evaluating the capacity of the models to learn to stay on targets once found. The results show that the PAM model learns fast both with and without noise and reaches a steady state value. On the contrary, the BASE model with noise learns more slowly and achieves lower performance than the PAM model: we shall see in section 8.3.5 that this lower performance is caused by the fact that, since the BASE model does not have the memory of the potential action map, once it foveates a distractor it looses the information given by the cue. Moreover, the BASE model with no noise achieves a very low performance: as we shall see in section 8.3.4 this is due to the fact that the model is not capable of producing the experience necessary to learn to stay on the target.

After training, the four trained models were systematically tested on 50,000 blocks of various versions of the two tasks: 1-cue/0-dis, 1-cue/5-dis, 1-cue/10-dis, 2-cue/5-dis, 2-cue/10-dis. The results are reported in Table 8.1 which shows various interesting facts. First, the test run with the 1-cue/0-dis task confirm that the the BASE no-noise model does not fully learn the task. Second, they indicate that in the 1-cue/x-dis tasks with five or ten distractors the PAM model outperforms the BASE model. As we shall see in section 8.3.5, this higher performance is due to the PAM model's capacity to retain in memory the information provided by the cues when

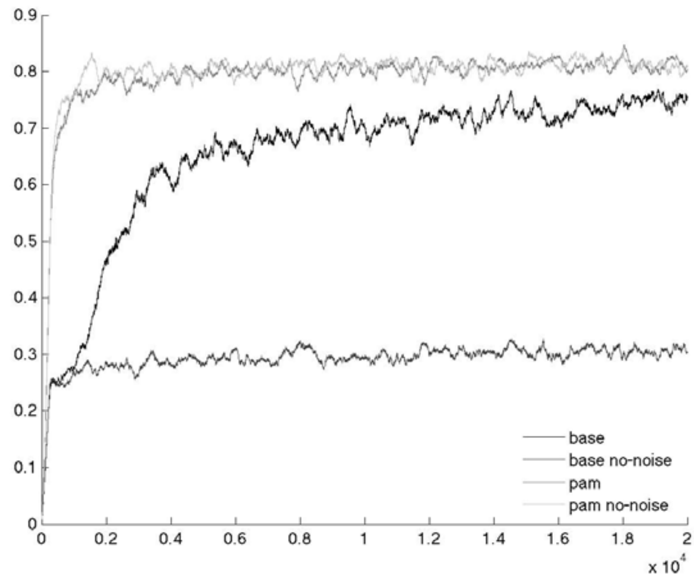


Fig. 8.4 Learning curves of the BASE and the PAM model trained with the presentation of 20,000 blocks of the 1cue/0dis task. The x-axis reports the blocks of eight images presented in sequence to the models whereas the y-axis reports the models' performance computed as the percent of times in which the systems' "eye" was on the target (the values reported by the curves correspond to this performance filtered with a moving average having a window size of 256 blocks).

the system foveates distractors (note that the difference in performance with the BASE model increases with a higher number of distractors, e.g. passing from five to ten). Third, in 2-cue/x-dis tasks with five or ten distractors the PAM model's performance is much higher than the BASE model's one, and this difference increases with a higher number of distractors, for example with five and ten distractors, the PAM model's performance is respectively 139% and 211% (i.e. more than double) of the BASE model's performance. As we shall see in section 8.3.6, this is due to the PAM model's capacity of integrating information given by the two cues so as to be capable of precisely locating the target notwithstanding the presence of distractors in areas suggested by the cues taken alone.

8.3.2 Bottom-Up Attention: Periphery Map and Inhibition-of-Return Map

This section analyses the functioning of the bottom-up components of the models, namely the periphery map and the inhibition-of-return map. For this purpose, Fig. 8.5 shows an image used in a block of the 1-cue/10-dis task and the activation of the saliency and inhibition-of-return maps of a MAP model after five exploration steps of the image. The data have been collected with a system that has not yet been

Table 8.1 Performance of the BASE and PAM models (with and without noise) in 50,000 blocks of five variants of the two tasks. The last two rows of the table report respectively the relative performance of the PAM and BASE models and the theoretical performance (not considering the distractors). Each cell of the first four rows reports the performance of the models both in terms of the fraction of steps in which the models got rewarded and as the fraction of such steps with respect to fraction of steps of the theoretical performance. Note that a performance higher than the theoretical one in the 1-cue/0-dis task is due to minor implementation biases with respect to the ideal test (e.g., the models found themselves already on the target in the first image of some blocks).

Architecture	1-cue/0-dis	1-cue/5-dis	1-cue/10-dis	2-cue/5-dis	2-cue/10-dis
BASE	0.8892 - 1.02	0.6792 - 0.78	0.4537 - 0.52	0.4980 - 0.57	0.3094 - 0.35
BASE no-noise	0.4999 - 0.57	0.1732 - 0.20	0.0945 - 0.10	0.1340 - 0.15	0.0748 - 0.08
PAM	0.8913 - 1.02	0.7445 - 0.85	0.5503 - 0.63	0.6909 - 0.79	0.6525 - 0.75
PAM no-noise	0.8906 - 1.02	0.7177 - 0.82	0.4979 - 0.57	0.6989 - 0.80	0.6487 - 0.74
PAM/BASE	1.00	1.09	1.21	1.39	2.11
Theor. perf.	0.8750	0.8750	0.8750	0.7500	0.7500

trained so that there are no top-down influences on the saliency map. This implies that the saliency map's activation reflects only the input from the periphery map and the inhibition-of-return map. This also implies that a BASE model would have had a similar behaviour as the one described in the following. Note that in this sections and the following ones, the noise of the saliency map was set to zero to have clearer analyses of the models and to show their intrinsic exploratory properties.

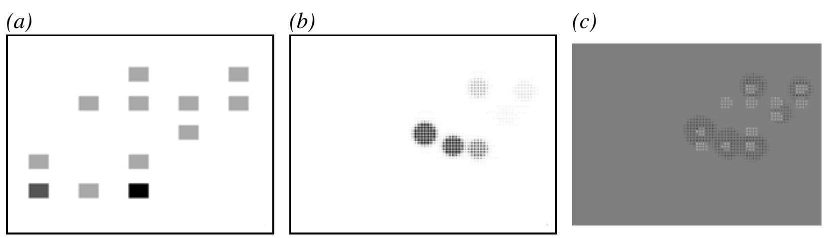


Fig. 8.5 (a) an image used in one block of the 1-cue/10-dis task; (b) activation of the inhibition-of-return map: black dots are proportional to the activation of the corresponding neurons in the map; (c) activation of the saliency map: white dots indicate activations above 0.5 (this are caused by the periphery map) whereas black spots indicate activations below such value (this are caused by the inhibition-of-return map). The sizes of the dots are proportional to the activation of the neurons.

The figure shows that while the eye explores the image (Fig. 8.5a) it generate clusters of inhibited neurons with an inhibition that decreases with elapsing of time (Fig. 8.5b). The figure also shows that the model explores only regions of the image where there are objects thanks to the bottom-up effects of the periphery map (note how clusters of inhibited activity, caused by eye's visits, fall only on spots where there is a bottom-up excitation from objects, see Fig. 8.5c). The interplay between the bottom-up saliency of the image's elements and the inhibition of return generates a rather efficient exploratory behaviour.

8.3.3 Analysis of the Vote Maps

Fig. 8.6 shows the activation of the vote map of the BASE and PAM models when the systems foveate either a target, a cue or a distractor. The activations are similar for the two models, so they can be explained together. The graphs show that when the systems foveate a target, the vote map has these activations: (a) a cluster of neurons activated above 0.5 in correspondence to the centre: this lead the eye to stay on the target; (b) clusters of neurons activated below 0.5 in the remaining places: this bias the eye to avoid moving to any other place. When the systems foveate a cue, the vote map has these activations: (a) a cluster of neurons activated below 0.5 in correspondence to the centre: this bias the eye to move away from the cue (this strengthens the effect of inhibition of return); (b) clusters of neurons activated above 0.5 in correspondence to the row and column of the cue: this bias the eye to move on objects on such column and row and captures the regularity related to the probabilistic spatial relations existing between the cue and the target; (c) cluster of neurons activated below 0.5 in the remaining places: this bias the eye to ignore distractors located in such positions. When the systems foveate a distractor, the vote map has these activations: (a) a cluster of neurons activated below 0.5 in correspondence to the centre: this bias the eye to move away from the distractor (this strengthen the effect of inhibition of return); (b) scattered clusters of neurons with activation above or below 0.5: the function of these activations is at the moment unclear and is still under examination.

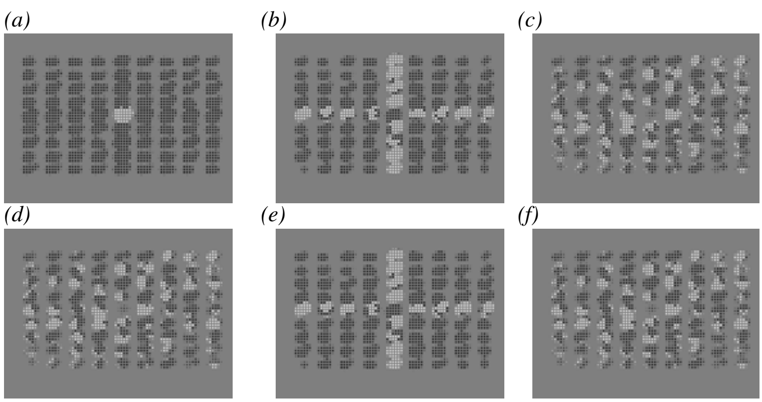


Fig. 8.6 Activation of the vote map of the trained BASE model (a, b, c) and of the trained PAM model (d, e, f) when the systems foveate a target (a, d), a cue (b, e) or a distractor (c, f). White dots indicate activations above 0.5 of the corresponding neurons, whereas black spots indicate activations below such value. The size of the dots is proportional to the activation of the neurons.

8.3.4 Capability of Learning to Stay, and of Staying, on the Target

We are now in the position to explain in detail why the BASE no-noise model fails to learn the 1-task/x-dis task (with one to five distractors), and as a consequence has a low performance in other (variants of the) tasks, whereas the PAM model (both with and without noise) quickly learns to stay on the target once found and hence has a high performance in all variants of the tasks. Direct observation of the behaviour of the BASE no-noise model during learning shows that when it finds the target it always moves away from it in the following step. On the contrary, the PAM model goes from the cue to the target and then stays on it. Fig. 8.7, which shows the average rewards that the two models get in the eight steps of the blocks, presents a quantitative account of these behaviours. The figure shows that, contrary to the PAM model, the BASE no-noise model tends to visit the cue in the first step of each block, then finds the target, but then goes back to the cue and repeats this behavioural pattern till the end of the block.

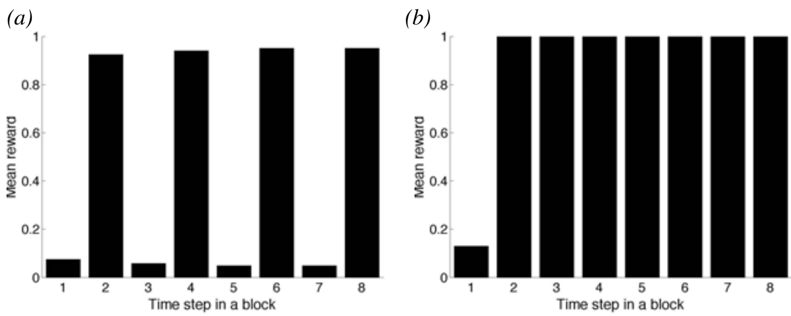


Fig. 8.7 1-cue/0-dis task: mean reward (y-axis; averaged over 50,000 blocks) that the BASE no-noise model (a) and the PAM model (b) receives in correspondence to the eight time steps of blocks (x-axis).

An analysis of the vote maps of the two models and of the related saliency maps, reported in Fig. 8.8, explains the reason of this different behaviours. Inhibition of return tends to have a negative effect when the eye is on the target, as it is highest for the currently foveated location, and so bias the eye to other locations. The BASE no-noise model is not capable of compensating this bias and so moves away from the target (Fig. 8.8d). The point is that, not being capable of remaining on the target, the system is not capable of producing the necessary experience needed to learn to stay on it. Fig. 8.8a shows that, as a consequence, the vote map learns a sub-optimal strategy: since the model is not capable of staying on the target, it votes to go back to the cue as this is followed by another rewarded target. On the contrary, the PAM model is capable of compensating the negative effect of inhibition of return thanks to the positive bias generated by the cue in the previous time step and still memorised in the Potential Action Map (this bias can be seen in Fig. 8.8e in terms of the clusters of units with an activation above 0.5 spatially arranged as a cross-shape). Fig. 8.8b

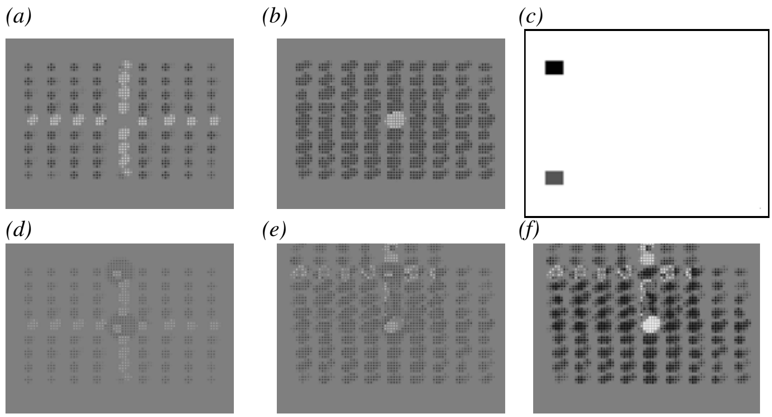


Fig. 8.8 Top graphs: activation of the vote maps of the BASE no-noise model (a) and the PAM model (b; same as Fig. 8.6a) when looking at the target of an image (c) of the 1-cue/0-dis task. Lower graphs: activation of the saliency map of the BASE no-noise model (d) and of the PAM model (e) as soon as they foveate the target, and activation of the saliency map of the PAM model after two steps (f).

shows that, as a consequence, the vote map learns the optimal strategy of staying on the target once found: this allows the model to remain on the target even when the cue's bias fades away from the potential action map memory and the inhibition of return on the cue decays to zero.

We have seen (Table 8.1) that the BASE model (i.e. the BASE model *with* noise) learns successfully to solve the 1-cue/x-dis task. The reason for this is that noise allows the system to occasionally overcome the negative effect of inhibition of return when the system is on the target. This allows the system to develop a bias to stay on the target similar to the one developed by the PAM model. However, it should be noticed that this solution based on noise is less powerful than the solution of the potential action map's cue-bias exploited by the PAM model that works each time the target is found and not only sporadically as the solution based on noise. This is demonstrated by the fact that the PAM model learns much faster than the BASE model (see Fig. 8.4).

8.3.5 Potential Action Map: An Action-Oriented Memory of Cue Information

This section explains in detail one of the main functions played by the potential action map, that is the capacity of memorising the information on the target rendered by a given cue in a format readily usable for guiding action. To show this function, the BASE and PAM models were tested with 50,000 blocks of the 1-cue/10-dis task. The results of the test reported in Fig. 8.9 give broad indications of the behaviour of the two models during the test. The graphs in the figure report the probability that the models' eye is on the target in the eight steps of the blocks. A comparison of the two models with respect to this aspect indicates that they do not differ with respect to the first step, usually involving the foveation of the cue, and the second step, usually involving the foveation of either the target or a distractor with a probability of respectively about 30% and 70%. Note that in the second step models can use both the bias from the cue to search the target on the“cross area” centred on the cue and the bottom-up information from the target or the distractors. Without the bottom-up information, they would have a chance of about $9\% = (1/11) \times 100$ of finding the target and $91\% = (10/11) \times 100$ of finding a distractor. The most relevant difference between the models happens on the third step. Here the probability that the PAM model moves to/is on the target is about 50% whereas the BASE model's one is 40% (this advantage is then maintained in the succeeding steps). The reason of this is that in the case the BASE model foveates a distractor after the foveation of the cue (second step), it completely looses the information on the target given by the cue. On the contrary, the PAM model is capable of continuing to search in the “cross-shaped” area indicated by the cue by exploring in sequence all the spots within such area marked by the bottom-up salience.

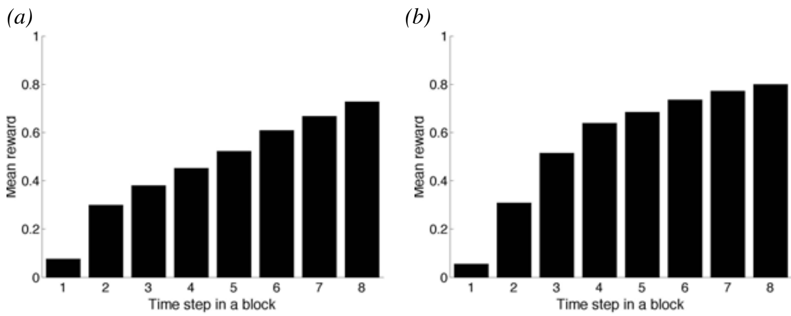


Fig. 8.9 1-cue/10-dis task: mean reward (y-axis; averaged over 50,000 blocks) that the BASE model (a) and the PAM model (b) receive in correspondence to the eight time steps of blocks (x-axis).

Fig. 8.10 supports this interpretation by furnishing a further analysis of the behaviour of the two models. The figure shows the ten most frequent sequences of objects foveated by respectively the BASE model and the PAM model during the 50,000 blocks of the test. As it can be seen, the sequence where the models first foveate the cue and then the target has the same frequency for both models (see the first most frequent sequence). However, the PAM model's second and third most frequent sequences are those where the system foveates one or two distractors after the cue and before the target. These are the best moves the model can perform when it fails to find the target in the first step after the cue. On the contrary, in the case of the BASE model, these two sequences have only the third and fifth rank in frequency and, what is more important, have a much smaller absolute frequency with respect to the PAM model (in particular, a frequency of about 0.05 and 0.04 respectively in the case of the BASE model versus 0.19 and 0.10 in the case of the PAM model).

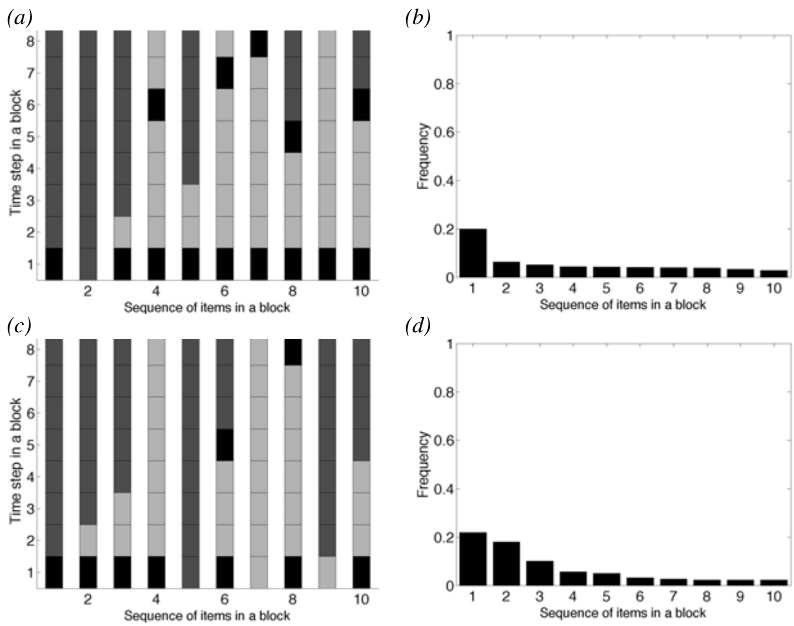


Fig. 8.10 1-cue/10-dis task: ten most frequent sequences of objects foveated by the BASE model (a) and histogram of related frequencies (b), and analogous data related to the PAM model (c and d respectively). In graphs (a, c) the vertical axis reports the sequences of object foveated by the models during the eight time steps forming blocks (black: cue; dark grey: target; light grey: distractors), whereas the horizontal axis reports the different sequences. In graphs (a, c) the vertical axis reports the fraction of blocks in which the sequences are used during the whole experiment (50,000 blocks).

Fig. 8.11 shows why the two models exhibit such behaviour after they encounter a distractor after the cue, in particular why the PAM model can still have a good performance after such "mistake" occurs. The figure shows the activation of the saliency map of the two models when they are on the cue, and the activation of the same map when the models visit a distractor after the cue. The activation of the map of the two models is similar when the cue is foveated, but differs when the models foveate a distractor in the following time step. In particular, contrary to the BASE model, when the PAM model foveates the distractor it maintains an activation corresponding to the potential positions where the target might be as suggested by the cue (see the clusters of highly activated units spatially organised as a cross-shape in Fig. 8.11d).

8.3.6 Potential Action Map: Capacity to Integrate Multiple Sources of Information

This section explains another important function played by the potential action map, that is the capacity of integrating the information on the target rendered by more

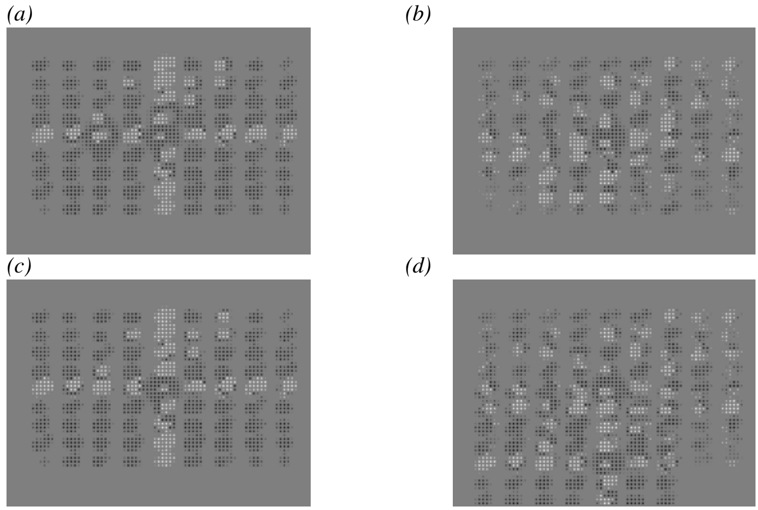


Fig. 8.11 Activation of the saliency map of the BASE model when it foveates a cue (a) or a distractor after the cue (b). Graph (c) and (d) show analogous data relative to the PAM model. Note in (d) the cluster of white dots (units activated positively) spatially organised to form a cross-shape centred on the position of the visited cue: the activation of these units is caused by the potential action map.

than one cue. To illustrate this function, the BASE and PAM models were tested with 50,000 blocks of the 2-cue/10-dis task. Fig. 8.12 reports some results of this experiment that, together with a direct observation of behaviour, furnish a general idea of the strategies used by the two models to tackle the task. In the first and second step, the two models get zero reward as only the cues are visible. In the third and fourth steps the models's performance diverges dramatically: the PAM model is on the target about 70% and 90% of the times respectively, whereas the BASE model only 25% and 30% of the times respectively.

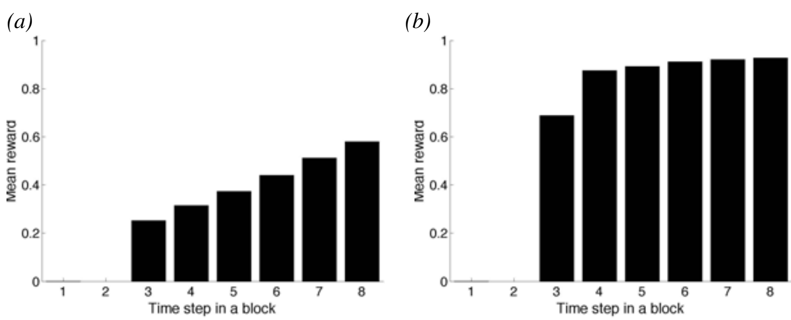


Fig. 8.12 2-cue/10-dis task: mean reward (y-axis; averaged over 50,000 blocks) that the BASE no-noise model (a) and the PAM model (b) receives in correspondence to the eight time steps of blocks (x-axis).

The reasons of this higher performance can be understood considering the most frequent sequences of moves performed by the two models, presented in Fig. 8.13. In this regards, the figure indicates that the PAM model's first choice is the optimal sequence formed by the two cues followed by six targets (this sequence is selected about 70% of the times), whereas this sequence is only the second choice of the BASE model (selected only about 25% of the times). The reason of this is that the PAM model can find the most likely position of the target by integrating the information from the two cues (Fig. 8.13c-d), and when it makes a mistakes (i.e. foveates a distractor after the cue) can often recover in the successive moves (see sequences 2, 4, 5, 6 in Fig. 8.13c, which include an increasing number of distractors before the target and have a decreasing frequency). On the contrary, the BASE model cannot exploit the information returned by more than one cue, so it has lower chances of finding the target after the two cues. Moreover, if it foveates a distractor instead of the target after the two cues it "gets lost" as it has not retained in memory the information from the cues, and so starts to search the target with a random search (see sequences 3-7 in Fig. 8.13a, which have a similar frequency, Fig. 8.13b)

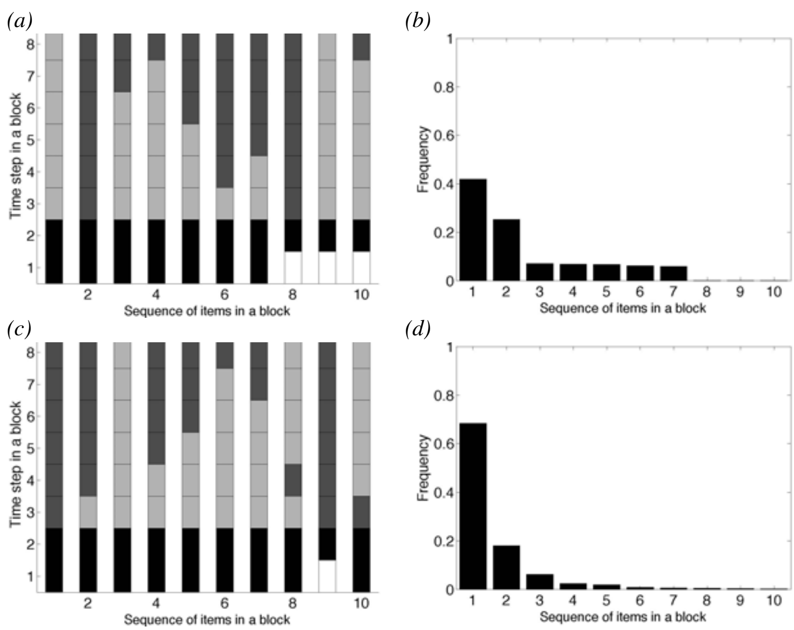


Fig. 8.13 2-cue/10-dis task: ten most frequent sequences of objects foveated by the BASE model (a) and histogram of related frequencies (b), and analogous data related to the PAM model (c and d respectively). In graphs (a, c) the vertical axis reports the sequences of object foveated by the model during the eight time steps forming blocks (black: cue; dark grey: target; light grey: distractors; white: saccade out of any object), whereas the horizontal axis reports the different sequences. In graphs (b, d) the vertical axis reports the fraction of blocks in which the sequences are used during the whole experiment (50,000 blocks).

This interpretation is corroborated by the data presented in Fig. 8.14, which show the activation of the saliency map of the BASE model and PAM model when they foveate the first cue and then the second cue. The graphs clearly show that the PAM model records information given by the two cues and so has a high probability of searching the target on the two intersections of the two cross-shaped areas suggested by the cues, whereas the BASE model looses the information about the first cue when it foveates the second cue.

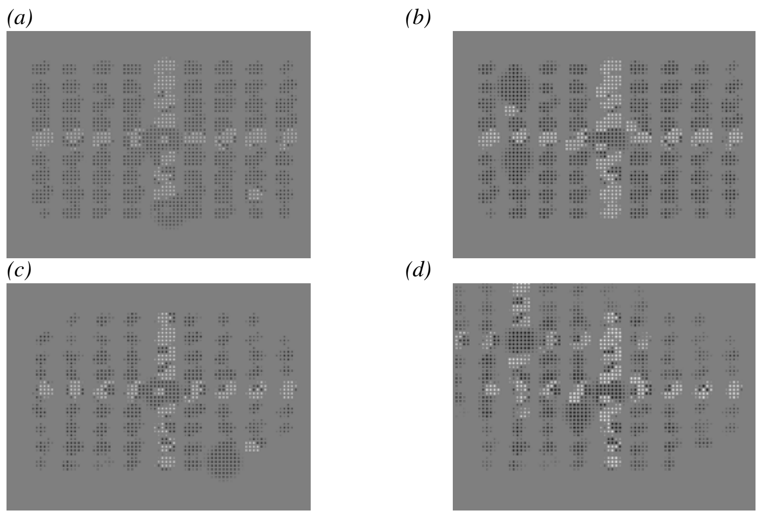


Fig. 8.14 Activation of the saliency map of the BASE model when it foveates the first cue (a) and then the second cue (b). Graph (c) and (d) show analogous data relative to the PAM model. Note in (d) the cluster of white dots (units activated positively) spatially organised to form a cross-shape centred on the position of the two visited cues: the activation of these units is caused by the potential action map.

8.4 Conclusions

This paper presented an architecture that combines a basic bottom-up attention system, analogous to systems proposed within the literature on visual attention, and a novel top-down component, the *Potential Action Map* (PAM), which uses reinforcement learning to learn to attend to rewarded stimuli. This map functions as a memory that accumulates evidence in favour of the locations where rewarding targets might potentially be located with respect to foveated cues.

The architecture, and in particular the potential action map component, have a number of appealing features. Some of these were investigated within the work reported here whereas others will be tested in the future.

The first strength of the architecture is its capacity to integrate bottom-up and top-down processes. The architecture shares this feature with other models, for example the one proposed by Balkenius and Johansson (2007). However, it should be noticed that that architecture does not integrate bottom-up attention, inhibition of return and top-down attention in a flexible way. In particular, it integrates them with a simple summation which makes the contributions from the various components rather rigid. A possible solution to this problem would be to let the system learn the relative contribution that the various components should give to the final action decision taking place at the level of the saliency map (see for example Balkenius et al., 2004). Indeed, in the current architecture's implementation it was not straightforward to tune the contributions of the various maps to the saliency map by hand so as to enhance learning. For example, the balance between the inhibition-of-return map, that tends to drive the eye away from foveated objects, and the top-down attention map, which drives the eye to remain on targets once found had to be carefully adjusted.

Another strength of the architecture is that it fully works in eye-centred coordinates. Eye-centred representations not only reflect empirical findings on neural representations used in real brains (Dominey and Arbib, 1992; Shadmehr and Wise, 2005), but they also have various computational advantages. A first advantage, mentioned in the introduction, is that complex visuomotor transformations can be implemented in later phases of the sensorimotor flow where information is usually encoded in more abstract forms. For example, here all visual processes (inhibition of return and bottom-up and top-down attention processes) were implemented on the basis of eye-centred reference frames. The transformation of information to a body/environment-reference frame, needed to issue the desired gaze command to the motor eye system, took place at a later stage in the form of a summation of the current eye posture and the desired eye displacement.

An alternative strategy would have been to use the retinal images to build a body/environment-centred representation of the environment, to apply the visual processes to it, and then to activate the saliency map according to these. However, this would have required the system to apply computationally shifts to the retinal images, which would be much heavier than the shifts applied here to the action-oriented memories implemented by the inhibition of return and potential action maps.

Another advantage is that eye-centred representations are "deictic" in the sense that they encode information "by pointing" with respect to the context in which they are used (e.g.: "move to direction x with respect to currently foveated point"). Deictic representations simplify computations and enhance the generalisation capabilities of systems (Ballard, 1991). For example, here the actor could learn the *relative* spatial relations existing between the cue and the target by representing these relations with respect to the currently foveated cue. This allowed the actor to have a simple structure, to learn fast, and to automatically generalise its knowledge with respect to the absolute position of the cue/target couples.

Another advantage of the potential action map becomes apparent when it is used in partially observable environments (as the one consider here; (Whitehead and Lin, 1995)) or stochastic environments. In the case of partial observability of environments the potential action map makes the model capable of testing various available options at a given uncertain state without the need of explicitly encoding such state. For example, assume that when the system is in state A or state B it has the same perception S_{AB} to which it associates two different areas in which to search the target (one learned when the system visited A in the past and the other when it visited B). In this case, the system might decide to first greedily visit the one of the two areas which is most promising in terms of reward. The point is that in case of failure of this greedy search, the system might still visit the other area by directly moving from the new state to such area on the basis of the information collected at S_{AB}, that is without the need of going back to it. Notice that the system can use a similar solution also in the case when the world is stochastic and the system selects an action that leads to an undesired state: it can still select an action that attempts to have the same effects of the previous one, that is, it can still use the information given by the previous state (eventually integrated with that of the new state).

In the future, it might be interesting to evaluate if the idea of the potential action map, and the mentioned advantages, might be extended to other control domains. For example, in the control domain of robotic arms engaged in reaching tasks the potential action map might represent the arm's potential actions within a neural map encoding the arm's "equilibrium points" (that is, the desired postures: see Ognibene et al., 2006; Herbort et al., 2007).

Notwithstanding the aforementioned strengths, the current implementation of the model has various limitations that might constitute the starting point for future work. A first limitation is the simplified bottom-up attention component. However, as mentioned in section 8.2.4, this is not a general drawback of the architecture as this component might be easily substituted with a more sophisticated component capable for example of performing detection of edges, colours, motion, etc. (Itti et al., 1998; Itti and Koch, 2001a; Balkenius et al., 2004).

A second limitation is the simplified object-recognition component, currently based on a simple colour-detection device. Again, as mentioned in section 8.2.3, this is not a general drawback of the architecture as this component might be substituted with mechanisms capable of implementing more sophisticated feature-extraction processes (see for example Rao and Ballard, 1995; Riesenhuber and Poggio, 1999).

A last limitation is that some important mechanisms used by the architecture are currently hardwired. These mechanisms are the reset of the memories (inhibition of return and potential action maps) when the scene changes, and the shift of their activation when the eye moves. The first mechanism, implementing the reset of memories, might be substituted with neural mechanisms that reset memories only locally on the basis of abrupt changes of the activation of the neurons of the bottom-up attention maps. A similar mechanism seems to operate in real brains for inhibition of return (Klein, 2000) and was used by Balkenius (2003) to reset a memory for visual context. The second mechanism, the shift of memories, seems to play an important role in visuomotor transformations implemented in real brains (Gnadt and Andersen, 1988; Shadmehr and Wise, 2005). Various neural models of this mechanism have been proposed in the literature, for example based on dynamic networks (Dominey and Arbib, 1992; Zhang, 1996) or "gain fields" (Casarotti et al., 2003; Shadmehr and Wise, 2005 for a review). Some of the algorithms used in these models might be suitably used to substitute the currently hardwired shift mechanism in future implementations of the architecture.

Chapter 9
Anticipation by Analogy

Boicho Kokinov, Maurice Grinberg, Georgi Petkov, and Kiril Kiryazov

9.1 Introduction

Why do we expect to find intelligent creatures on other planets in the Universe? Is there a general low stating it? Is there a theory that predicts it? Do we have many examples in order to generalize from them? No. Our anticipation to find intelligent beings is based on analogy with the only example we know - the planet Earth. Moreover, if we analyze the description of these potential creatures in the science fiction literature and movies, we will find out that all out imagination about these creatures is based on analogies with the human race or other animals on the Earth. Similarly, when the Europeans arrived for the first time at other continents, they expected the population there to have a culture analogous to the European and when this turned out not to be true, they announced ”the other” to be ”less developed”, exactly as the Romans declared all non Romans to be Barbarian. The same happens to each of us when traveling to a new country - we always expect to find patterns of behavior that are analogous to the patterns we are familiar with in our own culture and are highly surprised when these anticipations turn out to be wrong. Young children expect an object to breath or feel pain to the extent to which they consider it analogous to the human being ((Inagaki and Hatano, 1987). All these examples demonstrate that when human cognition faces a new situation it usually uses analogy with a previously experienced or familiar situation to make predictions of what to expect. These analogies do not necessarily lead to correct predictions, but this is often the best the cognitive system can do under the given circumstances (especially in a new domain, where little knowledge is present).

From computational perspective analogy-making is a good heuristics that makes effective short-cuts in the exhaustive search of the state space. Suppose you have to search for your keys. Theoretically they can be everywhere at your place or even outside it. So, if you want to be sure you will find them you should search the whole space and look under, inside, and behind every single object around you. This will, however, keep you busy forever. Analogy-making will provide you with a heuristics where to look first (rather than searching for the key randomly) - it might be worth

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 185–213, 2008.
© Springer-Verlag Berlin Heidelberg 2008

exploring the place where last time you found your keys - it is very probable that you put them there again. There is an assumption here that there is some regularity in your actions, that you do not behave randomly, which assumption makes sense.

Finally, from a cognitive perspective analogy-making is a very central and basic mechanism that is present from very early age if not from birth (Holyoak et al., 2001), (Hofstadter, 2001), (Goswami, 2001), so it is very likely that it is used for such an important function like anticipation. In addition, in this chapter we will explore the mechanisms of analogy-making and how they emerge from simpler mechanisms like spreading activation, marker passing, and elementary transfer and anticipation in the AMBR model. Thus the relation between analogy-making and anticipation is two-fold in this model: on one hand, analogy-making is used for high-level anticipation, on the other hand, low-level anticipations are used in the process of analogy-making, and more specifically in the process of representation-building (or perception).

9.2 The Anticipation by Analogy Scenario

Let us consider the simplest scenario in which a dog-robot (AIBO) is searching for a bone hidden somewhere behind some object in one of the rooms of a house (Figure 9.2).

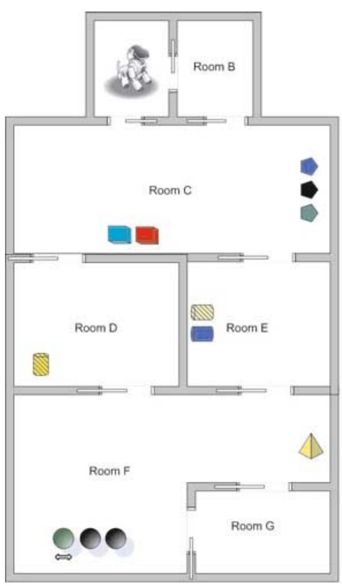


Fig. 9.1 The search scenarios: The AIBO robot is looking for the bone hidden behind an object in one of the rooms of a house. This is analogous to a human being searching for the keys somewhere in the house

The search scenarios: The AIBO robot is looking for the bone hidden behind an object in one of the rooms of a house. This is analogous to a human being searching for the keys somewhere in the house.

Let us consider an even simpler scenario: the dog-robot is searching for the bone hidden behind an object in the room (Figure 9.2). The first step would be to perceive the environment and build a representation of the room and the objects in it (Figure 9.3). As we will see later on in the chapter even here an analogy with a similar episode in the past might be helpful, i.e. old episodes assist us in representation-building by suggesting (anticipating) some attributes and relations in the environment and thus focusing our attention towards them. The next step would be retrieving an episode from long-term memory (LTM) that could potentially serve as a base for analogy (Figure 9.4). Then a mapping between the old episode and the current episode is being established and by transferring the knowledge where the bone was in this old situation, an anticipation is formed where the bone might be now (Figure 9.5). Finally, the robot moves towards the objects and searches for the bone in the anticipated place.

Fig. 9.2 The AIBO robot is in a room with several objects and the bone behind one of them.

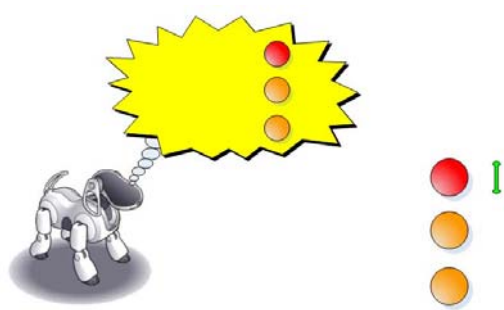


Fig. 9.3 The AIBO robot perceives the scene and build a mental representation of it.

Unlike the linear description above, the actual AMBR model works in parallel, i.e. all the above processes run in parallel and interact with each other. Thus even though the perceptual process must somehow start by a bottom up process very early on, it is later on guided by the bases already retrieved from memory. The mapping process is also further influencing the retrieval of episodes, etc.

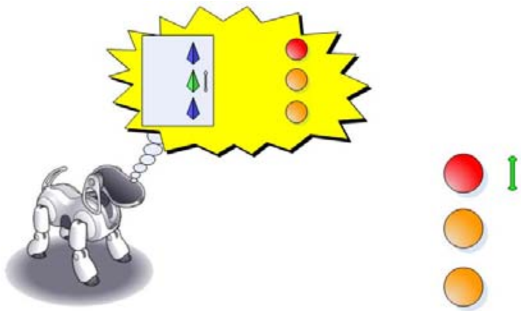


Fig. 9.4 The AIBO robot retrieves from long-term memory an old episode that seems analogous to the current one.

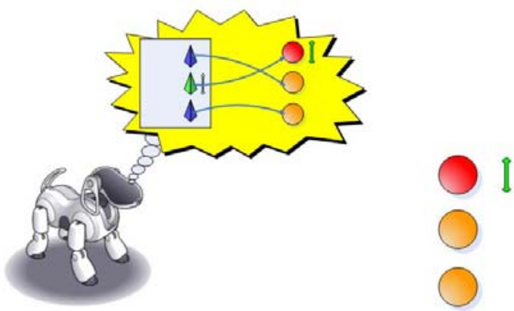


Fig. 9.5 The AIBO robot maps the two episodes and transfers the knowledge about the bone in the old episode to the new one and thus forms a specific anticipation. Then it moves to the anticipated place and searches for the bone.

9.3 Models of Analogy-Making

In an attempt to implement such a scenario a search for appropriate models of analogy-making was done. In this section we will present a brief review of a number of models.

The classics in the field is the Structure-Mapping Engine (SME) (Falkenhainer and Forbus, 1989) - a model fo analogical mapping based on the Structure Mapping Theory (SMT) by Dedre Gentner (Gentner, 1983). This model assumes that the mapping process is completely based on structural grounds and no semantic or pragmatic considerations play any role during this mapping phase. It postulates that only the relations are taken into account, i.e. the attributes are completely ignored. Moreover, only identical relations are mapped to each other. A very important constraint is the one-to-one mapping between the elements of the base domain and the target domain, which ensures parallel connectivity which means that when two relations are put into correspondence their arguments should also be put into correspondence. Another important constraint comes from the systematicity principle: when there is a choice between two alternative mappings, the more systematic one is preferred. This is in synchrony with the idea that analogy is mapping of systems of relations, not of isolated relations. All these main postulates are more or less shared by all models of analogy-making proposed afterwards. There are, of course, many disagreements in the field. For example, some researchers would not agree that only the relations play a role and that attributes are ignored. For example, in the scenario from the previous section the colors of the balls might be crucial for the analogy as we shall see in the later sections. Thus this model cannot be directly applied for this scenario. Other points of disagreements would be the need to map only identical relations. Again in this scenario we may need to make more abstract analogies where the relations in the two domains are different, but somehow similar. The SME will not allow us to make such analogies even if we take into consideration the more recent developments of the model where re-representations are possible. Another related model is MAC/FAC (Forbus et al., 1995) developed by the same team of researchers. This is a model of analogical retrieval, i.e. how an appropriate base for analogy is retrieved from Long-Term Memory (LTM). This model has two phases: the first one being fast and relatively cheap and selecting a few potential bases for analogy on superficially similarity grounds (the highest number of shared attributes and relations), and the second one being a version of the SME which selects the best structural mapping between the potential bases and the target. This model assumes duplicating representations of each episode in memory, one of the representations used in the first phase, and the other in the second one. The computations requirements for such a model are unrealistic from the perspective of the limitations of human working memory.

Another pair of models of analogy-making is the ACME and ARCS models (Holyoak and Thagard, 1989), developed by a group of researchers led by Keith Holyoak and Paul Thagard. These models overcome some of the limitations of SME by allowing non-identical relations to be mapped and by involving semantic judgments and pragmatic consideration to play a role in the mapping process. However, this model falls into other problems requiring even bigger load on human working memory and being too flexible, i.e. finding mappings that human beings would not consider appropriate. This model is based on the Multiple Constrain Theory which states that the analogical mapping is result of the competition between many potentially contradictory weak constraints and the constraint satisfaction mechanism which finds the near optimal solution. This model has also limitations like the n-ary restriction which allow only relations with the same number of arguments to be mapped on each other, which is not always possible in real life. Another serious limitation is the fact that the semantic similarity judgment is not part of the model but is feeded into the model from outside.

Both these pairs of models are considered to be unappropriate for the robot task described above also for other reasons. In all these models the descriptions of the target and of the bases are hand made and fed into the model, while we need the robot to perceive and encode the scene on its own. We also do not like the pipeline approach of these models where each nest stage is considered separately and as being encapsulated.

The CopyCat and TableTop models (Hofstadter and the Fluid Analogies Research Group, 1995),(Mitchell, 1993),(French, 1995) developed by Douglas Hofstadter and his group seem more appropriate from that perspective, because these are the only models of analogy-making which are building their own representations. Moreover, we fully agree with the statement that analogy-making and perception are highly interconnected and should be understood within a single model (Chalmers et al., 1992), where no separate phases will be isolated. The problem is that these models are dealing only with some aspects of high-level perception, they were developed in abstract domains like letter strings, and so no real visual input was presented to them. This limits the applicability of these models to the robot case, where real camera input from the physical environment is needed. Thus we have incorporated some ideas from CopyCat and TableTop into the model we are developing, but had to extend it further with other mechanisms to allow real vision. Another limitation of the CopyCat and TableTop models is that they lack episodic memory, i.e. these models do not have long-term memory for old experiences, no retrieval from LTM, etc. Thus we preferred to incorporate the perceptual ideas into another existing model of analogy-making, which does have episodic memory and mechanisms for analogical retrieval.

More recently the LISA model (Hummel and Holyoak, 1997) and its derivate DORA were developed by John Hummel and Keith Holyoak. LISA is an attempt to build a model of analogy-making which will be closer to the computational characteristics of the human brain. Thus LISA is a connectionist model based on a combination of localistic and distributed representation. This is a really important step into the right direction, however, it is still far from the real brain mechanisms and at the same time it can hardly deal with the complex analogies that the older models could handle. This model is more realistic from the point of view of the limitations of human working memory and naturally explains them, it has also overcome the n-ary restriction which is present in the above models. This model also integrates the memory retrieval and analogical mapping into a single integrated model. All these are important characteristics. At the same time this model has its own limitations. The main one is the problem with dealing with complex analogies where the specific propositions need to be feeded into the model by hand one by one. Thus the performance of the model will depend on the human assistance to the model. Also the retrieval and mapping phase are still separated and sequential.

Thus after making this review of the existing models we came to the conclusion that the most appropriate approach would be to extend our own AMBR model in order to meet the new requirement of the task. The AMBR model has the advantage that integrates mapping and retrieval in a parallel way thus allowing them to interact with each other. Developing perceptual mechanisms that will be integrated in the same parallel fashion will allow perception to be influenced by the analogy, and not only the analogy to be influenced by what is being perceived. We also rerelied on the IKARUS system for the low-level visual processing, a system developed by Christian Balkenius and hus colleagues at Lund University, which is biologically inspired. Also AMBR has the advantage of being highly context-sensitive and thus efficient in computing only those aspects that are considered relevant at a particular occasion.

The next section describes briefly the AMBR model and then its extensions are presented. Extensions that were developed in order to face the real world robot scenario described earlier.

9.4 AMBR Model of Analogy-Making

AMBR model for analogy-making (Kokinov, 1994a), (Kokinov and Petrov, 2000), (Kokinov and Petrov, 2001) is a multi-agent system, which combines symbolic and connectionist mechanisms. Whereas the behaviour of the model emerges from the local symbolic operations of a huge number of relatively simple interconnected DUAL-agents, the *relevance* of these agents to the current context is represented by their connectionist's activation level. Thus, the meaning is separated from the relevance. From one side, each agent stands for something - an object, a relation, a property, a concept, a simple proposition, a procedural piece of knowledge. There are various types of agents for concepts; tokens; hypotheses for correspondence; anticipations, etc. From the other side, there are two special nodes that are the sources of activation - the INPUT and GOAL nodes - which are representations of the environment and the goals, respectively. Activation (relevance) spreads from these two nodes to other DUAL-agents (typically instances of objects and relations from the target scene), then to their respective concepts and further upwards the concept hierarchy, then back to some of the concept instances and prototypes. Thus, the relevance of the items changes dynamically in response of the changes in the environment and goals and the overall pattern of activation through the knowledge based is assumed to represent the *context*. There is no separation between the semantic and episodic memories - they are strongly interconnected. The active part of the long-term memory forms the Working Memory of the model. Only active agents can perform symbolic operations like for example sending short messages to their neighbours, adjusting their weights, or creating new agents. This is a very important interaction between the symbolic and the connectionist parts of the model. The speed of the symbolic operations depends on the activation level of the respective agent. Thus, the most relevant (active) agents work faster, the less relevant - more slowly, and the irrelevant ones do not work at all. Table 9.1 and 9.2 summarize the main mechanisms and agent-types used in AMBR and describe the role and the routine of each of them in order the model to perform analogies. It is important to note, however, that there is no any central executor, all agents work locally and a-synchronically, and all these mechanisms run in parallel and influence each other.

9.5 Integrating Visual Perception and Motor Control in AMBR

The AMBR model has been tested successfully with various tasks for high-level analogy-making . However, one shortcoming of the model was that always the representation of the target episodes were constructed manually. In addition, the set of the winner hypotheses had to be analyzed by the programmers in order to be

Table 9.1 AMBR basic mechanisms

Spreading activation	The activation of the agents represents their relevance to the current context. It spreads just like in a neural network. The sources of the activation are two special nodes – INPUT and GOAL. The AMBR agents that represent the environment are attached to the INPUT, whereas the representation of the target is attached to the GOAL.
Marker emission and passing	Each *instance-agent* (representing a concrete token) emits a marker that spreads to the respective *concept-agent* (representing type) and then upward to the class hierarchy. When a marker from the target situation comes across a marker from a memorized situation, a *hypothesis-agent* between the two marker-origins is created. The hypothesis-agents always connect two agents and represent the inference that these two agents are analogical.
Structural correspondence s	There are various mechanisms for structural correspondence that create new hypotheses on the basis of old ones. For example, if two relations are analogical, their respective arguments should also be analogical; if two instance-agents are analogical, their respective concepts should also be analogical, etc.
Constraint satisfaction network	The consistent hypotheses support each other, whereas the inconsistent ones compete with each other. Thus, dynamically, a constraint satisfaction network of interconnected hypotheses emerges. After its relaxation, a set of *winner-hypotheses*, which represent the performed analogy, is formed.

interpreted as an 'answer' of the model. Thus, implementing mechanisms for perception and motor control was a big challenge. The tasks and scenarios for robots in rooms, defined in the previous sections, require implementing a connection of AMBR with the environment - which can be real or simulated. However, our motivation was not only to add two separate modules for perception and action, but to use the constraints, defined by the AMBR principles, in order to integrate vision, analogy-making, anticipation formation, and action performing in a single, merged system. The main ideas behind the integrated system is that the ability for analogical mapping is at the core of cognition; that cognition, including perceptual representation of scenes, is highly dynamic and context sensitive; that perceptions, reasoning, and actions are not separate modules, but influence each other in both directions. Thus, vision is assumed as a process that starts from a very poor initial representation. Then, on the basis of large number of initial mappings, anticipations about the possible relations and arrangements are transferred from memory and later on verified. After on more and more complex mappings emerge on the basis of these firstly transferred and verified relations, in turn more and more abstract anticipations are transferred, etc. Actually there is no clear boundary when and where perceptual process is finished and reasoning starts. In the same way, a certain action can be triggered immediately when it seems appropriate for solving the task. For the computer implementation that is organized in a three layers fashion - the world layer, the middle layer, and the reasoning layer. Each layer is implemented by different

Table 9.2 Main types of AMBR agents

Instance-agent	Represents tokens, i.e., particular exemplars. The instance-agents can represent objects, as well as aspects and relations.	*Examples:* bone-1, red-21, behind-3…
Concept-agent	Represents types, i.e., classes of similar exemplars. Again, can represent objects or relations.	*Examples:* bone, color, behind…
Hypothesis-agent	Always connects two elements – one from the target situation and one from a memorized one. Represents an inference that there is something in common between the two elements – they have common super-class or they are the respective arguments of corresponding relations.	*Examples:* bone1<-->bone-3, left-of<-->right-of, red-12<->green-8…
Winner-hypothesis	Represents an already established analogical correspondence between two elements. The hypothesis-agents become winners or fizzle out.	The same form like the hypothesis-agents

independent modules. This allows that any of the three layers can be upgraded or replaced as requirements or technology change, in this way limiting the impact on the others parts of the system. The different layers and their interaction are shown on the next illustration (Figure 9.6).

The World Layer

It can be either simulated, using appropriate software like Webots, or realized with a robot with sensors and motor abilities, living in a real world environment.

Middle Layer

The general purpose of the middle layer is to serve as a mediator between the other two levels, effectively translating and filtering out the information from one layer to the other. Thus, it becomes possible to use different robots or simulators without changing the reasoning layer. In addition, any conceptual improvements on the reasoning layer do not influence directly the world layer. Some of the operations, performed on this layer, are extraction of the needed relations between the objects in the reduced scene representation (see bellow); re-representation of the low-level

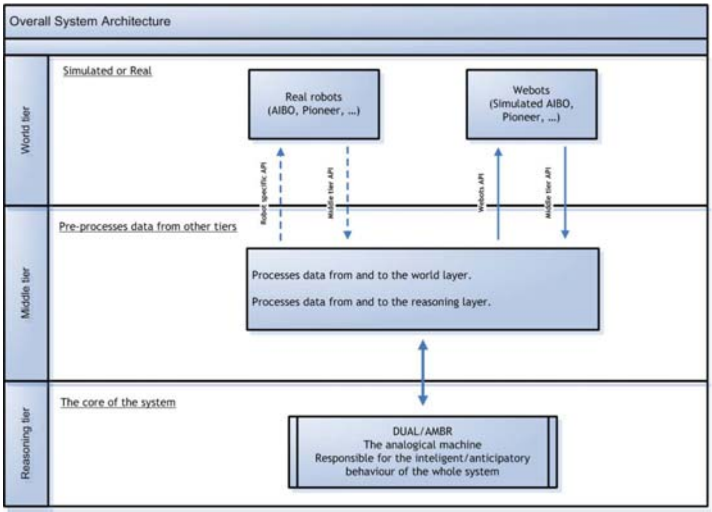


Fig. 9.6 Three layer architecture.

information in a suitable for the next layer form; the inverse operation - transformation of the data from the reasoning layer (the plan) into sequences of low level commands; ensuring the communication between the other layers.

Reasoning Layer (AMBR)

The reasoning layer is the cognitive part of the system. All the information coming from the middle layer is processed here and high level commands are issued backward to it. Here is where the anticipation is built based on a very poor initial description of the current situation from one side, and knowledge of past situations in memory from other side. Here is where the representation of the scene is created dynamically, on the bases of many created anticipations about relations and interpretations, and their serial verifications. Here is also where large analogies emerge, possible solutions are generated and actions are planned. In order to make AMBR a model of the mind of a robot several new mechanisms were developed and implemented. Most importantly, several analogical transfer mechanisms have been developed which will allow robust predictions based on analogy. The present development is related to the extension of the model with a dynamic top-down perceptual mechanism, a mechanism for control of attention, mechanisms for transferring parts from a past episode in memory towards the now perceived episode (analogical transfer), and mechanisms for planning and ordering actions based on that transfer. All new mechanisms are summarized in Table 9.3 while the new AMBR agent's types are given in Table 9.4. Note that all these mechanisms overlap in time and influence each other. It should be stressed that there is no central executive in AMBR. Instead, the AMBR agents interact only with their neighbours and perform all operations locally, with a speed, proportional to their relevance to the current context.

Table 9.3 Advanced ABMR mechanisms

Perceptual Anticipation (top-down influence on perception)	By a series of messages, the instance-agents from memorized situations inform the relevant relations in which they participate for all their hypotheses. If a certain relation collects the hypotheses for all its arguments, it creates an *anticipation-agent*. The anticipation-agents are copies of their mentor-relations but all their arguments are replaced with the respective analogical elements from the target situation.
Attention	The attention mechanism monitors all anticipation-agents, sorts them by their activation (i.e., relevance), and at fixed time intervals asks the perceptual system to check the relation represented by the most active one.
Goal-related Anticipation (Transfer of the solution)	When a certain hypothesis transforms itself into a winner-hypothesis, it informs its base element. The latter, in turn, informs the relations, in which it participates. The respective relations erase all anticipations and hypotheses that are inconsistent with the new winner. Thus, in reality, the anticipation mechanism creates many different possible solutions of the problem that compete with each other, whereas the transfer mechanism works by deleting most of them on the basis of the best analogy. As a final result of the transfer mechanism only the solution that is most consistent with the performed analogy remains.
Action	The *cause-agents* (representing causal relations) are equipped with a special routine. Via special messages, the agents, attached to the GOAL node inform the cause-relations, in which they participate, that the latter are close to the goal. After a period of time, if such 'close-to-goal' cause-agent receives information that it participates in a winner-hypothesis, it checks its antecedents for *action-agents* (representing description of a certain action or movement). If all these conditions are met, the action mechanism sends an order for executing the respective action.

More detailed description of the new AMBR mechanisms follows below.

9.5.1 Top-Down Perception

It is known that when executing a task in order to achieve a specific goal, top-down mechanisms are predominant (Duncan, 1984), (Chalmers et al., 1992). This finding is implemented by making AMBR the initiator of information acquisition actions.

At first, the robot looks at a scene. In order for the model to 'perceive' the scene or parts of it, the scene must be represented as an episode, composed out of several agents standing for objects or relations, attached to the input or goal nodes of the architecture. It is assumed that the construction of such a representation starts by an initial very poor representation (Figure 9.7) built by the bottom up processes. This includes, usually, only symbolic representations of the objects from the scene without any description of their properties and relations. These are attached to the

Table 9.4 Specialized AMBR agents

Anticipation-agent	Represents expectation that a certain relation is present in the environment.	*Examples:* ?red-cube-12?, ?behind-bone-cylinder-12?...
Cause-agent	Represents a certain casual relation. It always has antecedents and consequences. One cause-agent can be instance-agent or anticipation-agent.	*Example:* Cause1 -antecedents: move-12, behind-2 -consequences: find-8
Action-agent	Represents the description of a certain action or movement. The presence of an action-agent in the target situation does not mean that it will be executed. In order for AIBO to execute the respective action, a special procedure for this should be triggered.	*Examples:* Move (AIBO, cylinder-12),

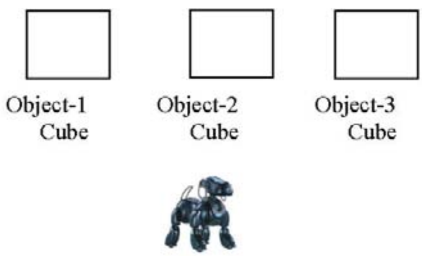


Fig. 9.7 Initial representation of the scene (bottom-up perception).

input of the model (in the example, *object-1*, *object-2*, and *object-3)*. Later on, some of the properties and relations, but only the relevant ones, would be dynamically added to the representation.

The representation of the goal is attached to the goal node (usually *find*, *AIBO*, *bone*). During the run of the system some initial correspondence hypotheses between the input (target) elements and some elements of the memory episodes (bases) emerge via the mechanisms of analogical mapping (Figure 9.8).

The connected elements from the bases activate the relations in which they are participating. The implemented dynamic perceptual mechanism creates predictions about the existence of such relations between the corresponding objects in the scene. As shown in the example of Figure 9.9, Object-1 from the scene representation has been mapped onto Cylinder-11 in a certain old and reminded situation. The activation mechanism adds to working memory some additional knowledge about Cylinder-11 - e.g. that it is yellow and is positioned to the left of Cylinder-23, etc. (Figure 9.9) The same relations become anticipated in the scene situation, i.e. the system anticipates that Object-1 is possibly also yellow and could be on the left of

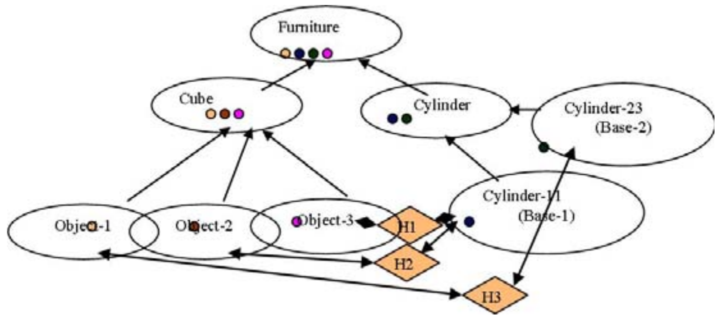


Fig. 9.8 Creation of hypotheses (H1, H2, H3) on the basis of marker intersections.

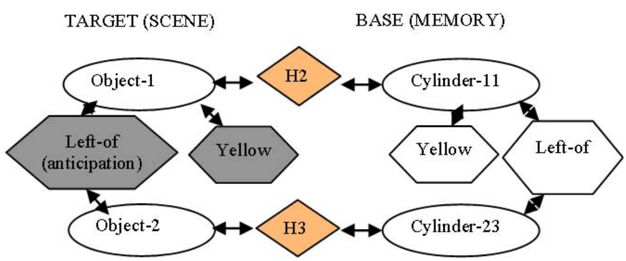


Fig. 9.9 Formation of anticipation agents in AMBR on the basis of missing in the scene arguments of already mapped relations. (Note that 'left-of' relation is asymmetric and the order of arguments is coded in AMBR although it is not shown in the picture)

the element, which corresponds to Cylinder-23 (if any), etc. However, some other mappings between Object-1 and other memorized objects would trigger creations of alternative anticipations. Thus, various anticipation-agents emerge during the run of the system.

9.5.2 Attention

The attention mechanism deals with the anticipations generated by the dynamic perceptual mechanism, described above. With a pre-specified frequency, the attention mechanism chooses the most active anticipation-agents and asks the perceptual system to check whether the anticipation is correct (e.g. corresponds to an actual relation between the objects in the scene). Middle layer, as described earlier, simulates the perceptions of AMBR based on input from a real environment or simulated one. It receives requests from AMBR and simply returns an answer based on the available symbolic information from the scene.

The possible answers are three: 'Yes', 'No', or 'Unknown'. In addition to colours ('color-of' relations), spatial relations, positions, etc., it also generates anticipations like "the bone is behind 'object-1'", or "if I move to 'object-3'", I will find the bone". Those relations play a very important role for the next mechanism - the transfer of the solution (i.e. making a firm prediction on which an action will be based) -

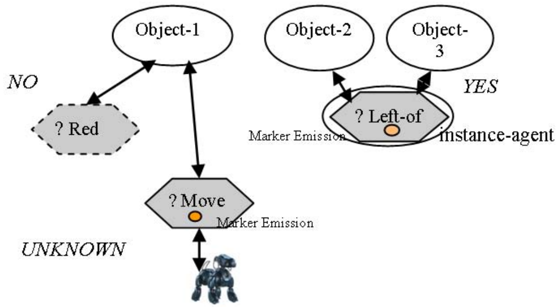


Fig. 9.10 Processing of the different types of answers of relation questions.

as explained below. After receiving the answers, AMBR manipulates the respective agent (see Figure 9.10). If the answer is 'Yes' it transforms the anticipation-agent into an instance-agent. Thus the representation of the scene is successfully augmented with a new element, for which the system tries to establish correspondences with elements from old episodes in memory. If the answer is 'No', AMBR removes the respective anticipation-agent together with some additional anticipation-agents connected to it. Finally, if the answer is 'Unknown', the respective agent remains an anticipation-agent but emits a marker and behaves just like a real instance, waiting to be rejected or accepted in the future. In other words, the system behaves in the same way as if the respective prediction is correct, because this prediction is still possible. However, the perceptual system or the transfer mechanism (see below) can discard it.

9.5.3 Transfer of the Solution

Thus, the representation of the scene emerges dynamically, based on top-down processes of analogical mapping and associative retrieval and of the representation in Middle layer and its functioning. The system creates many hypotheses for correspondence that self-organize in a constraint-satisfaction network (see Figure 9.11)

Some hypotheses become winners as a result of the relaxation of that network and at that moment the next mechanism - *the transfer of the solution* does its job. In fact, the transfer mechanism does not create the agents, which represent the solution. The perceptual mechanism has already transferred many possible relevant relations but now the task is to remove most of them and to choose the best solution. As in the previous examples, let's take a target situation consisting of three cubes and let the task of AIBO be to find the bone. Because of various mappings with different past situations the anticipation mechanism would create many anticipation-agents with a form similar to: "The bone is behind the left cube" [1] This is because in a past situation (sit-1 for example) the bone was behind the left cylinder and now

[1] Note, however, this statement is not represent with a single DUAL-agent, but with a large coalition of agents, following the main principles of the model. For purposes for simplicity only, often larger coalitions are described in the text with single statements.

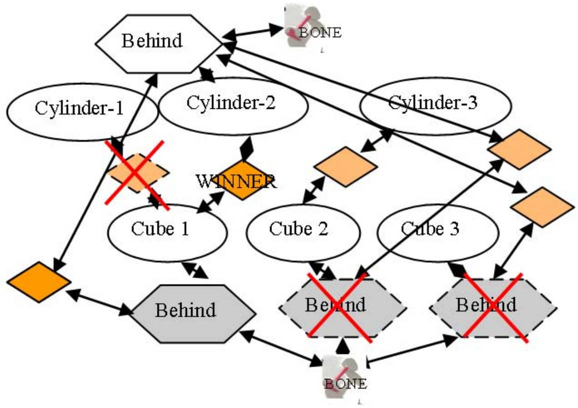


Fig. 9.11 Constraint satisfaction network between hypotheses. Winner hypotheses remove many of the inconsistent anticipations until only few anticipation-agents remain.

the left cylinder and the left cube are analogical. Because of the analogy with another situation, for example, the anticipation that ”the bone is behind the middle cube” could be independently created. Another reason might be generated due to which the right cube will be considered as the potential location of the bone. Thus many concurrent possible anticipation-agents co-exist. When some hypotheses win, it is time to disentangle the situation. The winner-hypotheses care to propagate their winning status to the consistent hypotheses. In addition, the inconsistent ones are removed. In the example above, suppose that sit-1 happens to be the best candidate for analogy. Thus, the hypothesis ’left-cylinder¡–¿left-cube’ would become a winner. The relation ’behind’ from the sit-1 would receive this information and would care to remove the anticipations that the bone can be behind the middle or behind the right cylinder. As a final result of the transfer mechanism, some very abstract causal anticipation-relations like ”if I move to the cube-1 this will cause finding the bone” [2] become connected with the respective cause-relations in the episodes (bases) from memory via winner-hypotheses.

9.5.4 Action Execution

The final mechanism is sending an action command (see Figure 9.12). The cause-relations that are close to the GOAL node trigger it. The GOAL node sends a special message to the agents that are attached to it, which is in turn propagated to all cause-relations. Thus, at a certain moment, the established cause-relation ”if I move to cube-1, this will cause finding the bone” will receive such a message and when one of its hypotheses wins, it will search in its antecedents for an action-agents. The

[2] Represented as a part of the all memory - with one agent for the relation ”cause”, two agents for its arguments ”move” and ”find”, respectively, other agents for the arguments of the arguments - ”I”, ”behind”, ”I”, ”bone”, etc.

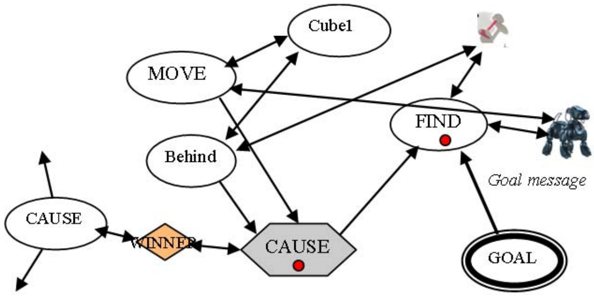


Fig. 9.12 Hypotheses of the cause-relations receive markers from the GOAL node. If the consequents satisfy the goal, then the actions from the conditions are executed.

final step of the program is to request the respective action to be executed and this is done again via a message to Middle layer.

9.6 Running the Simulated Model and Comparing It with Human Data

In the first series of simulations we used only the simulated version of the robot and the environment (Petkov et al., 2007), thus excluding perception and exploring only the role of selective attention. The robot faces several objects in the room and has to build their representation in its mind. Then the task of the robot is to predict behind which object would the bone be and then finally to go to the chosen object and check behind it (Figure 9.13)

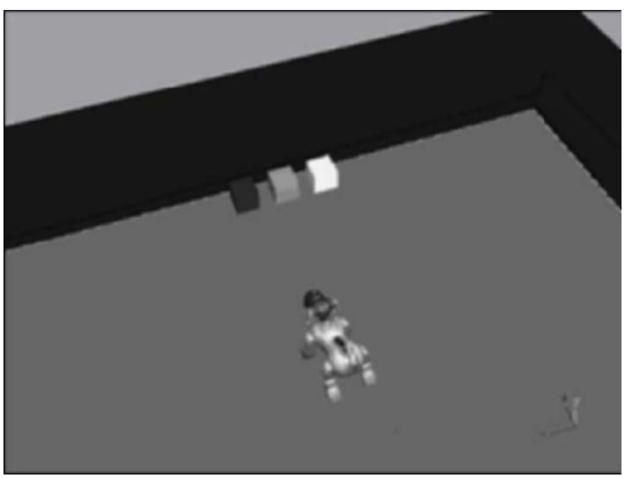


Fig. 9.13 Find an object scenario in a simulated environment with Webots.

We used Webots software to simulate the environment and the body of the robot. Webots is professional mobile robot simulation software, which allows the simulation of physics properties such as mass repartition, joints, friction etc. It is possible to control the robot by setting the position of its body parts (with regard to the joints), as well as to obtain information from the robot's sensors. Each simulated robot can be modeled individually. Webots comes with several models of real robots, like AIBO and Pioneer. This enables transferring tested behavior to real robots. In addition, Webots allows connection with external software modules. The results so far include:

- creation of worlds with physical laws
- creation of object with any shape
- creation of robots and endowing them with behaviour (i.e. programming their behaviour).

Thus there is a representation building part of the model, which target representation is then used for recalling an old episode which could be used as a base for analogy, a mapping between the base and the target is built, and the place of the hidden object in this old episode is used for predicting the place of the hidden bone in the current situation. Finally, a command to go to the chosen object is send. It is important to emphasize that all these processes emerge from the local interactions of the micro-agents, i.e. there is no central mechanism that calculates the mapping or retrieves the best matching base from memory. In the simulations described here the AIBO robot had four specific past episodes encoded in its memory, presented in Figure 9.14 . In all four cases the robot saw three balls and the bone was behind one of them. The episodes vary in terms of the colors of the balls involved and the position of the bone.

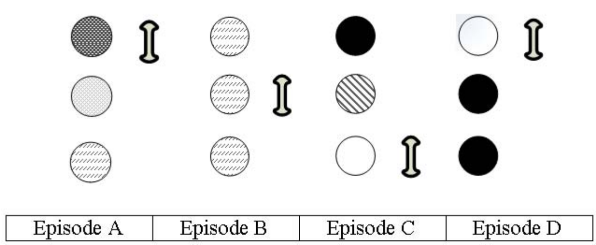


Fig. 9.14 Old episodes in the memory of the robot (different colors are represented with different textures).

The robot was then confronted with eight different new situations in which it had to predict where the bone might be and to go and check whether the prediction was correct (Figure 9.15). The situations differ in terms of colors and shapes of the objects involved.

In Figure 9.16 one can see the representation of the target situations that is extracted from the description of the simulated environment. (Representation build-

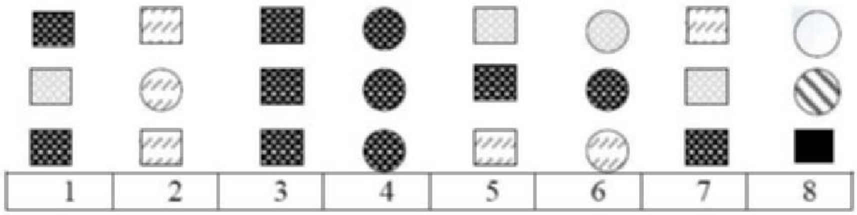


Fig. 9.15 New tasks that the robot faces.

ing for perceived real environment is described in the next section.) For the first series of simulations, however, the representation involves relations known to the robot such as color-of (object-1-sit001, red1), same-color (object-1-sit001, object-3-sit001), unique-color (object-2-sit001), right-of (object-2-sit001, object-1-sit001), instance-of (object-1-sit001, cube), etc. . The relations are in turn interconnected in a semantic network. For example, same-color and same-form are both sub-classes of the higher-order relation same.

In the simulations described above the attention of the robot was simulated by connecting only some of these descriptions to the input list which results that even though all relations, properties, and objects will be present in the Working Memory (WM) of the robot, only some of them will receive external activation and thus will be considered as more relevant. Thus different simulations with the same situation, but focusing the attention of the robot towards different aspects of the given situation, could result in different outcomes.

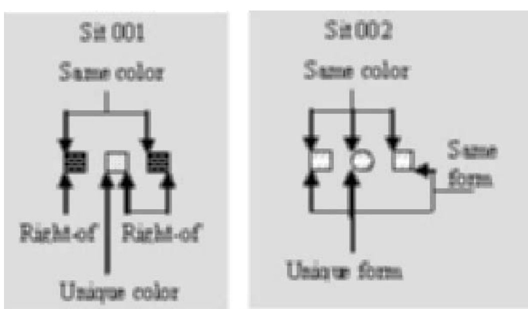


Fig. 9.16 Representation of the target situations 1 and 2.

In each case there could be various solutions: different analogical bases could be used on different grounds and in some cases for the same base several different mappings could be established that will lead to different predictions (See Figure 9.17 and Figure 9.18 for the specific mappings established and the predictions made).

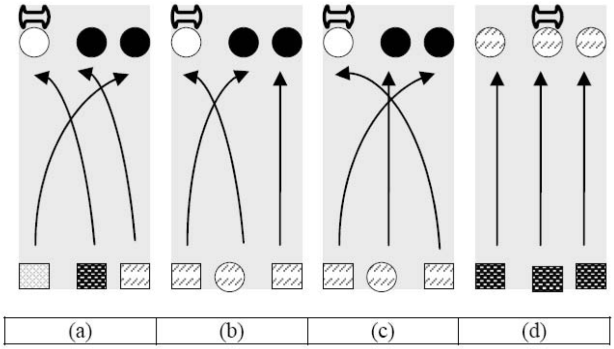


Fig. 9.17 Examples of mappings established with changing the attention from form (a) and (b) to color (c) and (d).

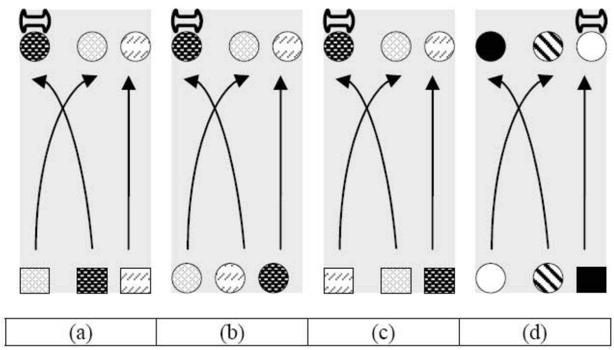


Fig. 9.18 Examples of mappings based on the superficial color relations

9.6.1 Comparing with Human Data

After running the first series of simulations several times varying only the focus of attention to see whether the mapping changes; we conducted a psychological experiment. We showed the bases to the participants, changing the AIBO robot and the bone with a cover story about a child who has lost its bear-toy. We asked the participants to predict where the bear-toy would be in the given new situation. The data from the human experiment are given in Figure 9.19a. As one can see there is a variety of answers for almost each target situation. Still there are some dominating responses. In order to be able to test the robot's behavior against the human data, 50 different knowledge bases have been created by a random generator that varies the weights of the links between the concepts and instances in the model . After that the simulation has been run with each of these knowledge bases in the "mind" of the robot.Figure 9.19b reflects the results. They show that the model has a behavior which is quite close to that of the participating human subjects in terms of the dominating response. The only major difference is in situation 2 where hu-

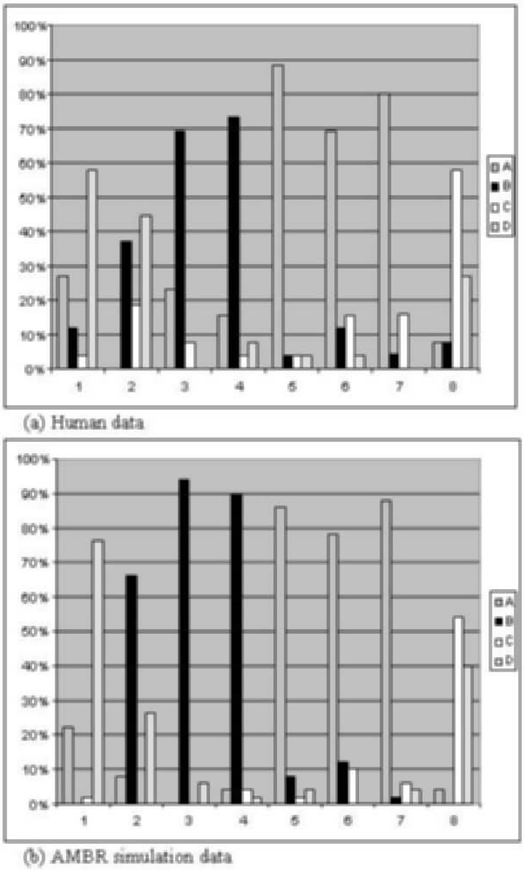


Fig. 9.19 Comparing human and simulation data: which base has been used for analogy with each target situation and how many times.

man subjects are ”smarter” than AMBR: they choose an analogy with situation D (same-form goes onto same-color) much more often than AMBR. Still AMBR has produced this result in 25% of the cases. This means that AMBR is in principle able to produce this result, but it would requite some tuning of the model in order to obtain exactly the same proportion of such responses.

9.7 Running the Real Robot Model in the Real World

As in the simulated version we define our tasks and test the model in a house-like environment and in a ”find-an-object” scenario. There is one room with some objects in it. The objects vary in shape- cubes , cylinders and in colors. We used Sony AIBO robots (ERS-7). The goal of the robot is to find a bone (or bones) hidden under an object. All objects are visible for the robot. (Figure 9.20).

Fig. 9.20 Real world scenario -"Where is my bone?".

The AIBO robot has to predict where the bone is hidden based on analogies with past episodes and go for it. The episodes are manually built for the moment, but we work on the learning process by which the newly perceived situations will remain in LTM.

In order to simplify the problems related to 3D vision we decided to have one camera attached on the ceiling having a global 2D view of the scene. There is a colour marker on the top of the AIBO to facilitate its recognition. A web camera server sent the image data via TCP/IP to the network camera module of IKAROS. All software is run on remote computers and communicate with the robot through wireless network.

In order to connect AMBR with the real world several new modules were developed (Kiryazov et al., 2007). A major step was to build a real world perceptive mechanism with active vision elements. Several modules of IKAROS system (Balkenius and Moren, 2003), (Balkenius et al., 2007), (Kiryazov et al., 2007) related to perception were successfully integrated in order to carry out the difficult task of bottom-up visual perception and object recognition. Another module - AMBR2Robot - was developed to mediate between the perception modules of IKAROS and ABMR. AMBR2Robot supports the selective attention mechanisms, which were described above.

The resulting architecture consists of the following modules (see Figure 9.21):

- AMBR - the core of the system, it is responsible for attention and top-down perceptual processing, for reasoning by analogy, for decision making, and for sending a motor command to the robot controller.
- IKAROS module - a low-level perception module performing bottom up information processing.
- AMBR2Robot - a mediation module, the link between AMBR and IKAROS and the robot controller.
- AIBO robot.
- Camera attached to the ceiling.

The camera takes visual information of the environment. It is received by the IKAROS module. The visual information is processed and symbolic information about objects in the environment is produced. This symbolic information is used from ABMR2Robot to provide AMBR with bottom-up perception information and also to handle the top-down requests which are described below. ABMR2Robot also waits for a ”do-action” message from AMBR, which when received makes the module to control the robot and guide it to the target position using AIBO Remote framework. AMBR does the substantial job of making predictions about where the bone is hidden based on the representation of the current situation and making analogy with past situations. AIBO Remote Framework is a Windows PC application development environment which enables the communication with and control of AIBO robots via wireless LAN.

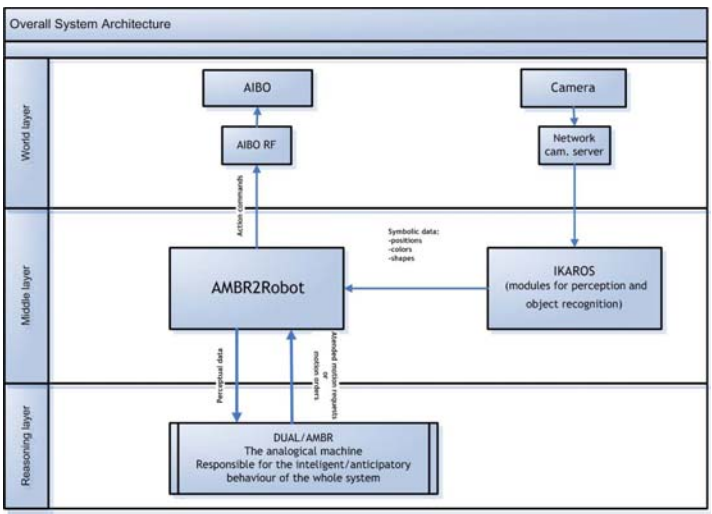


Fig. 9.21 Main modules and data flow between them

The newly developed modules are described in details below.

9.7.1 Ikaros

IKAROS is a platform-independent framework for building system-level cognitive and neural models (Balkenius and Moren, 2003), (Balkenius and Johansson, 2007) (see also www.ikaros-project.org). The system allows systems of cognitive modules to be built. The individual modules may correspond to various cognitive processes including visual and auditory attention and perception, learning and memory, or motor control. The system also contains modules that support different types of hardware such as robots and video cameras. The modules to be used and their connectivity are specified in XML files that allow complex cognitive systems to be built by the individual modules in IKAROS. Currently, there are more than 100 different modules in IKAROS that can be used as building blocks for different models.

In the present work, IKAROS was used for visual perception and object recognition . An IKAROS module receives images from a camera while another module segments the image into different objects based on colour. The result is sent to AMBR2Robot for further processing.

The object recognition module proceeds in the following stages:

1. RGB image is mapped onto RG- chromacity plane to reduce effects of different illuminations and shadows on the objects.
2. Edge filter is applied to find the contours of the objects.
3. Colours of the image are normalized between edges to produce homogenous colour regions (Figure 9.22).
4. Pixel are segmented into clusters with different colours (each colour is defined as a circle sector around the white-point in the RG-chromaticity plane) (Figure 9.22).
5. Each object is localized in a rectangular region. A histogram of edges direction is produces in that region. The histogram is processed to find the object shape (Figure 9.23).

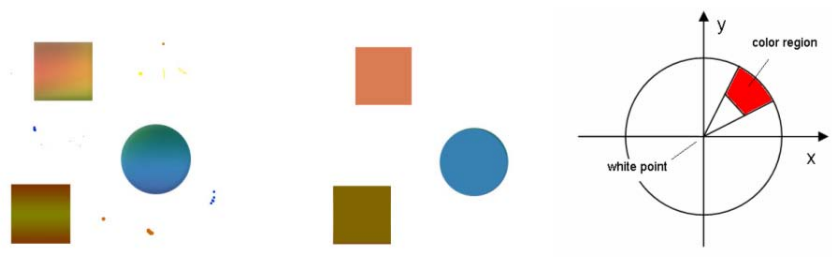


Fig. 9.22 Left. An initial image. Middle. Colours after normalization between edges. Right. A colour region in the RG-chromaticity plane

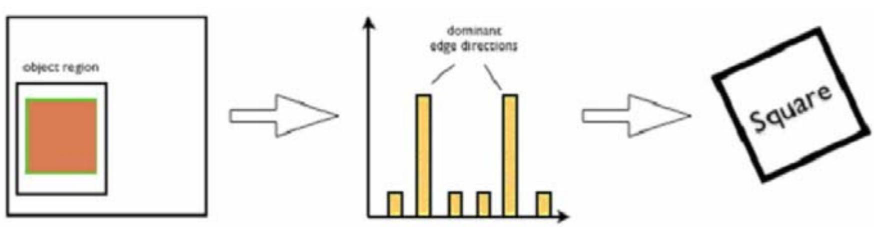


Fig. 9.23 The recognition of shapes. Left. A region is formed around the colour cluster. Middle. A histogram of edge orientations is calculated. Right. The distribution of edge orientations is used to determine the shape.

Note that we are not trying to find the complete contours of the objects. Instead, the method is based on the distribution of different edge orientations which is a much more robust measure. The different processing stages were inspired by early visual processing in the brain but adapted to efficient algorithms. The mapping to the RG-chromaticity plane discards the illuminant and serves the same role as the interaction between the cones in the retina (Dacey, 1996). The detection of edges is a well known function of visual area V1 (Hubel and Wiesel, 1968). The colour normalization within edge elements was inspired by theories about brightness perception (Paradiso and Nakayama, 1991) and filling-in (Grossberg, 2000).

9.7.2 AMBR2Robot

AMBR2Robot provides AMBR with perceptual information from IKAROS and also serves for implementing the selective attention mechanism in the model. The other main purpose of this module is receiving the action tasks from AMBR and executing them using AIBO-RF. We could say that it simulates the link between the mind (AMBR) and the body (perception system, action system). The work of the module AMBR2Robot formally can be divided into three sub- processes:bottom-up perception, top-down perception, performing actions.

Bottom-Up Perception

At this stage just a small part of the scene-representation is sent to AMBR. As described above information is further transformed into the form used for knowledge representation in AMBR by creating a set of AMBR agents with appropriate slots and links and connecting them to the so-called input of the architecture.

Top-Down Perception

As mentioned above AMBR sends top-down requests in the form of questions about the presence of properties and relations about the identified objects. These requests are received by AMBR2Robot and are answered based on visual symbolic information provided by IKAROS. Relations represent the most important information for analogy-making and are extracted by AMBR2Robot from the scene description which does not contain them explicitly but only implicitly (e.g. in the form of coordinates and not spatial relations).

The main types of top-down perception requests are for:

1. spatial relations: right-of, in-front-of, in-front-right-of, etc. ...
2. sameness relations: same-colour, unique-shape, etc. ...
3. colour properties: orange, blue, red, etc. ...

The spatial relations are checked based on the objects' positions as described by their coordinates and size and with respect to the gaze direction of the robot. Figure 9.24 shows how the above example relation request (left-of object-2 object-3) is processed. Positions of all objects are transformed in polar coordinates respective to

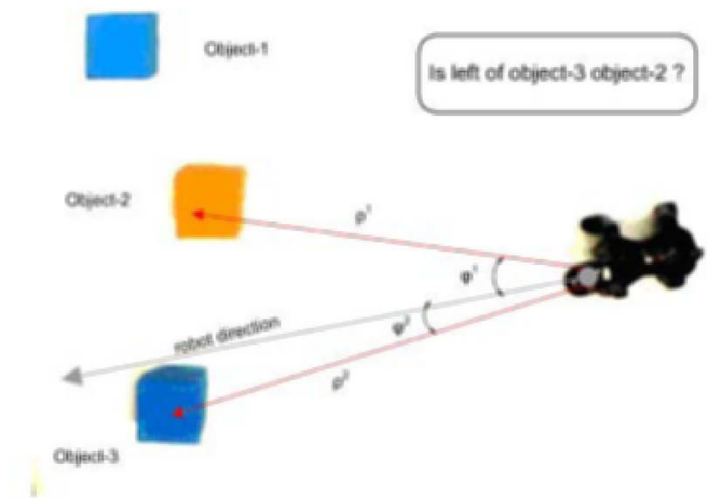


Fig. 9.24 Processing the spatial relation requests.

a robot-centric coordinate system. Then some comparison rules are applied to the coordinates. For processing the sameness relation the relevant properties (shape or colour) of all the visible objects are compared.

Action

AMBR2Robot receives action commands from AMBR and, knowing the positions of the target object and the robot, navigates AIBO by sending movement commands via the AIBO Remote Framework (see Figure 9.21). During the executed motion, IKAROS is used to update the objects' position in the scene (actually all object except the robot are stationary) The robot is guided directly to the target object without any object avoidance (to be implemented in the future in more sophisticated examples). After the robot has taken the requested position, it is turned in the appropriate direction to push and uncover the target object. At the end it takes the bone if it is there or it stops otherwise.

9.7.3 Tests

The results from a single run of the system will be described below. There are two past situations in the robot's memory (Figure 9.25a) The robot faces the situation showed in Figure 9.25b

The image (from the global camera) is sent to the IKAROS module. In Figure 9.26 the visual field after RG-chromaticity transformation and the edge histogram for the one of the recognized objects are shown.

The IKAROS module recognizes the objects and produces the input:

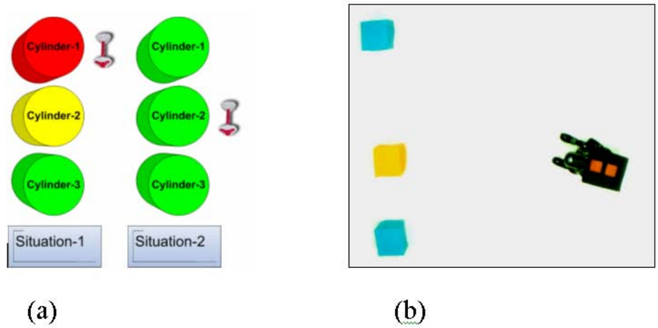


Fig. 9.25 (a) Old episodes in memory (b)AIBO is in a room with three cubes with different colours.

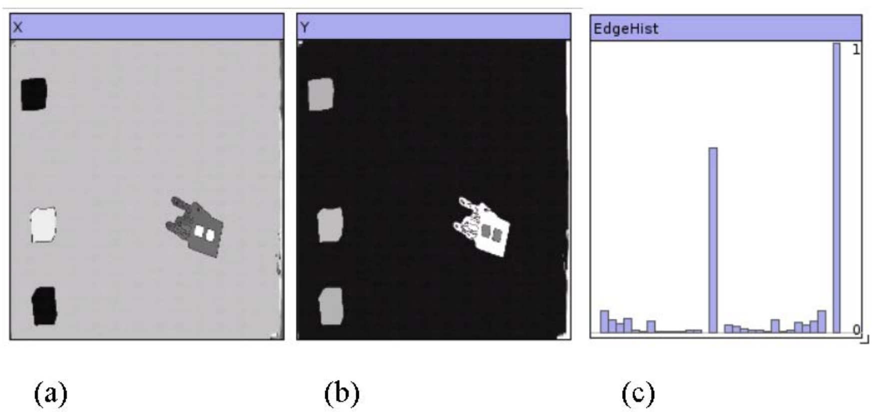


Fig. 9.26 (a),(b) RG-chromaticity transformation (c)Edge histogram for the upper cube

```
object-1, shape=cube, position= (50,85)
object-2, shape=cube, position=(71,303)
object-3, shape=cube, position=(75,429)
aibo-I, position=(438,301), direction=3.387
```

Processing that information AMBR2Robot sends part of it as bottom-up perceptual information to AMBR:

```
object-1: cube, object-2: cube, object-3: cube
```

In the top-down perception part, lots of relation requests are generated from AMBR. Here we show some of them, including the answers from AMBR2Robot:

```
? behind: bone-t -> UNKNOWN
? left-of: object-1 object-2 -> YES
? blue: object-3 -> YES
? green: object-2 -> NO
? in-front-right-of: object-3 aibo-I -> NO
? move: aibo-I object-2 -> UNKNOWN
```

Some of the already created anticipation agents are turned into instance agents according to the answer:

```
anticip-blue-situation-1
anticip-left-of-1-situation-2
```

The name of the agents is formed as follows. 'anticip'- stands for anticipatory . After that follows the name of the relation which it "anticipates". This relation can belong to one of the situations in robots memory (situation-1 or situation-2 in this case) Note that after transforming an anticipation agent into instance one its name remains the same. After some time AMBR make an analogy with situation-1. Some of the winner hypotheses (in the order they emerge) are:

```
object-3 <-> right-cylinder-situation-1
anticip-blue-situation-1 <->blue-situation-1
aibo-I <-> aibo-I-situation-1
anticip-move-situation-2 <->move-situation-1
object-2 <-> midle-cylinder-situation-1
object-1 <-> left-cylinder-situation-1
```

Many other agents are mapped. After the mapping of the cause agent, the action mechanism is triggered, which sends a motion command to AMBR2Robot.

```
move-to object-3
```

AMBR2Robot guides the robot to the target. Once arrived, the robot uncovers the object and tries to get the bone. Figure 9.27 shows some images of the robot moving.

9.8 Mechanisms for Active Vision

Dynamic Representation of the Objects

A step toward improving the model would be to create more plausible mechanism for attentual moving. In the current version of the system it is assumed that all

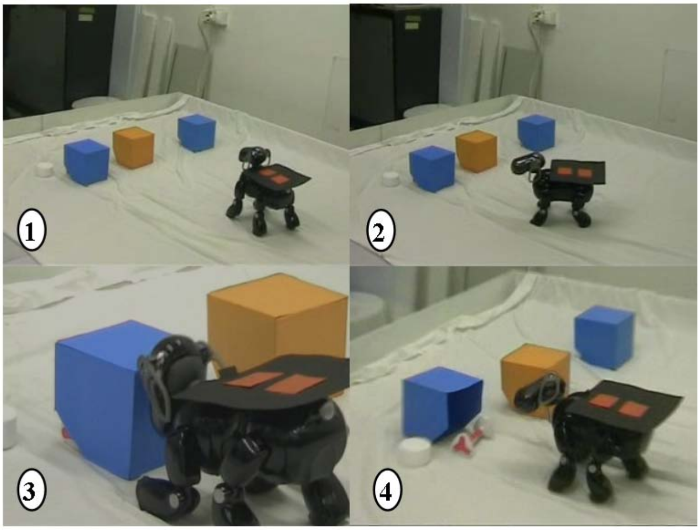


Fig. 9.27 1. AIBO is starting to move. 2. approaching the target. 3. uncover object. 4. get the bone

objects are pre-attentively and simultaneously represented. Thus, actually, the objects serve for primitives of the system. This unrealistic assumption should be replaced with a system for gradually context-dependent exploration of the objects in the scene, together with the relations between them. Thus, some objects that are outside of the robot's visual angle, as well as the occluded objects should be not initially represented. (Figure 9.28)

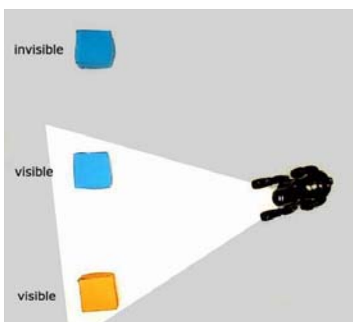


Fig. 9.28 Filtering the visual field related to robots first point of view

Bottom-Up Perception

The coin, however, has its opposite site too. By making the initial scene representation poorer and poorer, the overall behavior becomes more and more dependent on the goals only. Without decreasing the importance of the top-down processes in vision, some direct bottom-up mechanisms for perceptual grouping should be added, as well as their interaction with the top-down processes should be studied.

Active Vision

The ability to robot to move continuously and the dynamic of the input information makes the process more complex and difficult to be implemented. Sometimes, however, probably the additional constraints can make the problems easier to be solved. For example, in the case of an object, partially occluded from another object, recognition is impossible from the static two-dimensional picture. On the contrary, if the robot moves, the task can be much easier. Thus, an important step for improving the system should be to design more action plans for vision only and additional usage of the influence on actions on reasoning and perception.

9.9 Discussion and Conclusion

This chaper presented an integrated approach to anticipation starting from perceiving the scene all the way down to making a decision and undertaking an action. In this case the anticipation is based on analogy with a previously experienced situation. This is a new approach to anticipation and at the same time a new challenge for the AMBR model of analogy-making which has been extended with new mechanisms related to perception and anticipation which are integrated into the whole architecture and following the same basic principles.

Chapter 10
Anticipation in Coordination

Maurice Grinberg and Emilian Lalev

10.1 Introduction

The games defined in formal game theory (like e.g. the Prisoner's Dilemma game) are widely used to model social interactions (Colman, 2003). Recently, several influential research efforts (e.g. Axelrod (1984) and Epstein and Axtell (1996)), based on Multi-Agent Simulations (MAS), have been carried out successfully in order to explain (and even try to influence) such important aspects of societies like cooperation and competition. The typical framework of such approaches consists of the use of simple agents interacting with an environment via simple rules or game playing. Although the phenomena arising in such environments are important enough to deserve detailed investigation, we have adopted a different approach here. We have been interested in cognitively plausible agents whose performance can be compared against experimental data from human participants.

The use of cognitively sophisticated agents can be regarded as a development of the opposition of standard game theory and the bounded rationality framework (Colman, 2003). In standard game theory, players are described as perfectly rational and possessing perfect information about the game including knowledge about the possible moves and payoffs, and opponents. On the other hand, the bounded rationality view on cognition states that people are almost never perfectly rational (Colman, 2003) due to limitations in perception, time, thinking, and memory. Moreover, people tend to minimize the cognitive effort while making decisions. Finally, the results of experiments involving games demonstrate that people rarely play as prescribed by the normative game theory. One such famous example is the Prisoner's Dilemma (PD) game which will be dealt with in this Chapter.

The influence of cognitive constraints and mechanisms on decision making in the Iterated Prisoner's Dilemma Game (IPDG) and thus on the simulations describing social interactions has been studied for instance in a series of investigations (see Hristova and Grinberg (2004) and Lalev and Grinberg (2007) and the references therein). Moreover, the use of of cognitively plausible agents can insure that the information gained by using them in simulations of complex social interactions will

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 215–236, 2008.
© Springer-Verlag Berlin Heidelberg 2008

take into account specific cognitive mechanisms which are essential for the explanation of the phenomena observed.

One such cognitive mechanism is anticipation and its role in explaining cooperation and coordination will be the focus of our results and discussions in this Chapter. Special attention will be devoted to the use in MAS of the anticipation model proposed by Lalev and Grinberg (2007), where the role of anticipation on cooperation in IPDG has been investigated. The detailed analysis of the model features and the comparison with previous experiments with human participants demonstrated the importance of prediction for adequate description of the behavioral data on cooperation. These results were obtained in the experiments and in the theoretical frameworks by using individual playing against a tit-for-tat opponent focusing on individual decision making. Here, we want to present results which demonstrate the role of anticipation in small societies of agents. The key characteristics monitored will be cooperation and coordination as related to the essence of social interaction as discussed in other chapters of these book.

10.1.1 The Prisoner's Dilemma Game

The Prisoner's dilemma is a two-person game and a famous example of a social dilemma game. The payoff table for this game is presented in Table 10.1. The players simultaneously choose their move - C (cooperate) or D (defect), without knowing their opponent's choice.

Table 10.1 Payoff table for the PD game. In each cell the comma separated payoffs are the Player I's and Player II's payoffs, respectively.

		Player II	
		C	*D*
Player I	*C*	R, R	S, T
	D	T, S	P, P

In Table 10.1, R is the payoff if both players cooperate (play C), P is the payoff if both players "defect" (play D), T is the payoff if one defects and the other cooperates, S is the payoff if one cooperates and the other defects. The payoffs satisfy the inequalities T $>$ R $>$ P $>$ S and 2R $>$ T + S. This structure of the payoff matrix of that game offers a dilemma to the players: there is no obvious best move. The dominant D move (T $>$ R and P $>$ S) would lead to lower payoffs if adopted by all the players (payoff P) although this is the choice prescribed by standard game theory. Cooperation seems to be the best strategy in the long run (R $>$ P) but at the risk of one of the opponents to start to defect and the other to receive the lowest payoff S. This quite complicated situation is at the heart of the dilemma in this game and is the reason for the on-going interest in this game over the past 50 years and continuing today. The importance of the possibility to predict the opponents' moves is obvious especially in the iterated version of the game. Reliable prediction would lead in some cases to trust in the opponent and higher cooperation while in other cases to 'punishment' of expected defection. In any case, anticipatory agents playing IPDG will be involved in specific interactions, which have to be investigated.

Rapoport and Chammah (1965) proposed the quantity CI = (R-P)/(T-S), called cooperation index, as a predictor of the probability of C choices, monotonously increasing with CI. In Table 10.2, two examples of PD games with different CI 0.1 and 0.9, respectively are presented.

Table 10.2 Examples of PD game matrices with different CI - 0.1 and 0.9, respectively. The first payoff in each cell is the payoff of the 'row' player and the second of the 'column' player.

CI=0.1		Player II	
		C	*D*
Player I	*C*	56, 56	0, 60
	D	60, 0	50, 50

CI=0.9		Player II	
		C	*D*
Player I	*C*	56, 56	0, 60
	D	60, 0	2, 2

10.2 Related Research

A common assumption is that, in the long run, people build mental models of themselves and of other people they interact with. Such models include grasping typical aspects of their behavior and this may result in establishing relations of trust or distrust. It is also agreed that more or less pure instances of IPDG can be observed in real-life social relations. As long as these relations matter for the well-being of people, they try to make models of the other players and of the environment. These models are guided by past experience and the actions assigned to interactions comply with predictions about the other agents' actions. In other words, past experience and predictions about events based on this experience are factors which cannot be neglected in the understanding of human social interactions (e.g., Leydesdorff and Dubois (2004)).

In simulated societies of anticipatory agents, effects of the social interactions such as reputation formation and strategic teaching in social dilemmas (Camerer et al. (2002), Taiji and Ikegami (1999), Isaac et al. (1994)), or coordination between agents (Dittrich et al., 2003) were observed. Leydesdorff and Dubois (2004) gave examples of how information and meaning processing become interweaved in a society of very simple agents with anticipatory capabilities.

10.2.1 Fictitious Play

Awareness of the presence of an opponent in IPDG consequently implies that the player tries to make a model of the opponent's strategy which leads to a more complicated behavior related to attempts to utilize this model (see Sutton and Barto (1981a)). Provided players are able to predict the opponent's move, they may want to maximize their payoff by choosing the most profitable move given the predicted opponent's move by fictitious play (Brown, 1951). Fictitious play is a behavior in which a player evaluates reinforcements from situations that did not actually happen but were only imagined (see Camerer and Ho (1999)). Agents choose to play the action (i.e., the pure strategy) that maximizes their expected payoff given their estimate of the opponent's strategy (Lipson and Leyton-Brown, 2006).

Fictitious play may refer to the past interactions with the opponent like in Camerer and Ho (1999). In this case, its purpose is to explore possible strategies that may have earned higher foregone rewards. Provided fictitious play refers to the future of game interactions, it serves for the players to make the optimal decision about their future behavior like in Taiji and Ikegami (1999).

MAS with a variety of reinforcement learning agents (conducted by Asher Lipson and Kevin Leyton-Brown, 2006) revealed that fictitious play agents were best at converging to a Nash equilibrium (Lipson and Leyton-Brown, 2006).

10.2.2 Strategic Teaching and Reputation Formation

When players assume their opponents are adaptive and are influenced by their strategy, players may try to create reputation for themselves (Camerer et al. (2002), Taiji and Ikegami (1999)). Reputation can be thought of as an incentive to guide the coordination between two interacting agents. Reputation in social dilemmas is a kind of abstract resource during iterated play. It assures that the players on the other side will behave well as it is a factor for higher cooperation rates.

Reputation may also be considered a currency to other games (Milinski et al., 2002) that are played in an alternating mode at the same time by the same players. The knowledge of being recognized as the same individual in both scenarios motivates players to invest in their reputation. Sophisticated relations, including theory of mind and reputation formation, are related to anticipation in decision making (Rosen, 1985). They provide explanation of cooperation in IPDG based on forward-looking decision-making.

A model that uses reputation is the Experience-Weighted Attraction (EWA) model (Camerer and Ho, 1999) in its ”sophisticated” version. The EWA model combines elements of two approaches. First, realization of the belief learning approach in EWA includes considering foregone fictitious play payoffs. The second approach in EWA is the standard reinforcement learning.

In the sophisticated EWA (Camerer et al., 2002), players use this model to forecast what the other players will do and choose strategies with high expected payoffs given their forecast. Because the model assumes that sophisticated players think others are sophisticated (and those others think others are sophisticated, ...), it creates a whirlpool of recursive thinking which nests equilibrium concepts (Camerer et al., 2002). Sophisticated players playing iteratively with the same opponents usually have an incentive to ”teach” adaptive players by choosing strategies with poor short-run payoffs which will change what adaptive players do, in a way that benefits the sophisticated player in the long run. This ”strategic teaching” gives rise to repeated-game equilibria and reputation formation behavior through the interaction between the players (Camerer et al., 2002).

Another example of reputation formation may be seen in the the Best-response with signaling (BRS) model that was suggested by Isaac et al. (1994) for the play in the Public Goods social dilemma game. It incorporates rational decisions with features of forward-looking best-response and signaling of cooperative intentions. According to the BRS model, players may benefit from signaling their cooperative intentions to other players. Signaling is an increase in the cooperation rate of a player, and it serves to propose to the opponents to cooperate more in oncoming games.

Taiji and Ikegami (Taiji and Ikegami, 1999) proposed a connectionist model architecture especially designed to investigate cooperation among anticipatory players in IPDG. Strategic teaching and reputation formation naturally emerge in the model provided that it possesses first or second order intentionality. The model is able to recognize the dynamics of the opponent’s strategy, and to make predictions about the opponent’s future moves with the help of a dynamic recognizer (Pollack, 1990). Therefore, it has an internal model of its opponent and uses it in predicting its behavior.

The model of Taiji and Ikegami exists in two versions which are called ”Pure Reductionist Bob” and ”Clever Alice”. Pure Reductionist Bob anticipates the opponent’s strategy by the dynamic recognizer. He believes that the opponent behaves according to simple algorithms like finite automata and is trying to infer these algorithms from her behavior.

Clever Alice in turn assumes that the opponent behaves like Pure Reductionist Bob. She knows that the opponent is making her model and he decides the next action based on that model of hers. In other words, she builds the model of herself and treats this model as her image in the opponent. In order to decide the player’s next action, a prediction of the opponent’s future action based on the dynamic recognizer apparatus is used in both versions of the model. The algorithm of Pure Reductionist Bob is that he chooses his forward actions up to several fictitious future games. He can predict the opponent’s actions from his forward actions, and the expected score is evaluated. The process is repeated for all possible strings of actions of a given length and Bob chooses the action with the highest score as the best action.

Clever Alice chooses her forward actions and predicts the opponent’s actions assuming that he behaves like Pure Reductionist Bob. Again the process is repeated for all possible variations of future strings of C and D moves and she chooses the action string with the highest score as the best one. In other words, she predicts her image in the other person and tries to educate him to have a favorable image of her through her actions.

In simulations where the model played against the same model (not against simple computer strategies), the IPDG play always converged to mutual defection after some time. In other words, in this case anticipation was not of help for players to play cooperatively in the long run. However, the model managed to discover the cooperation strategy as the most profitable one against a Tit-For-Tat computer opponent.

10.2.3 Social Order and Coordination

A key concept for Dittrich et al. (2003), that used communicating agents to model and simulate the origin of social order, is the situation of double contingency (Luhmann, 1984). Social structures, social order, or social systems are first of all structures of mutual expectations. Every entity expects that the other entity has expectations about its next activity (Dittrich et al., 2003). There are two factors into the decision process of an agent in their simulated societies: The first is its expectation about the future and the second is its expectation about the other agent's expectation (called "expectation-expectation" by Luhmann (1984). Simulation experiments of the model reveal that social order (coordination) appears as a rule in the dyadic situation. In this case, agents succeed to establish good coordination between each-other.

On the other hand, in a population of many interacting agents, the order usually disappears. Scalable order in larger societies only emerges for very specific cases. One case is if agents generate expectation-expectations based on the activity of other agents, not only on their own activity. A second case is if there is observation of others (Dittrich et al., 2003).

Coordination in the society can lead to the transition from a more actor-oriented perspective of social interaction to a systems-level perspective.

10.2.4 Anticipation and Information Processing in Societies

With much simpler anticipatory agents Leydesforff (see Leydesdorff and Dubois (2004) and Leydesdorff and Dubois (2004)) managed to demonstrate complicated effects and derived inferences about information processing in the society. The idea implemented is that anticipation is a process opposite to recursion called incursion (Dubois, 1998). The direction of this process is opposite to the time line - the predicted state influences the current one. Agents, due to their anticipatory properties, may choose one of two paths when they appear in a bifurcation. Such behavior resembles the binary choice decisions made by agents in IPDG (Lalev and Grinberg, 2007). Thus, the degree of insecurity in the whole system is reduced.

Leydesdorff and Dubois (2004) defines two contradictory couples of entities: recursion and information versus incursion and meaning. In social anticipatory systems, observations and expectations can be exchanged between agents. The meaning of the current action of an agent becomes evident through anticipation. Meaning is provided to observations from the perspective of hindsight, while information processing follows the time axis (Leydesdorff, 2005).

10.3 Agent Architecture and Decision Making Model

Most of the models presented above are quite schematic and rely on specific pre-specified mechanisms. From a cognitive modeling point of view the challenge is to understand the decision making mechanisms that would lead to the results observed in the experiments with human participants taking account all of the important characteristics (e.g. the dependence of cooperation on CI or of cooperation on the level of predictive capabilities). We are convinced that an adequate agent model should have a minimal but sufficient level of complexity and should perform in a environment similar to the environments of human experiment participants (e.g. they should perceive the payoff matrix of the game before making a move and take into account the opponent's moves and game outcomes). In the same time human players rely on past experience and on predictions of future events.

The model presented here, following Lalev and Grinberg (2007), is aimed at complying with these requirements. It has taken into consideration the results from extensive recent theoretical and experimental research on the cognitive processes involved in decision making in IPDG (see Hristova and Grinberg (2004) and Lalev and Grinberg (2007)) using different approaches involving psychological experiments, eye-tracking studies, and modeling and simulations.

10.3.1 The Model

The core architecture of the model (hereafter referred to as Model A following Lalev and Grinberg (2007)) is an Elman recurrent neural network (Elman, 1990) depicted in Figure 10.1. In Taiji and Ikegami (1999), a recurrent network has also been used to model the behavior of PD game players. However, the model used here has a much more complicated structure and includes as input the game payoff matrix, the players' previous moves and the received payoffs (related to the specific last game outcome). The network consists of eight inputs, thirty hidden-layer, and six output nodes (see Figure 10.1). The activation functions of the hidden layer and of the output layer are tan-sigmoid and log-sigmoid functions, respectively. Because of the logistic output activation function, some of the network's outputs are interpreted as probabilities.

Inputs and Outputs All the inputs of the network were rescaled to be within the range [0, 1]. As can be seen from Figure 10.1, the values of the payoffs from the current game matrix (excluding the payoff S which was always zero in this specific set of games, taken to be the same as the one in the experiment), as well as the payoff received in the past game, the player's and opponent's moves in the previous game were presented at the input nodes at each cycle.

The past moves were recoded as [0,1] - for C and [1,0] - for D moves, so that activation would always come from any of the two couples of input nodes, no matter what the moves were - C or D. The values of the T, R, and P payoffs from the current game had to be reproduced as an output by the model thus implementing an in-built auto-associative memory. There were two reasons to decide to include this

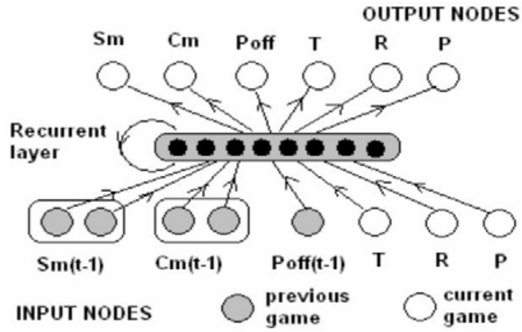


Fig. 10.1 Schematic view of the recurrent neural network and its inputs and outputs/targets. Notation: Sm and Cm are respectively the simulated subject and computer opponent (probability for) moves; Poff(t) is the model's received payoff at time t.

component in the network architecture. The first was to force the network to establish representations of the games in its hidden layer which are supposedly crucial to account for the game payoff structure in the decision making process. The second one was related to the anticipatory decision mechanism of Model A where the output nodes concerning T, R, and P were used as predictions of the possible game' payoffs in the fictitious playing mechanism explained later.

At the output, the player's move ('Sm' node) and the computer-opponent's move ('Cm' node) nodes were interpreted as the probability for cooperation for the player and the prediction about the probability of cooperation of his/her opponent in the game at hand. The payoff ('Poff') node represented the expected gain from the current game.

Training PD games with varying CI - from 0.1 to 0.9 - were presented to the neural network (T was always equal to 1, and S was always 0, R and P were distributed in this interval depending on the CI of a particular game). The games were randomized with respect to CI in the same way as in the experiments with human participants (see Hristova and Grinberg (2005b) and Hristova and Grinberg (2005a)) in order to allow for the comparison with experimental results (see Section 10.4). The network was trained using back-propagation on an input consisting of sequences of overlapping five games - the current game and the four previous games. Such sequences are further called micro-epochs.

In the very beginning of the IPDG, the length of micro-epochs was increasing with each next completed game until it reached five games. The very first inputs were as follows: the first game matrix, the player's move and the prediction of the opponent's move generated with probability 0.5. The first received payoff ('Poff') was obtained from the averaging of the payoffs of the games.

The small number of games the network dealt with at a time implies sensitivity to local changes in the game and to memory constraints we assumed to exist in real game playing. On the other hand, the micro-epochs were long enough so that specific events in the history of IPDG were encoded in the recurrent hidden layer.

The values at the six output nodes were used as predictions when the network was trained within the current micro-epoch. The 'T', 'R', and 'P' output nodes were expected to reproduce the corresponding input values in the input payoff matrices. The output of the 'Sm' node was the model-player's probability for cooperation in the current game. The output at the node 'Cm' was the prediction for the cooperation probability of the opponent, and the output at the 'Poff' node meant the expected game payoff. When both, player and opponent had made their moves and the target values for the output layer were known, the new target micro-epoch was updated and the network was trained. For all of the output nodes the training signal is supplied by the game (payoffs) with the exception of the model-player's move probability. The latter has to be supplied either from experimental data with a human player (if the model is used to fit the behavior of a real player) or by explicitly modeling the evaluation of the game outcome. Here, we will present results along the latter line.

10.3.2 Judgment and Decision Making

In order to build a realistic model able to make decisions comparable to the ones made by human subjects, we need a good evaluation mechanism for the outcomes of the player's moves. Hereafter, we present the mechanism adopted in Lalev and Grinberg (2007) which implements fictitious playing based on predictions of the recurrent network used in received payoff maximization.

The model uses the predictive properties of the recurrent network in order to "guess" how the game would proceed if its current move were either C or D. An anticipatory module was implemented in the model, so that two sequences of five games predicted by the neural network were produced before making a move. The first sequence began with a C move, and the second one with a D move. Only the first player's move was fixed in any sequence (C or D, respectively). The recurrent network had as first inputs the current game input (together with the other four games from the micro-epoch) including the values of the T, R, and P payoffs, and the players' moves and payoff from the previous game. This is a simpler mechanism than the one used in Taiji and Ikegami (1999), where all the possible strings of C and D moves are taken into account. Here the first move is chosen and everything else is based on the network predictions.

As the player's move is known in the first fictitious game (C or D), the opponent's move is generated with the probability predicted by the network. The payoff for the player is based on the moves of both players according to the rules of PD game and the payoffs T, R, P or S as obtained on the output layer.

In the second fictitious game the input micro-epoch is updated so that the new T, R, and P values are taken from the output layer of the neural network and considered as predictions about the fictitious game payoffs. The 'Poff(t-1)' node activation gets the value of the fictitious payoff from the previous game and the previous moves nodes (the 'Sm(t-1)' and 'Cm(t-1)' nodes) activations were the fictitious previous game moves. In the next iterations everything is repeated as in this but using the player move generated with its predicted probability at the output layer and not a fixed move as in the first iteration.

In this way, the work cycle is complete and the model can predict several future games and related moves and outcomes. Then the payoffs from both sequences - PoffC for initial move C and PoffD for initial move D - are considered. The obtained payoffs from the five fictitious games for each initial move choice were evaluated using a discount factor as follows:

$$Poff_{C,D} = \sum_{t=1}^{5} Poff_{C,D}(t)\beta^{t-1}, \qquad (1)$$

where PoffC,D(t) is the value of the payoff at moment t, for initial move C or D and β is the usual discount parameter that indicates to what extend the remote future game payoffs are important for making decisions at present. If β is 0, only the first fictitious payoff would matter, and if β was 1, all the five payoffs will be considered as equally important. Further the quantities PoffC and PoffD are normalized so that their sum is equal to one preserving the ratio between them. The probability for cooperation for the current move of the model is then calculated using a soft-max function:

$$P(C) = \frac{e^{Poff_C/k}}{e^{Poff_C/k} + e^{Poff_D/k}}, \qquad (2)$$

where P(C)is the calculated cooperation probability and k is a parameter for the sensitivity of the function to the difference between PoffC and PoffD. Smaller value of k correspond to larger sensitivity to the difference between the C and D alternative choices.

10.4 Game Simulations with Individual Agents: Comparison with Experimental Results

In this Section the most relevant results from Lalev and Grinberg (2007) will be presented as they are the basis of the MAS simulation presented in Section 4.

10.4.1 Comparison of the Model with Experimental Results

The agents play individually against a probabilistic Tit-for-two-Tats (Tf2T) computer strategy. Their moves depend on the player's two previous moves, thus being adaptive to their temporal cooperativeness without being easily predictable. The computer opponent probability for cooperation thus obtained is respectively: 0.5 for [C, D] and [D, C], 0.8 for [C, C], and 0.2 for [D, D]. This choice of a computer opponent is the as the one in the experiments reported in (Hristova and Grinberg, 2004) and allows for a comparison with the experimental results.

The results presented in this section are based on 30 IPDG sessions of two-hundred games against the Tf2T computer strategy. For the comparisons with the experiment the first 50 games are taken to match the number of games played by human participants (see Hristova and Grinberg (2004)). From the experiment reported in Hristova and Grinberg (2004), only data from the first part and for the control condition was used in the comparison. In this experiment (see Hristova and Grinberg (2004) for details) 30 participants played 50 PD games against the computer opponent described above. After each game the subjects got feedback about their and the computer's choice and could permanently monitor the total number of points they had won and its money equivalent. The subjects received information about the computer's payoff only for the current game and had no information about the computer's total score. This was made to prevent a possible shift of subjects' goal - from trying to maximize the number of points to trying to outperform the computer. In this way, the subjects were stimulated to pay more attention to the payoffs and their relative magnitude and thus indirectly to CI. Games of different CI, ranging from CI = 0.1 to CI = 0.9, were presented both to participants and in simulations with models A and B. Games were coming at random regarding their CI.

The best fit of the experimental results was obtained with the following parameters (see eqs. (1) and (2)): $\beta = 0.7$ and k = 0.05.

10.4.1.1 Mean Cooperation and Payoffs

The results for the mean cooperation and payoffs for the model and human participants' experimental data taken from Hristova and Grinberg (2004) are respectively presented in Figures 10.2 and 10.3. No statistically significant differences are found between the model simulation data and the experiment.

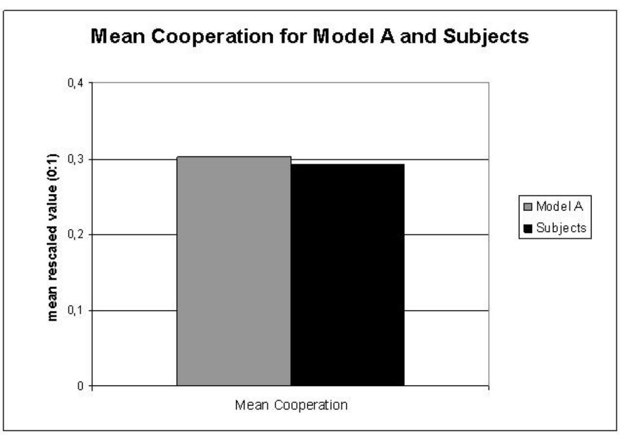


Fig. 10.2 Comparison of the mean cooperation between the model and the experimental data from (Hristova and Grinberg, 2004).

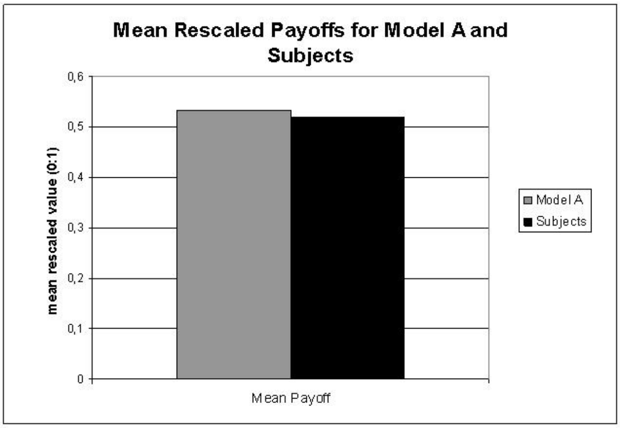


Fig. 10.3 Comparison of the mean payoffs between the model and the experimental data from (Hristova and Grinberg, 2004).

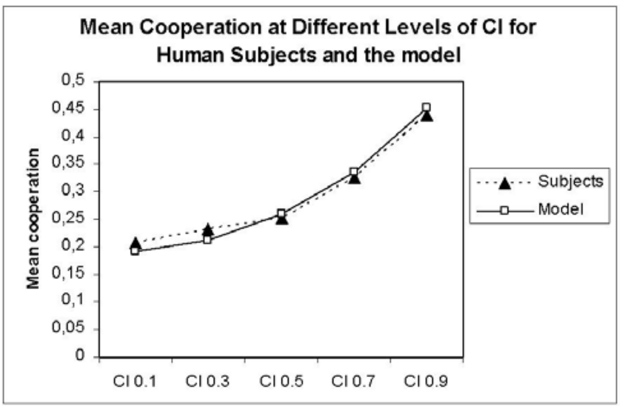


Fig. 10.4 Influence of CI on cooperation rates for the model and in the experiment from Hristova and Grinberg (2004).

10.4.1.2 Dependence of Cooperation Rate on CI

The adequacy of the model can be further seen from the comparison of the influence of CI on cooperation displayed by the model and by human subjects (see Figure 10.4; main effect observed with $F=16.908$ and $p<0.01$).

In Figure 10.4, a detailed comparison, concerning the cooperation rate dependence on CI, between the predictions of the model and the experimental results is shown. It is seen from Figure 10.4, that the model gives a good description of the experimental results with no statistical differences between the mean cooperation of subjects and the model at all CI levels, and no main effect of the type of player (model or human) on cooperation ($F = 0.386$, $p = 0.856$).

As stated earlier, our main interest is related to the CI dependence of the cooperation rate. The ability to reproduce such details in the experimental data seems very important to us in order to assess the model's validity. The simulation by the model of possible games and moves and outcomes involves the prediction about the payoff structure of the game and thus indirectly of the CI. The main effect in the CI dependence found in the simulations comes from the specific anticipatory form of evaluation of the best move involving the payoffs of the game at hand and of anticipated payoffs reflecting the structure of the current game (see Lalev and Grinberg (2007) for details).

10.4.1.3 Comparison of Game Outcomes

In Figure 10.5, the distribution of the possible types of game outcomes for the model and the subjects were compared and no significant difference was found. This statistics is very important as it shows not only the cooperation rate but gives information on the specifics of the interactions between players. Of special interest is the outcome CC in which both players cooperate.

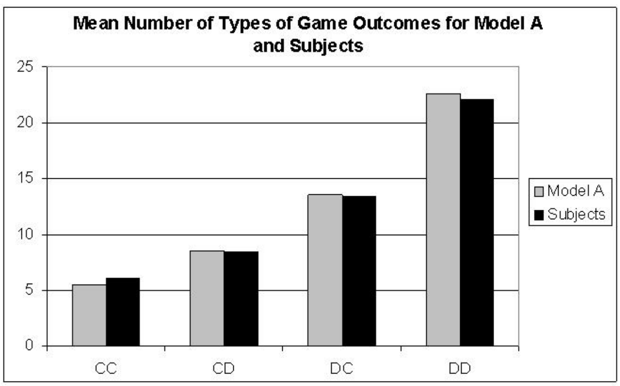


Fig. 10.5 Comparisons of types of game outcomes for the model with human subjects experiment taken from ref. (Hristova and Grinberg, 2004) (with all values for the CI).

10.5 Multi-Agent Simulations

As seen from the comparison with experiments with human participants, the model presented in the previous sections gives a good account for human playing in IPDG against a Tf2T player. In this section, we present the results from simulations of the interactions in a society of artificial players implementing such a model. The aim of the simulations was to investigate what is the role of anticipation in a society of payoff-maximizing agents on cooperation and coordination among them.

10.5.1 Agent Societies

For this purpose, groups of ten agents, with different parameterization of the model, played IPDG in simulated social environments. They played against each-other in randomly assigned couples. The length of the IPDG interaction sessions was 100 games for a pair of players. The PD game payoff matrices used in the simulations were identical to the ones used in the previous sections i.e. with CI from 0.1 to 0.9 (see the description in Section 10.4). In a society, only one pair of agents at a time played a whole game session. The pairs were chosen randomly with replacement so it was possible that one or both players from the previous IPDG session also play in the current one. There were 50 sessions in a simulation. After the end of a session, agents kept their trained network weights from their play with the opponent and these weights were kept as initial weights of the agent when it started a new IPDG with the next opponent. The sequences of last inputs and targets were also kept for each particular agent as experience from the previous session. These served as initial inputs and targets in the next IPDG sessions for the agents. When a new session began in the sequence of inputs the values of the new PD game's payoffs were used in the input vector along with the values for the last payoff, last own and opponent's moves. The overall performance of all players in the society determined its specific states and processes. When starting a new IPDG session, each player was influenced by its experiences in previous sessions with other opponents from the same society. In these simulations no mixing of agents from different societies has been done. This simulation scheme was chosen to have some common basis for comparisons with the simulations with Model A alone and with the experimental results reported earlier.

In order to investigate the role of anticipation, several parameters of the agents were varied like the number of the recurrent network's hidden units, the training method and the importance and number of fictitious games used for move evaluation (the parameter β in eq. (1). We considered five societies of agents by varying their capabilities to predict future opponent's moves and received payoffs:

1. Agents without anticipation of payoffs and opponent's move beyond the present PD game, i.e. β=0 in eq. (1) (Further referred to as Low-Anticipation society);
2. Agents implementing exactly "Model A" (30 hidden units) from Lalev and Grinberg (2007) used in the comparison with the experimental data in Section 10.4. (Further referred to as Model-A-30 society);
3. Agents with a larger number of hidden units (50 hidden units) which should increase the predictive power for the model (Further referred to as Model-A-50 society);
4. Agents with 50 hidden units and the pseudo rehearsal training method used (see Ans et al. (2002) for details). The method circumvents the neural networks' catastrophic interference problem and improves the learning and therefore the predictive capability of the model by a rehearsal procedure using pseudo training vectors. The agents trained by using this method are very sensitive to the learned in the past in IPDG sessions with other opponents which makes their behavior difficult to predict (Further referred to as Pseudo-rehearsal society);

5. Agents with 50 hidden units and strengthened anticipation predispositions: the number of fictitious games was set to 10 (twice as more as in Model A) as well as the importance of remote games was increased by setting the discount parameter from β=0.7 to β=0.9 (Further referred to as High-Anticipation society).

10.5.2 Simulation Results and Discussions

In order to compare the five societies of agents formed on the basis of their anticipation capabilities, we have concentrated on the following characteristics: cooperation rate, payoffs, type of games outcomes, and coordination in cooperation (sequences of games in which both agents cooperated).

Cooperation Rates In a simulated society, agents played ten IPDG sessions on average. With each next session the experience of players grew. In Figure 10.6, the cooperation rates for all agents in a society are averaged over their subsequent playing sessions from the first to the tenth. For example, the cooperation rates for all agents from their first IPDG session in the simulation are averaged, then for the second and so forth till the tenth.

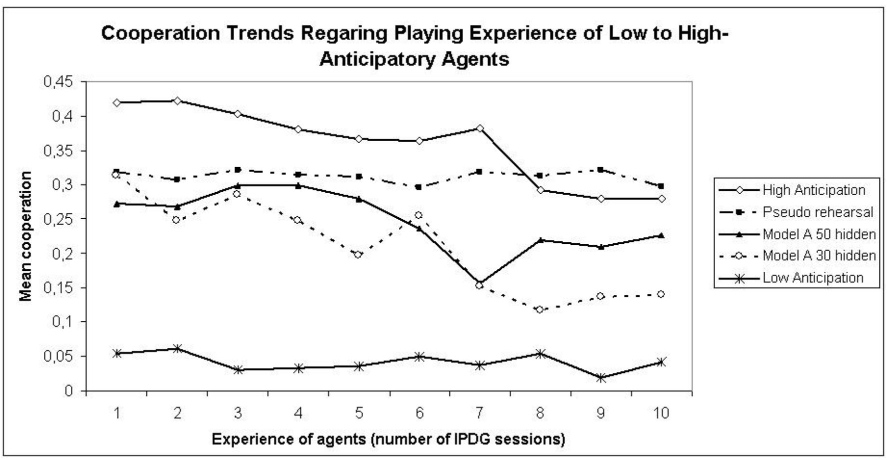


Fig. 10.6 Comparison between agent societies of the mean cooperation rates as a function of experience measured in terms of the number of IPDG sessions.

There was no significant difference between the mean cooperation of the High-Anticipation and Pseudo-rehearsal simulations (F=1.45, p=0.231) (see Figure 10.7). These two societies had the highest cooperation rates among the societies as there was a significant difference between the mean cooperation of the Pseudo-rehearsal society and the Model-A-50 society (F=18.72, p<0.01). There was also no difference in the cooperation rates between the agents from the Model-A-30 and Model-A-50 societies (F=1.93, p=0.168). Their mean cooperation rates were higher than the mean cooperation in the Low-Anticipation agent society as in the comparison between Model-A-30 and the Low-Anticipation societies F=69.95 and p<0.01 (see Figure 10.7).

Overall, the results presented in Figure 10.6 show that the anticipatory capabilities of adaptive players in social settings may be basic for sustaining reasonable levels of cooperation over time. Only in simulated societies where agents accounted to a higher extent for previous experience and used it to predict further behavior a stable level of cooperation among its players could emerge at least during the first ten IPDG sessions (as in the Pseudo-rehearsal society). In all other cases there was a tendency towards gradual decrease in the cooperation rate with time or low cooperation rate for all sessions.

In Figure 10.7 the mean cooperation in the agent societies is presented. It is seen that cooperation increases with anticipation capabilities and reaches about 0.3 for the High-Anticipation and Pseudo-rehearsal society while in the Low-Anticipation society it is below 0.05.

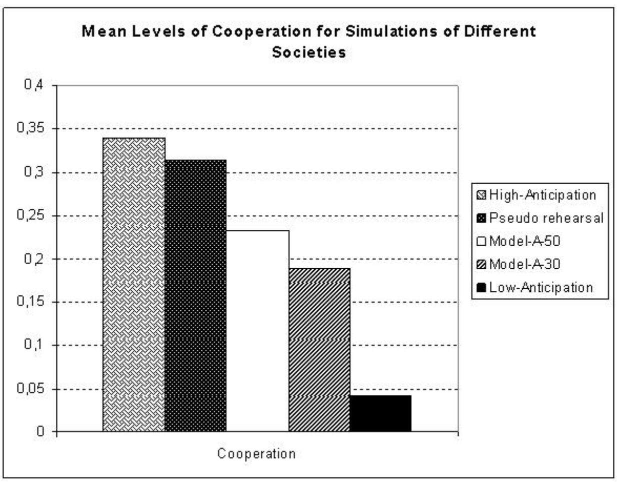


Fig. 10.7 Mean level of cooperation in simulations.

Payoffs The mean payoff received by agents is another interesting characteristic because the agents use maximal payoff-based evaluation mechanism (see Figure 10.8). The High-Anticipation and the Pseudo-rehearsal societies did not differ in the mean payoffs that were received (F=0.004, p=0.953). They got payoffs higher than the Model-A-50 society: the difference between the Pseudo-rehearsal and Model-A-50 societies was significant (F=7.82, p<0.01). The payoffs of society Model-A-30 did not significantly differ from those of society Model-A-50 (F=2.36, p=0.128). The Low-Anticipation society got the lowest payoffs as its payoffs were lower than Model-A-30 society's (F=62.21, p<0.01).

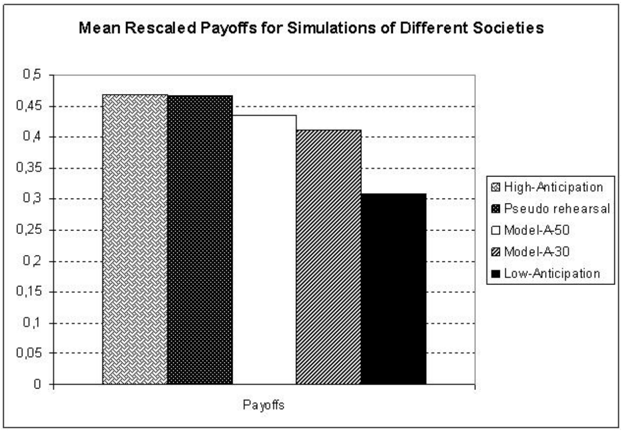


Fig. 10.8 Mean level of payoffs in simulations.

As a whole, comparison of both the analyses of cooperation and payoffs (Figures 10.7 and 10.8) reveal a rule according to which in the simulations higher cooperation rates corresponded to higher payoffs.

Again, as with mean cooperation, the High-Anticipation and the Pseudo-rehearsal societies showed the largest number of CC games and the smallest number of DD games (see Figure 10.9). The number of CC games was not different for these two simulated societies ($F=0.74$, $p=0.39$). On the other hand, the DD game outcomes were more for the Pseudo-rehearsal society than in the High-Anticipation society ($F=4.99$, $p<0.05$).

The number of CC games was significantly lower for each next society (as follows, in Model-A-50, Model-A-30, and Low-Anticipation societies), and in the Low-Anticipation society they had the smallest number (see Figure 10.9). Concerning the mutual defection (DD) game outcomes the situation is inverse. In the High-Anticipation society the smallest number DD games was observed. The largest mean number of DD games per IPDG session (more than 90 percent of the games) was reached in the Low-Anticipation simulation. For the DD game outcomes there was no difference only between Model-A-50 and Model-A-30 societies ($F=1.8$, $p=0.183$).

A tendency of increase of the mean number of CC game outcomes per simulation is observed with increase of the anticipatory propensities of agents in the societies. The opposite is valid for the mean number of DD game outcomes per simulation regarding the anticipatory propensities of agents in the societies (Figure 10.9).

For each agent society, we calculated the mean cooperation rates of agents for games with a specific CI (see Figure 10.10). It was interesting to see if the dependence on CI will be preserved in games among the agents using only a recurrent network model and playing against a Tf2T opponent as it was the case in the experiment replication (see Section 3). In all societies the monotonously increasing dependence of cooperation on the CI is clearly observed except for the Low-Anticipation society. This confirms again the role of anticipation in getting this dependence as in the experimental results with human subjects (Rapoport and Chammah, 1965).

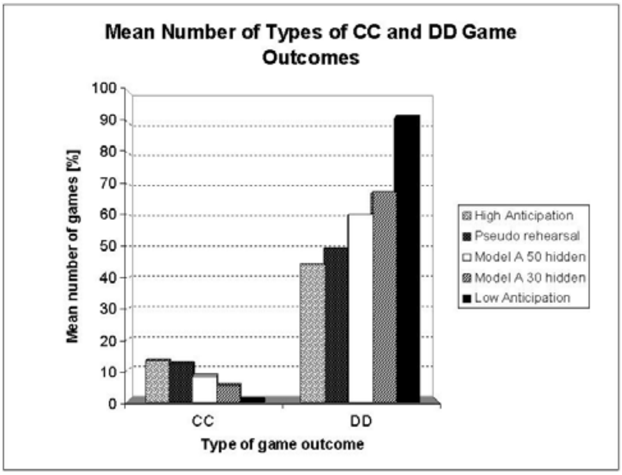


Fig. 10.9 Comparison of the mean number of CC and DD game outcomes calculated for the agent societies

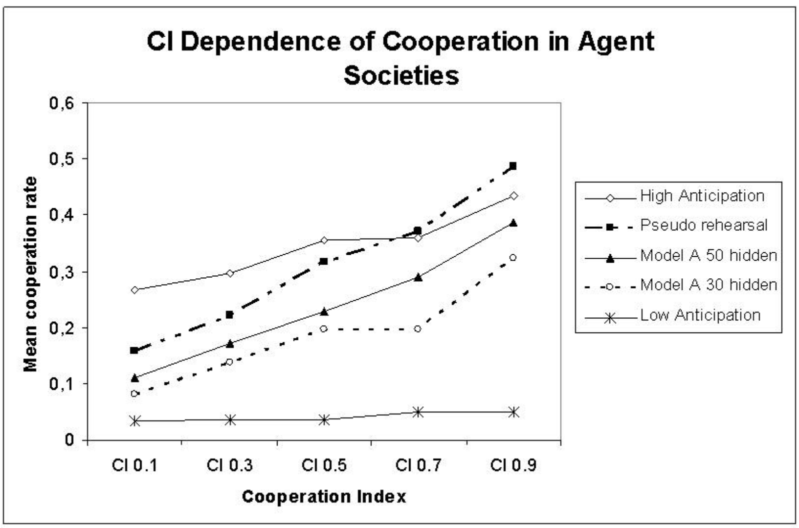


Fig. 10.10 CI dependence of the mean cooperation rate in the agent societies.

10.5.2.1 Coordination

We adopted as a first measure of the level of coordination between the agents the mean number of CC games played in a row per IPDG session. In Figure 10.11, the statistics for the agent societies are presented. The longest CC coordination lasted for five games and was present only in the High-Anticipation and Pseudo-rehearsal societies. Four-games-long sequences were observed also in the latter and in the Model-A-30 societies. In the Low-Anticipation society no sequences longer than

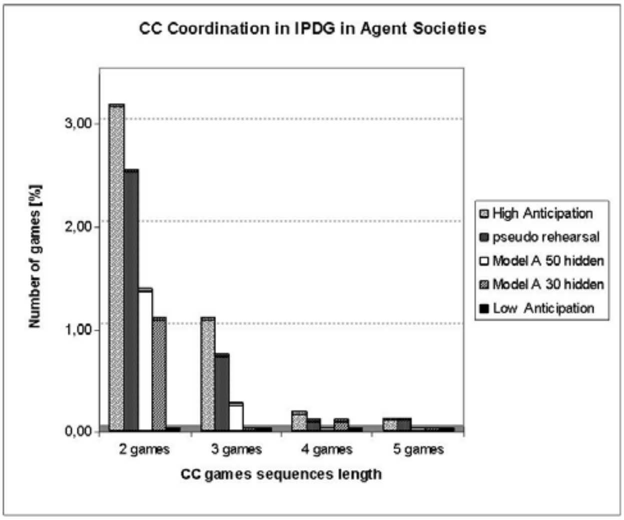


Fig. 10.11 Agents' coordination in terms of the mean length of the series of mutual cooperation (CC games) per IPDG session averaged over 50 IPDG sessions in each of the agent societies.

two were found. Although the sequences are not very long (especially compared to DD sequences some of which were 100 games long) the influence of anticipation is considerable.

This conclusion is confirmed by a related analysis we performed: the number of agents in a society that participated in a CC game sequence of given length (see Figure 10.12). It is seen from Figure 10.12 that for example only 70 percent of the agents from the Low-Anticipation society ever played a CC game whereas for all other societies this percentage equals 100. Moreover, a considerable number of agents with sequences of CC games longer than two are observed only in societies with anticipation.

The agents from the High-Anticipation and Pseudo-rehearsal societies did not differ in the mean prediction errors of their opponents' moves. But these agents were harder to predict by their opponents than this was in the Model-A-50 society. For example, there was statistical difference between the High-Anticipation society and the Model-A-50 society ($F=6.8$, $p<0.05$). The difference between Pseudo-rehearsal and Model-A-50 societies was even bigger ($F=11.64$, $p<0.01$). There was no significant difference between the prediction errors for the opponent's move in Model-A-30 society and all High-Anticipation, Pseudo-rehearsal, and Model-A-50 societies. In order that an agent be able to predict well the opponent's move, it has to possess good predictive capabilities. On the other hand, an agent with better prediction becomes more complicated and harder to predict.

These facts might account for the difficulties of more complicated players to guess the others' move intentions and, in the opposite case, simple agents to be easier to predict by the same simple agents. In the Low-Anticipation society opponents are considerably and significantly easier to predict($F=80.46$, $p<0.01$, compared to

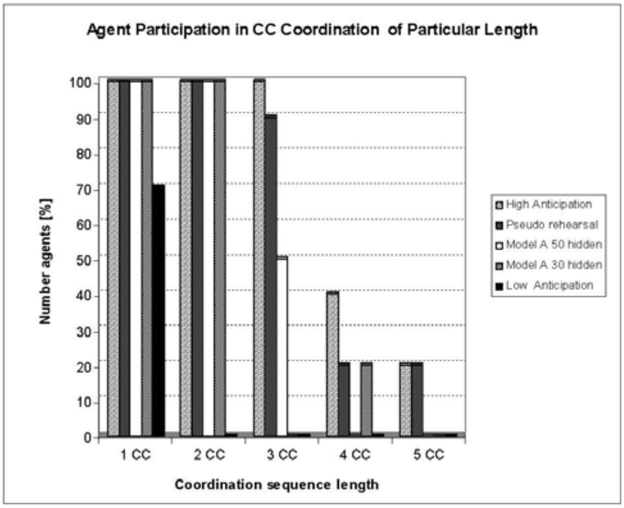


Fig. 10.12 Number of agents which played a series of CC games of a given length for each agent society.

Model-A-50 society) than in all other societies because in this society agents defect almost all of the time.

Following the difficulties to predict the opponent's move in the High-Anticipation condition, it was also more difficult for the agents to predict their payoffs in the same simulation than it was in the Pseudo-rehearsal society ($F=9.97$, $p<0.01$) and there was no such difference between the High-Anticipation and Model-A-50 societies. The prediction errors of the latter also did not differ from those of the Pseudo-rehearsal simulation. The Model-A-30 society only differed in the prediction of their payoffs from the Low-Anticipation society ($F=53.93$, $p<0.01$) and from the High-Anticipation society ($F=10.31$, $p<0.01$).

There were no differences in the auto-association of the payoff 'Temptation'. But the different types of agents were not equally capable of predicting the values of the 'Reward' and 'Punishment' payoffs. Best at predicting the R payoff were the agents from the Pseudo rehearsal simulation and those from the Low-Anticipation simulation (see Figure 10.13) whose mean errors did not significantly differ. But for the P payoff the Low-Anticipation society had better performance than the Pseudo-rehearsal society ($F=28.55$, $p<0.01$). For the R payoff there was also no difference between the High-Anticipation and Model-A-50 societies. In all other cases the mean prediction errors for R differed between societies as can be seen in Figure 10.13. The High-Anticipation and Model-A-50 agents did not differ in their ability to predict the game payoff matrix and they took medium position in this among all other agents. The best payoff matrix predictions were made by the Low-Anticipation society maybe because there was little external information to interfere with these predictions: the play in this society often converged to the DD Nash equilibrium. It was followed by the Pseudo-rehearsal society which payoff matrix prediction er-

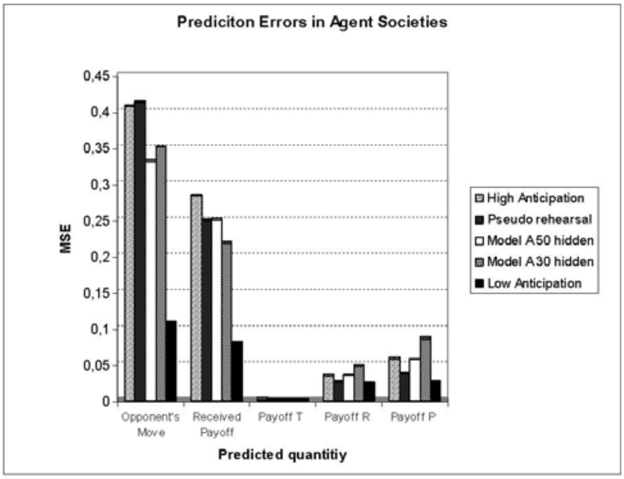


Fig. 10.13 Agents' prediction errors related to opponent move, expected payoff and game structure.

rors were lower than the remaining three simulations. A possible explanation for the good payoff predictions of this society is the enhanced training method (Ans et al., 2002) of its version of the model player. The worst in predicting the PD payoff matrix were the Model-A-30 society. Perhaps, this is due to the low memory/prediction capacity of its recurrent network architecture and various other information interfered with the auto-association of the game matrix.

10.6 Conclusion

In this chapter, multi-agent models of social interaction based on anticipation were presented and discussed. Special attention was devoted to a recurrent neural network model used to simulate IPDG playing in a society of agents. The model has been validated by comparison with human subjects experiments in a previous paper (see Lalev and Grinberg (2007)) in which participants played individually against a computer opponent. Several interesting characteristics could be reproduced which gave confidence that this model could be used in a multi-agent simulation in which the role of anticipation on cooperation and coordination could be investigated. Once the adequacy of the model was established we set up a MAS modeling a small society of agents interacting among themselves by playing the PD game. We were interested in the role of anticipation for two essential for successful social functioning characteristics - cooperation and coordination. The agents were distributed in five types of societies based on their anticipatory abilities - from agents with low predictive ability to agents with high predictive one.

The simulations showed clearly that anticipation is decisive for high level of cooperation and higher coordination. The results show that the higher the anticipatory ability is, the higher the cooperation rate and the coordination in cooperation between agents are.

In the same time, anticipatory agents opposed to each other get involved into entangled, sophisticated behavior making mind-reading difficult. A possible way out of this problem could be the existence and influence of shared norms which could lead to more coherent and 'transparent' behavior. But as human cooperation in IPDG is close in rates to the cooperation of our anticipatory agents, the prediction is that coordination series among human subjects may be in close ranges to those, observed in the simulations.

In general, there are no many investigations involving anticipatory agent societies. Further research, e.g. based on the PD game and/or other games, is needed in order to explore the full importance of anticipation for social functioning.

Chapter 11
Endowing Artificial Systems with Anticipatory Capabilities: Success Cases

Giovanni Pezzulo, Martin V. Butz, Cristiano Castelfranchi, Rino Falcone, Gianluca Baldassarre, Christian Balkenius, Alexander Förster, Maurice Grinberg, Oliver Herbort, Kiril Kiryazov, Boicho Kokinov, Birger Johansson, Emilian Lalev, Emiliano Lorini, Carlos Martinho, Maria Miceli, Dimitri Ognibene, Ana Paiva, Georgi Petkov, Michele Piunti, and Vin Thorsteinsdottir

11.1 Introduction

This book has provided various theoretical perspectives on anticipatory processes in natural and artificial cognitive systems. Advantages have been proposed and confirmed in various detailed case studies, which may have given the reader detailed insights into anticipatory processes and their importance in various cognitive systems tasks. To wrap up these advantages and give a concluding overview of various current anticipatory process advantages, this final chapter highlights a concise collection of precise success stories of anticipations in artificial cognitive systems. We survey fourteen case studies, which were developed during the EU project **MindRACES**[1]. In these studies, simulated or real robots were tested in different environmental tasks, which required advanced sensorimotor and cognitive abilities. These abilities included the initiation and control of goal-directed actions, the orientation of attention, finding and reaching goal locations, and performing mental experiments for action selection. All the studies have shown advantages of anticipatory mechanisms compared to reactive mechanisms in terms of increased robot autonomy and adaptivity. In some cases, anticipations even caused the development of new cognitive abilities, which were simply impossible without anticipatory mechanisms. For each case study, we indicate relevant associated publications, in which the interested reader may find further details on the relevant computational architectures, the involved anticipatory mechanisms, as well as on the analytical and quantitative results.

While the book as a whole has laid out the theoretical principles and design methodology for such advancements, this final chapter thus provides various possible starting points for further developments in both the surveyed system architectures and the presented solutions to the cognitive tasks addressed.

[1] MindRACES: From Reactive to Anticipatory Cognitive Embodied Systems; funded under grant FP6-511931 under the "Cognitive Systems" initiative from the EC.

G. Pezzulo et al. (Eds.): The Challenge of Anticipation, LNAI 5225, pp. 237–254, 2008.
© Springer-Verlag Berlin Heidelberg 2008

11.2 Flexible Goal-Directed Arm Control: The SURE_REACH Architecture

SURE_REACH (a loose acronym for *Sensorimotor, Unsupervised, REdundancy-REsolving control ArCHchitecture*) is a hierarchically structured, self-supervised learning control architecture (Butz et al., 2007a; Herbort and Butz, 2007). Initially, SURE_REACH explores the interactions of its associated body with the environment by means of random motor babbling. During these explorations, it forms associative models of motor-dependent sensory correlations, which is referred to as a sensorimotor model, and of bodily inverse kinematics. The developing knowledge of SURE_REACH about its body and environment is represented by population-encoded spatial body representations, and associative structures that correlate the body representations with each other.

SURE_REACH has been applied to the control of a three-degree of freedom arm in a two-dimensional environment. Due to this setup, each potential target position can be reached with various arm goal postures and on various paths from the current posture to these postures. Given a goal location, SURE_REACH consequently determines not only the most suitable target posture but also the currently optimal path to that posture online. Due to the population encoding, the associative correlations, and the goal-directed, anticipatory behavior approach, it has been shown that SURE_REACH is able to represent multiple problem solutions implicitly in parallel and thus is able to adjust its behavioral policy highly flexibly to current task constraints and changing environmental circumstances.

In the original setup (Butz et al., 2007a; Herbort and Butz, 2007), SURE_REACH represents an *extrinsic hand space*, which encodes hand locations (x-y coordinates) with a uniformly distributed, partially overlapping 2-D array of neurons. Additionally, an *intrinsic posture space* similarly encodes arm postures with an uniformly distributed, partially overlapping 3-D array of neurons (shoulder, elbow, and wrist angles). SURE_REACH learns to correlate these two spaces with each other in a *posture memory*, which associates hand with posture space neurons—effectively encoding an inverse kinematics model—and a *sensorimotor bodyspace model*, which associates postures with each other action dependently. The resulting system is able to predict which posture is reached given a current posture and chosen motor command. More importantly, though, it is also able to deduce the posture that *preceded* a given posture given some action was executed. This latter capability enables the system to choose motor commands when given a current posture and a goal posture. If the goal posture(s) are not in the immediate vicinity, then dynamic programming can be used to generate potential fields in the representation that are able to guide the arm to the goal posture by means of closed-loop control.

The representation together with the goal-directed behavior control processes enable highly flexible and adaptive anticipatory behavior control. Essentially, due to the anticipatory, redundant encoding of behavior alternatives, the system is able to **initiate goal-directed actions highly effectively and context-dependently** (Butz et al., 2007a). Essentially, the anticipatory architecture learns to reach goal postures as well as hand goal locations highly reliably. Moreover, the system can flexibly adjust its behavior to various task constraints: it can avoid obstacles that block the shortest path to the goal; it can compensate for broken joints and prefer particular joint movements over others; it can combine multiple goal constraints—such as a hand goal location with a particular position of a joint (as long as the hand goal is still reachable). These capabilities were achieved by simple multiplicative influences on the architecture.

Moreover, the system can also account for future goal priorities while executing reaches to current goal locations, essentially preparing for a **faster and smoother reach** to a subsequent goal (Herbort and Butz, 2007). Additionally, in collaboration with other MindRACES partners, the architecture was coupled with goal state selection mechanisms. These selection mechanisms were based on reinforcement learning and particular the actor-critic method. In the examples investigated, the goal selection mechanisms chose to reach goals inside a goal region that were the further away from a punishment region the closer the punishment region and the stronger the punishment. Essentially, these experiments showed **improved decision making** due to the anticipatory representations and the implicitly anticipatory goal location choice, which emerged from the reinforcement learning architecture. Besides the general effectiveness of the action choice, the data also closely reproduced data from psychological experiments (Herbort et al., 2007).

In conclusion, SURE_REACH is a control architecture that effectively stores alternative (redundant) behavior means and goals. When a particular goal is desired, it effectively constrains the representation and issues control commands that yield the behavior that is maximally suitable given the current goal and (optionally) further current constraints. The achieved behavioral flexibility is only possible due to the effectively encoded sensorimotor model and the anticipatory, goal-dependent control structure.

11.3 Learning Cognitive Maps for Anticipatory Control: Time Growing Neural Gas

The time-growing neural gas (TGNG) approach builds spatial representations linking sensory codes time- and motor-dependently (Butz et al., 2008b). Essentially, TGNG learns sensorimotor spaces from scratch via random motor babbling. It forms a neural network representation of the experienced space and the connectivity within that space. Distances in the space are represented by the motor activity necessary to reach one spatial location from the other. Sensory proximity between close spatial representations is not required. The exploration essentially leads to the generation of a 'cognitive' map of an environment, which is represented by the growing neural network. The representation is related to place-cell and head-direction cell encodings found in the hippocampus in rats (Wiener et al., 2002). Each node in the growing network essentially encodes a place cell, which is activated by appropriate sensory input. Each edge, which connects two place cell nodes, associates the average motor command that was executed to move from one node to the next one. TGNG is complementary to SURE_REACH: it may replace the currently uniformly distributed bodyspace encodings in SURE_REACH by the flexibly developing spatial encodings.

TGNG was used to control a robot vehicle in a maze environment, which the robot initially explores by random movements. The developing map enables the invocation of goal-directed control commands that move the vehicle to currently activated goal locations. This behavior is initiated by the activation of the network nodes that correspond to the goal locations, the subsequent propagation of that activity via dynamic programming principles, and the final closed-loop movement to the goal. Behavior is controlled by activating those motor commands that are associated with the edge that leads to the next higher activated node from the node best representing the current robot location. The spatial anticipatory encoding showed to yield **highly effective and context-based action initiation** and **fast and smooth behavior execution**. Particularly, the system reached goal locations reliably and effectively and it was able to flexibly adapt its behavior given additional task constraints, such as preferred movement directions (Butz et al., 2008b).

In sum, TGNG showed that cognitive maps can be learned by simply associating growing perceptual nodes motor-dependently. As long as the problem is Markov—so that each perceptual node maps to a unique state in the environment—the resulting representation is very suitable to generate flexible, anticipatory, goal directed control commands. The advantage of connecting states in time is that states can be connected that may strongly differ in how they are perceived as long as they are easily reached from each other. As a consequence, the cognitive map is rather independent of sensory proximity. It represents proximity directly motor dependently, that is, which motor effort is necessary to reach a location in space from another one. The consequence is that direct anticipatory, highly flexible behavioral control is possible.

11.4 Learning Effective Directional Arm Control: The Evolutionary System XCSF

The XCSF classifier system is well-known in the evolutionary computation community for its robust capability of iteratively learning function approximations accurately and reliably (Wilson, 2002). The system has been applied to approximation tasks of up-to seven dimensions and it has shown to be machine-learning competitive in several respects (Butz et al., 2008a). In behavioral tasks, the system has also been applied to the problem of learning generalized Q-value functions in real-valued domains (Lanzi et al., 2006).

Recently, during the MindRACES project, XCSF was further developed to yield directional, anticipatory control structures. In a first application of the resulting architecture, a directional control structure was learned for an arm with three degrees of freedom in a two-dimensional environment. The architecture mapped the arm posture space by evolving a population of overlapping, piece-wise (linear) classifier approximations. Each classifier encoded the control commands necessary for a directional movement in a particular arm posture subspace of the environment.

During learning, XCSF partitioned the posture space of the simulated arm to accurately predict how motor actions affect hand movements. The inversion of the predictions enabled goal-directed, closed-loop control of reaching movements. The system reached remote hand locations accurately, reliably, and effectively. Moreover, it was shown that the learning approach did not rely on the particular sensory inputs nor on a linear mapping. In fact, the evolving control map in XCSF is inherently non-linear. Due to the predictive learning approach and the consequent, inverse, piece-wise linear control approach, the system yields **fast and smooth behavior execution** patterns (Butz and Herbort, 2008).

In sum, an evolutionary approach was used to learn a forward-inverse piece-wise linear mapping of an arm control space, which was represented by a population of neural classifiers. Only the employed forward-inverse anticipatory representations of the classifiers coupled with goal-based, directional closed-loop control enabled the effective invocation of fast and smooth behavioral control.

11.5 Anticipatory Target Motion Prediction

To solve the task of predicting the movement of a visual target, a learning linear associator with memory, embedded within a Kalman filter was developed. While the Kalman filter takes care of the prediction of the target location the linear associator learns the model of the target motion (Balkenius and Johansson, 2007; Balkenius and Gardenfors, 2008). The memory component stores previous observations and allows the associator to train on a large number of observation in each iteration. This technique effectively emulated some of the advantages of batch-training methods within an on-line learning system.

The learning system has been applied to a number of task in which a tracking component is necessary, including the tracking of moving balls in partially occluded situations and the modeling of pursuit eye-movements. By combining the learning predictor with an inverse model of a robotic arm with three degrees of freedom, it became possible to catch a plastic toy fish that moved along a regular circular path. The system learned to predict the location of the target using color tracking in combination with the associator described above. The implemented system had a delay of approximately 500 ms from camera image to motor control. The prediction mechanism was essential for **successful tracking** as well as for the **manipulation of moving objects**. It also allowed for **faster and smoother behavior execution** since actions can be directed toward the future location of the target.

Prediction is an important ability that is useful as a component in many different applications. The results showed that motion prediction can be effectively included in many different tasks. Furthermore, it was investigated how different learning systems can be adapted for prediction by delaying inputs, outputs, and training data in different ways. A general conclusion is that any learning system can be adapted for prediction tasks in the way outlined and that predictive learning can be vary fast in simple situations.

11.6 Anticipatory Spatial Attention with Saliency Maps

A new approach of learning saliency maps was formulated that allows standard reinforcement learning techniques to be used in a number of attention tasks (see chapter 4). The approach is based on a novel and compact formulation of a saliency map, which allows many types of visual information to be combined in a coherent way. In a current implementation, feature-based and spatial attention was combined in a seamless way. The central idea is that the saliency map can be seen as an approximation of a value function for reinforcement learning. Unlike the standard action-value function in reinforcement learning, there is no state in this formulation. Instead, each location in the image corresponds to an individual action that directs attention to that location. Since all different sources of attentional signals eventually lead to attention that is spatially focused, the approach provides a common language for all such processes.

The mechanism has been applied to selective attention as well as priming in sensory processing. The mechanism is general enough to be used in any system that includes any form of sensory selection. In particular, the mechanism was used to select targets for visual tracking. It can also be used to **improve top-down attention** by tuning it to external reinforcement. Moreover, it **improves information seeking** by allocating larger processing resources to input data that resembles previously rewarded stimuli.

The new mechanism shows how it is possible to add reinforcement learning also to systems that are not use to control actions directly and suggests a general strategy for the marriage between reinforcement learning and perceptual processing (Balkenius and Winberg, 2008). Moreover, the proposed mechanism can be used as an important part of a complete reinforcement learning architecture to select stimuli as well as actions. To the best of our knowledge, this is the first computationally efficient implementation of a mechanism first suggested in (Balkenius, 2000).

11.7 Behavior Prediction in a Group of Robots

A combination of several techniques was used to anticipate the future behavior of a group of robots (Johansson and Balkenius, 2007). Kalman filters were used for short term prediction and correction of tracking data. Associative anticipatory attention mechanisms were used to learn where robot will reappear after they disappeared behind obstacles and to produce epistemic actions in the form of directed attention to gain optimal information about the behavior of other robots. The system also used internal simulation based on internal models of the other robots to anticipate how they will behave, in order to select appropriate actions in relation to the other robots. The robots also used a form of primitive joint attention through communication about their observations.

The combined mechanisms were used to investigate how a robot can control its own behavior depending on the anticipated behavior of another robot. A hiding scenario was implemented using a multi-robot set-up. There were four robot thieves and two guards that patrolled the environment in a regular fashion. The task of the thieves was to hide from the robot guards while navigating to certain places in the environment. The implemented system showed **improved decision making** by simulating the other robots' behavior. **Faster and smoother behavior execution** was achieved by employing anticipatory mechanisms at three levels: at a low level of motor control, at an intermediate level for avoiding collision with other robots, and at a higher level for reaching desired goals without interference with other robots.

The implementation shows how anticipation at a number of levels can be made to work together within a unified agent architecture and addresses many of the difficulties that arise as complex anticipatory systems are built. It is believed that this is the first architecture that combines this many aspects of anticipation in a coherent system and successfully controls the individual anticipatory behavior of a robot as well as the emergent interactions between a group of robots with similar or conflicting goals.

11.8 Enhanced Adaptivity in a Predator-Prey Scenario

Another study investigated how anticipations can enhance the adaptivity in a predator-prey scenario, in which a predator (usually the dog like Sony AIBO™robot) is supposed to catch a prey (such as a simpler robot with a reactive behavior, or in more sophisticated scenarios, a second AIBO™). First of all, the magnitude of the benefits by implementing anticipatory behavior for the predator depends on its physical abilities like velocity and agility, but seen with respect to the analogue capabilities of the prey. If, for instance, a predator can navigate much faster than the prey, then the simple behavior of heading towards the prey and then approaching it will usually be a sufficient strategy once the prey has been detected. This strategy can still be filed under reactive behavior, as the necessary anticipatory capabilities of learning the effects of necessary movements can be achieved quite easily.

A certain form of anticipation is necessary, if prey and predator possess comparable navigation abilities and both operate in the open field. The predator has obviously some advantages if it is able to predict the trajectory of the other robot. If additionally obstacles are present, which might occlude the opponent, then anticipatory capabilities deliver substantial advantages and it is easy to construct scenarios in which pure reactive behavior rarely succeeds, for example, if the prey is only visible for a short amount of time, which is not sufficient for a successful access.

Two approaches have been developed, one is based on Markov models (Lewandowski, 2007) and the second uses artificial immune systems (De Castro and Timmis, 2003). Both approaches operate in the space of observed sensor values and do not try to estimate robot locations within a world map. For the Markov approach, the current camera image is transformed into one of a final number of possible views. The elementary building blocks are the estimated probabilities of transitions between the current and the following state, given the current action, which are updated continuously while maneuvering. The ability to construct 1-step predictions could be used now to form chains of actions and to predict possible future outcomes, or to be more exact, to predict probability distributions over possible states for a given time point. The system knows in its representation every so far encountered state.

Actually, the algorithm implements goal-oriented behavior in a different way. A desired state is one already experienced view, for which the prey has been seen from the nearest so far observed distance. Multiple desired states are allowed to exist. A plausibility check for a necessary condition ensures that the desired goal states are probably achievable from the current state with the acquired knowledge of transitions. A backward induction algorithm (Puterman, 1994) is then applied in a recursive manner that finally a current action can be chosen that, based on the history of transitions, one of the goal states will be reached as fast as possible. The algorithm allows that the prey is temporarily hidden, and therefore it is not mandatory that each step reduces the distance to the prey.

The building blocks and systemic operations of the AIS approach are elements and procedures that are in line with the immune system metaphor. Simple atomic elements that code for a condition, an action and an expectation, in terms of the possible outcome of an action, correspond to the antibody that reacts to a certain degree matching environmental stimulus, the antigen. Sensor inputs and therefore the epitopes of the antigens are coded as strings and are compared to the condition parts of the artificial immune systems elements. The best corresponding element is chosen and its action is executed. The expected outcome of this single action can be predicted and this in consequence allows the anticipation of the outcome of future action sequences in relation to the current situation and therefore the prediction of the outcome after multiple time steps.

After each successful run, the population of elements is dynamically updated according to certain interdependency patterns that closely follow Niels Jerne's immune network theory (Jerne, 1974), with the consequence that the concentration of elements belonging to successful runs will grow higher and their elements will be chosen more likely to produce genetically varying offspring than less effective antibodies, which are subsequentially suppressed. Within AIS, the goal to catch the prey is therefore not explicitly given, but is implicitly coded in the distribution and concentration of the elements of the network's population. A main goal of the experiments with artificial immune systems was the evaluation of their capabilities in controlling a successful predator (and in some cases prey) in match to other control strategies. Among these were genetically evolved static strategies like differently complex subsumption architectures, as well as other evolutionary computation algorithms, including classifier systems (XCS) and even anticipatory classifier systems (Sigaud and Wilson, 2007). In all scenarios, from the simplest without any obstacles, to the most complex, where both players are controlled by anticipatory control mechanisms and a hiding place was introduced, the artificial immune system approach was able to solve the task of catching the prey (or escaping) very well. Additionally the behavior exhibited to the human observer in many cases appeared like a reasonable anticipatory strategy that one would observe also in biological hunter and prey scenarios.

In conclusion both approaches can be successfully applied, if the behavior of the prey under similar circumstances remains stable. When the presence of obstacles is allowed, they provide **substantial advantages over purely reactive models**. The behavior of the prey does not need to be deterministic, but it should not change abruptly. Thus, the algorithms in their current form are best applicable, if the prey's behavior is stationary.

11.9 Adaptive Navigation and Control with Anticipation

In a real robot task an omnidirectional robot learns to correct inaccuracies while driving, or even learns to use corrective motor commands when a motor fails, both partially or completely, to optimize the driving accuracy for soccer competitions (Rojas and Frster, 2006). The robot anticipates how its actions influence the environment. It uses this knowledge to choose the best action fulfilling its intention in the future. The robot also observes itself to detect drifting effects of its actions and adapts its own world-model accordingly.

A feed forward neural network with historic and current information of the robot's poses is used for learning the robot's response to the commands. The learned model can be used to predict deviations from the desired path and takes corrective actions in advance, thus improving the driving accuracy of the robot. The model can also be used to monitor the robot and assess if it is performing according to its learned response function. It was demonstrated that even if a robot loses some motor's power, the system can **relearn to drive the robot** in a straight path, even if the robot is a black box and we are not aware of how the commands are applied internally. The robot controller framework integrates action control by choosing actions based on its own self-model. It integrates attention and monitoring by observing its own hardware quality to change its behavior and monitoring. Without these three anticipatory mechanisms, the robot would still fulfill its task to some extent, but not sufficiently (in the sense of robustness and accuracy) to win an international multi robot contest.

11.10 Mental Experiments for Selecting Actions

A robotic task consisting in finding and moving to a randomly placed unique colored cup in the room illustrates that anticipatory mechanisms are essential to fulfill the assignment (Bakker et al., 2006; Zhumatiy et al., 2006). The key idea to solve the task is to use a forward model to translate sensor inputs to robot movement commands. This module anticipates a probabilistic world-model to estimate future rewards by *mental experiments*. The results of the mental experiments are then used to select the most promising actions for the motor controller.

The robot is equipped with a color camera and placed in a room, which contains the colored cup. The camera is mounted in front of the robot and looks a bit downwards. It has a very limited field of view in relation to the room. Therefore, the robot has to find the cup before it can move to the target position.

The controller of the robot translates sensor input data to robot movement commands. It is trained by various reinforcement learning methods. In (Zhumatiy et al., 2006) the mean position of all camera pixels in a specific color range of the target object is used as input for the reinforcement learner. To reduce the huge amount of memory for the policy, a Piecewise Continuous Nearest-Sequence Memory (PC-NSM) algorithm is used for general metrics over state-action trajectories. In (Bakker et al., 2006), the visual information from the camera is preprocessed into a 5x4 binary grid, which represents the position of the cup in the camera image, if the cup is visible. To reduce the long training time for reinforcement learning algorithms for real robots, a probabilistic anticipatory world-model is learned from comparatively few real robot experiments. This world-model is then used to conduct mental experiments to train the controller with Prioritized Sweeping, a modification of the standard Q-Learning algorithm. The policy is applied with a high repetition rate during the learning process of the mental model and with a real time repetition rate in the physical world.

The robot controller framework combines action control, active vision, attention, goal directed behavior, and monitoring. Removing one of these anticipatory modules makes it impossible to learn the whole task on a real robot within a reasonable time frame.

11.11 Anticipations for Believable Behavior

The emotivector (Martinho and Paiva, 2005) is an affective anticipatory mechanism situated at the agent sensory interface. Each emotivector is coupled with a sensor and generates affective signals resulting from the mismatch between sensed and predicted values. Inspired by the psychology of emotion and attention, our implementation of the mechanism showed how attention grabbing potential as well as elementary sensations can be automatically generated from the observation of the values flowing through the agent sensors, and be used to generate believable behavior for the agent.

Two tasks were performed to evaluate the effectiveness of the emotivector mechanism: one taking place in the virtual world, where Aini, a synthetic flower, helped users to perform a word-puzzle task; and another, taking place in the real world, where the iCat social robot played a game of chess against the user. In both tasks, the affective signals generated by the emotivector were used to directly control the affective expression of the synthetic character interacting with the human user. The results showed that behavior autonomously generated by the emotivector is perceived as believable and understandable by the users.

Anticipation (and the associated uncertainty) plays an important part in the generation of affective signals. In the emotivector mechanism, anticipatory mechanisms are used for: deciding whether a percept is salient; defining the quality of the elicited affective states; allowing the mechanism to be context-free, without any requirements related to the manual tuning of parameters. As such, besides being a crucial factor in the autonomous generation of attentive and affective signals, anticipations allow parsimonious design.

The research issue resulting from the gap between the scope of psychological theory and the engineering needs of the believable character community has only started to be addressed in a principled manner. The emotivector mechanism is located in this area of relevance and addresses the specific question of creating autonomous believable behavior to support the engineering of believable synthetic characters. The emotivector achieves this goal by fusing the fields of anticipatory computing and affective computing. By design, this approach is broadly applicable and provides practical means to significantly improve the capabilities of believable characters driven by such architectures.

11.12 Anticipatory Behavior in a Searching-for-an-Object Task

In section 9, the use of analogy making as a prediction mechanism was investigated. Analogy allows prediction making about the current situation based on one episode in LTM, which could be from a different domain. The analogy involved can be very superficial or very deep. The AMBR model of analogy making (Kokinov, 1994b; Kokinov and Petrov, 2000, 2001) was further developed and augmented with a transfer and evaluation mechanism, which allowed the implementation of a real robot scenario involving perception and action execution. These mechanisms allowed the usage of analogy as a selective attention mechanism and top-down perception mechanism, which directed attention of the robot to anticipated objects or properties. The scene representation was build gradually including only relevant objects and relations. The use of analogy as a basis for anticipatory mechanism is novel and proposed for the first time.

The task in which the model was tested was a ‘searching for an object task’. The AIBO robot had to find its hidden bone under an object in a room. The objects differed in shape and color. The robot analyzes the scene, makes a decision where the object could be, and goes to find it.

In such a task, using a non-anticipatory approach will lead to full search for the objects below all shapes. Although it cannot be guaranteed that the episode retrieved by analogy will lead to the correct solution, in many cases it does and the way the solution is found is unique and was only possible due to analogy making. The advantage of the analogy based prediction compared to other prediction methods is the ability to use just one positive trial in order to generate the prediction. The proposed methods for top-down perception and selective attention based on anticipation deal with the hard problem of processing visual input. The huge space of objects and relations is filtered, which allows the mind to handle only small but relevant aspects of the available information.

The analogy based anticipatory mechanisms seem very promising for finding solutions in some situations. They have to be further tested in richer environments in order to explorer their full potential and scalability. A promising further development seems to be the use of analogy based predictions as a basis for models of perception and action and work along this line is in progress.

11.13 The Role of Anticipation in Cooperation and Coordination

In section 10, the problem of cooperation as a basic principle underlying was investigated in the framework of Iterated Prisoner's Dilemma Game (IPDG). Results from simulations with a connectionist architecture of an anticipatory PD player allow the conclusion that anticipation is of high importance for cooperation to be present in 2x2 interactions as well as in simulated PD play in a society of agents. The architecture combines a simple recurrent neural network with an auto-associator and a forward looking evaluation mechanism (Lalev and Grinberg, 2007), implementing an anticipation-based decision making mechanism. In IPDG, the recurrent network processes the flow of available information: the structure of the PD matrix, the players' moves, and the payoffs obtained from the game. Due to the learning mechanism, the network correlates this information in time, whenever appropriate (for example, how players' moves correspond to gains from the game), and tries to infer information that is not available yet—such as the move of the opponent. In the case of a single couple of IPDG players, the model managed to reproduce results from experiments with human subjects by (Hristova and Grinberg, 2004). Anticipation, rather than backward-looking reactive behavior, was responsible for the cooperation of the model against the simple-strategy computer opponent also used in the experiment with human subjects. Manipulation of anticipation forward-looking parameters also revealed that the anticipatory properties of the model's decision making contribute most to the observed dependence of cooperation on the structure of the payoff matrix (the so-called cooperation index).

With instances of the validated model architecture, simulations of IPDG playing in small societies were conducted. The aim of the simulations was to investigate the role of anticipation in a society of payoff-maximizing agents on cooperation and coordination among them. For this purpose, small groups of agents with different parameterization of the model played IPDG in simulated social environments. The parameters were chosen to have five groups of players with increasing anticipatory capabilities. The analysis of the processes in each society was based on the overall level of cooperation, mean payoffs, as well as cooperative coordination. It turned out that the level of cooperation in the simulated IPDG societies grew with anticipation, starting from 5% in the first society and reaching up to 30% in the fifth society. Corresponding to their anticipation, the intermediate types reached intermediate levels of cooperative interactions.

A tendency of increased mean number of mutual cooperation cases per simulation was observed with increased anticipatory properties of the agents in the societies. The opposite was valid for the mean number of double non-cooperative choices (mutual defection) per simulation, as this number increased with diminishing anticipations in the societies. Higher mutual cooperation is considered an advantage as long as this is the most profitable outcome for the society in the long run. The summary payoffs, which were gained in each society, were also positively correlated with forward-looking capabilities: the higher the anticipation within a society, the higher the payoffs obtained by the members of the corresponding society.

As a first measure for the level of coordination between the agents, the mean number of mutual cooperation games played in a series per IPDG session was used. The longest mutual cooperation coordination was observed in the societies with highest anticipation. Although the sequences are not very long, the influence of anticipation is considerable.

Cooperation and coordination play a positive role in a society and represent a decisive advantage. In IPDG, for example, bilateral coordinated cooperation would result in higher gains for both players. On the level of a society, cooperation and coordinated actions will lead to high overall productivity and benefits. These simulations showed that anticipation is decisive for high levels of cooperation and higher coordination. According to the results, the higher the anticipatory capabilities are, the higher the cooperation rate and the coordination in cooperation between agents. As human cooperation in IPDG is close in rates to the cooperation of our anticipatory agents, the prediction is that coordination series among human subjects may be in close ranges to those observed in the simulations.

11.14 Anticipatory Effects of Expectations and Emotions

Recent computational models in the context of cognitive systems are providing simple affective states in terms of their functional effects on agent's behavior. Their roles are argued to enable adaptive and situated cognition and span from reactive methods of control (similar to those employed in primitive biological organisms) to the control of computational resources, attention, and decision making processes. Systems based on appraisal theory stressed different relations between emotions and cognition, arguing emotions as a causal precursor for the mechanisms to detect, classify, and adaptively cope with significant events and environmental changes. Typically, emotions are modeled as cognitive mechanisms to monitor goal pursuit in terms of functional appraisal of action achievement and failure. Besides, emotion signals may rule intelligent resource allocation, improve situated cognition, and generate goals, which are translated into purposive behavior.

Moreover, an approach was investigated that promotes the anticipatory effect of emotions as a main breakthrough. In so doing, we envisage to address either the route from emotion to anticipation, or the reverse one, from anticipation to emotion.

Emotions and Anticipation in Goal Directed Agents Our theoretical model of emotions have been developed by using mathematical and logical tools which have been developed in the field of decision theory and applied modal logic (Castelfranchi, 2005; Lorini and Castelfranchi, 2007). This level of specification has been a first foundational step towards the design of computational architectures for affective and anticipatory agents.

In realizing a computational model for emotions and anticipations, a multifaceted approach was adopted by distinguishing different processes behind goal oriented behavior. In so doing, practical reasoning, epistemic reasoning, and situated reasoning were distinguished and basic principles at the basis of emotions in terms of their cognitive ingredients were investigated. This promoted a clear disambiguation between slow, decisional processes, processed devoted to deal with knowledge and processes related to cope with situated events.

Whereas typical approaches in modeling cognitive agents are oriented at including graded primitives and temporal dimensions (i.e. belief on the future), a cognitive approach was adopted, which introduced expectations as emerging attitudes coming from epistemic and motivational states. A particular use of expectations enhancing problem solving and learning abilities has been modeled in the deliberation and goal selection phases. In addition, by pointing out the subjective character and the functional role of expectations, as intrinsic cognitive ingredients of many basic emotions (i.e. surprise, hope, relief, disappointment), a further kind of interaction between emotion and anticipatory activities was considered that consisted not only in predicting future events, but also in anticipating future emotions.

In (Piunti et al., 2007c,b,d), the quantitative influence of expectations upon the terms of a rational decision was investigated. In so doing, expectation driven decision making was introduced(Piunti et al., 2007c), which enabled agents to proactively take decisions either on the basis of anticipated events (i.e. trends of monitored signals) or on the basis of ongoing needs and desires.

As in appraisal-inspired models, emotions and mental states were provided to coordinate different computational and physical components required to effectively interact in complex environment. A clear methodological separation of concerns allowing the modeler by breaking down the work into two separate and independent activities was promoted: while the former was defined referring to the goal overview in the problem domain and clearly involves decisional processes (i.e. deliberation of alternative courses of actions), the latter can be defined through control frames to improve situated behavior. The following step was made to reinstate the two approaches by taking into account the correlations and the relative interactions occurring in system execution model. This allowed to integrate the low-level, situated reasoning to be used to inform higher decisional processes.

The emergent nature of affective states enables agent to adopt a mental frame while both expectations and emotions are conveyed to inform reasoning for redirecting resources and adopt long term strategies once a disturbing event is detected. To this end the contribute of Mental States is twofold: from the one side they can relieve the deliberative and the attentive processes from the burdens to process weakly relevant information in decision processes, excluding action alternatives that are likely to be less promising or have vanishing likelihood to be achieved. On the other side, Mental States provide ready to use action selection and resource allocation strategies that may relieve agent's need for resource-demanding and meta decision processes. An additional effect of modeling emotions through mental states is for agent's intention reconsideration. Traditional reconsideration strategies indicate an agent to abandon an intention when a related goal is achieved, when a goal become infeasible or when the agent relieve some inconsistencies between the world state and the external conditions necessary for goal achievement. Our model allows basic emotions to elicit an *interruption* on normal cognitive processes when unexpected events require servicing. Once based on expectations of future states, intention reconsideration becomes anticipatory and can be used to coordinate behavior with prediction of future states.

A further ability based on expectation processing is to allow agents to modify their courses of action in order to anticipatorily coordinate with the other agents behavior. At this stage, agents were investigated that are able to thwart some expected events if the expectation is threatful or to promote them if the expectation is promising (with respect of the ongoing goals) (Piunti et al., 2007a).

Surprise Signals as Filters for the Update of Relevant Beliefs Taking into account the above mentioned model, a novel approach in epistemic reasoning and active perception was proposed. A *surprise driven* belief update processes introducing a notion of information relevance based on goal processing was presented by (Lorini and Piunti, 2007). By considering a proactive and an anticipatory perceptive process, the proposed model implements a novel strategy for epistemic reasoning according to which agents search and filter information from their environment not by monitoring nor perceiving all the retrievable data, but according their ongoing needs, desires, concerns, thus filtering and assessing only what is expected to be relevant for pursuing their goals.

The surprise based filter mechanism allows agents to consider useful for their belief updates only those information related with their goals and expectations. Raised from a mismatch between agent's knowledge and his perceived facts, a surprise signal is sent back to the control system in order to trigger a belief update process. The filter mechanisms is then responsible for (1) signaling the inconsistency between beliefs and an incoming input which is relevant with respect to the current task and (2) the revision of Beliefs and Expectations on the basis of the incoming relevant information.

The proposed filter allow agents to realize as useless and unnecessary those additional costs spent for data processing. Hence, surprise governed attention enable agents to process and filter the perceived data according to the ongoing expectations, modeled on the basis of a knowledge model (the belief base) coupled with subjective goals importance (related to subjective desires, purposes and concerns). In so doing they acquire the capability to divide the overall set of perceivable data in a *relevant* and *irrelevant* subsets.

Our experimental analysis measures the costs for perception and belief updates in agents engaged in dynamic environments. The results show that to higher environment dynamism, the greater costs sustained for epistemic activities are not compensated by an enhancement of achieved tasks. This elicit an important, general strategy exploitable by all those agent engaged in information rich worlds, with big sized information sources to be reported in their belief base. In these conditions, only the relevant information was filtered from the environments that was guessed to be a critical issue for forthcoming cognitive systems (consider for example agents engaged in an information retrieval task in the context of open system applications).

11.15 On-Line and Off-Line Anticipation for Action Control

In a series of studies, the effectiveness of reactive and anticipatory control architectures in predator-prey scenarios were studied. In these scenarios, it was necessary to satisfy competing drives, such as hunger and the avoidance of predators (Pezzulo and Calvi, 2006a; Pezzulo, 2008b).

An important aspect of the undertaken investigations focused on the trade-off between accurate control of action and time spent to form predictions. Anticipatory mechanisms used on-line with action make it more accurate and permit forecasting possible dangers arising from it. At the same time, prediction is a costly operation that requires time, and in principle it can make situated agents less effective or less responsible to dynamics in their environment such as dangers. This trade-off is further complicated when anticipatory mechanisms are used off-line with action to predict multiple and/or distal events (for example, the long-term outcome of several alternative potential courses of actions). Here the (computational) costs for engaging in 'imagination' and 'planning' are higher and can prevent effective situated action. Due to the failure of the earlier AI systems to deal with this trade-offs, it is believed by several researchers that belong to the 'novel AI' (Brooks, 1991) that situated agents should better act than reason.

Advantages of Anticipatory Mechanisms used On-Line with Action In a first study (Pezzulo and Calvi, 2006a), the performance of two schema-based agent architectures were compared in a predator-prey scenario involving multiple entities (predators and preys), obstacles, and moving objects. The first agent architecture included multiple perceptual and motor schemas (for detecting and escaping predators, detecting and catching preys, etc.), having one inverse and one forward model (Wolpert and Kawato, 1998) each. These two (coupled) internal models were used for determining the motor action and for predicting its sensory effects respectively. Prediction errors (of the forward models) were used for action control and schema selection: schemas generating reliable predictions (and related to the current active drives) were selected for controlling action. The second agent instead lacked the internal forward models.

The results indicated that the first agent architecture demonstrates a better adaptivity, since it was better able to satisfy its multiple drives (Pezzulo and Calvi, 2006a). This shows that in such dynamical and demanding environments it is advantageous to use on-line anticipation for action control and schemas selection, despite the costs of running forward models in real time.

Overall, contrary to the view that situated agents need to be reactive (Brooks, 1991), this experiment indicates that internal modeling is highly advantageous for situated agents if it is done on-line with action and produces representations whose format is compatible with the agent's sensorimotor loop (as in the case of internal forward models).

Further Advantages of Off-Line Simulations In a second study (Pezzulo, 2008b), it was investigated how the same anticipatory mechanisms exploited in the first study can be exploited off-line to anticipate several steps in the future (an internal, 'mental' simulation of behavior) for the sake of (1) preventing dangers and (2) planning goal-directed action (that is, mentally generating and selecting sequences of actions to realize further). In this case, the possible advantages or disadvantages of engaging in 'imagination' during situated actions were tested.

Again, two agent architectures were compared, with and without the capability to re-enact schemas in simulation. In the first agent architecture, the same sensorimotor schemas adopted in the first study where used, but now they were allowed to run off-line in *simulation mode*, to predict the long-term sensory consequences of their motor commands. In simulation mode, motor commands were inhibited (not sent to the actuators), but fed as sensory inputs to the forward models, which then produced new sensory predictions that were used by the inverse models for generating new motor commands 'as if' the agents actually sensed the predicted future. The loop between forward and inverse models allowed generating long-term predictions for an arbitrary number of future steps. Two additional mechanisms were responsible for (1) stopping the current action if its predicted outcomes are evaluated as dangerous (a kind of 'somatic marker' mechanism, Damasio, 1994), and (2) evaluating, selecting, and activating the better 'plans' (that is, sequences of 'simulated' actions). The second agent architecture used forward models in the on-line control of action, but lacked the ability to run in generation mode.

The two agent architectures were tested in a tasks consisting in exploring a simulated house for finding a 'treasure' without being captured by guards (as in the first study, each agent architecture had concurrent drives to satisfy). Our results have shown that the first agent architecture, using anticipation both on-line and off-line (that is, in generation mode), outperformed the second one in terms of adaptivity thanks to its capability to predict possible future dangers and to plan from time to time (Pezzulo, 2008b). This happened despite the costs of simulating and planning.

Consistently with recent simulative theories of cognition (Damasio, 1994; Grush, 2004; Hesslow, 2002), the results of our experiments indicate that *mental simulation* is an effective strategy for avoiding dangers and planning in dynamic environments despite the fact that 'imagination' can in principle make an agent less efficacious in its current sensorimotor interaction. As argued in (Grush, 2004; Pezzulo and Castelfranchi, 2007), it is believed that off-line mental simulation is a suitable, embodied alternative to 'reasoning by symbol-crunching' of traditional AI systems, since it permits internal manipulation of (anticipatory) representations without losing grounding and situatedness. Thanks to anticipation, artificial systems can engage in mental operations that when performed by 'ungrounded' AI systems typically determine a poor performance in situated activities.

Overall, these studies have the potential to shed light on the role of mental simulations, how they enable increasingly complex cognitive capabilities, and how they open the passage from present-directed to goal-directed, purposive actions (Pezzulo, 2008a; Pezzulo and Castelfranchi, 2007). See also (Pezzulo et al., 2005; Pezzulo and Calvi, 2005, 2007a) for related studies.

11.16 Conclusion

We have surveyed fourteen case studies, which were developed during the EU project **MindRACES**. In each study, anticipatory mechanisms of (simulated or real) robots determined an advantage in terms of behavioral flexibility, adaptability, reliability, or social interaction. The presented case studies, however, are only few among a growing number of examples of how anticipatory mechanisms can enhance the cognitive and behavioral capabilities of artificial systems or can even develop novel capabilities that may be impossible without anticipatory mechanisms. Overall, these studies illustrate that our design methodology, based on a systematic endowment of artificial systems with anticipatory capabilities, determine huge advantages in terms of increased autonomy and adaptivity.

We hope that the design methodology we have illustrated throughout the book—and exemplified in this chapter with success cases—can inspire future research and real-world product development of cognitive system architectures. Taking the passage 'from Reactive to Anticipatory Cognitive Embodied Systems', this book may serve as the guideline to successfully design and develop further truly flexible, adaptive, effective, and interactive cognitive agents.

References

Adams, J.A.: A closed-loop theory of motor learning. Journal of Motor Behavior 3, 111–149 (1971)

Agre, P.E., Chapman, D.: Pengi: an implementation of a theory of activity. In: Proceedings of AAAI-87, Seattle, WE, pp. 268–272 (1987)

Ahn, H., Picard, R.W.: Affective cognitive learning and decision making: The role of emotions. In: 18th European Meeting on Cybernetics and Systems Research (EMCSR 2006), Vienna, Austria (2006)

Allport, A.: Visual attention. In: Posner, M.I. (ed.) Foundations of Cognitive Science, MIT Press, Cambridge (1990)

Amari, S.: Dynamics of pattern formation in lateral inhibition type neural fields. Biological Cybernetics 27(2), 77–87 (1977)

Anderson, J.R.: The Architecture of Cognition. Harvard University Press, Cambridge (1983)

Anderson, M.C., Spellman, B.A.: On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. Psychological Review 102(1), 68–100 (1995)

Ans, B., Rousset, S., French, R.M., Musca, S.: Preventing catastrophic interference in multiple- sequence learning using coupled reverberating elman networks. In: Proceedings of the 24th Annual Conference of the Cognitive Science Society (2002)

Arbib, M., Rizzolatti, G.: Neural expectations: A possible evolutionary path from manual skills to language. Communication and Cognition 29, 393–424 (1997)

Arbib, M.A.: The metaphorical brain 2: Neural networks and beyond. Wiley, New York (1989)

Arbib, M.A.: Schema theory. In: Shapiro, S. (ed.) Encyclopedia of Artificial Intelligence, 2nd edn., vol. 2, pp. 1427–1443. Wiley, Chichester (1992)

Arbib, M.A.: The mirror system, imitation, and the evolution of language. In: Dautenhahn, K., Nehaniv, C.L. (eds.) Imitation in animals and artifacts, pp. 229–280. MIT Press, Cambridge (2002)

Arbib, M.A.: Schema theory. In: The Handbook of Brain Theory and Neural Networks, 2nd edn., MIT Press, Cambridge (2003)

Arbib, M.A.: From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. Behavioral and Brain Sciences 28, 105–167 (2005)

Arkin, R.C.: Behavior-Based Robotics. MIT Press, Cambridge (1998)

Arkin, R.C., Ali, K., Weitzenfeld, A., Cervantes-Pérez, F.: Behavioral models of the praying mantis as a basis for robotic behavior. Robotics and Autonomous Systems 32(1), 39–60 (2000)

Arnold, M.: Emotion and Personality. Columbia University Press, New York (1960)

Atance, C.M., O'Neill, D.K.: Episodic future thinking. Trends Cogn. Sci. 5, 533–539 (2001)

Axelrod, R.: The Evolution of Cooperation. Basic Books, New York (1984)

Bakker, B., Schmidhuber, J.: Hierarchical reinforcement learning based on subgoal discovery and subpolicy specialization. In: Groen, F., Amato, N., Bonarini, A., Yoshida, E., Krose, B. (eds.) Proceedings of the 8-th Conference on Intelligent Autonomous Systems, IAS-8, Amsterdam, The Netherlands, pp. 438–445 (2004)

Bakker, B., Zhumatiy, V., Gruener, G., Schmidhuber, J.: Quasi-online reinforcement learning for robots. In: Proceedings of the International Conference on Robotics and Automation (ICRA-06), Orlando, Florida (2006)

Baldassarre, G.: Coarse planning for landmark navigation in a neural-network reinforcement learning robot. In: Proceedings of the International Conference on Intelligent Robots and Systems, IROS-2001 (2001)

Baldassarre, G.: A modular neural-network model of the basal ganglia's role in learning and selecting motor behaviours. Journal of Cognitive Systems Research 3, 5–13 (2002a)

Baldassarre, G.: Planning with neural networks and reinforcement learning. PhD thesis, Computer Science Department, University of Esssex (2002b)

Baldassarre, G.: Forward and bidirectional planning based on reinforcement learning and neural networks in a simulated robot. In: Butz, M.V., Sigaud, O., Gérard, P. (eds.) Anticipatory Behavior in Adaptive Learning Systems. LNCS (LNAI), vol. 2684, pp. 179–200. Springer, Heidelberg (2003)

Baldi, P.: Surprise: A shortcut for attention? In: Itti, L., Rees, G., Tsotsos, J. (eds.) Neurobiology of Attention, Academic Press, London (2004)

Balkenius, C.: Spatial learning with perceptually grounded representations. Robotics and Autonomous Systems 25(3-4), 165–175 (1998)

Balkenius, C.: Attention, habituation and conditioning: Toward a computational model. Cognitive Science Quarterly 1, 171–214 (2000)

Balkenius, C.: Cognitive processes in contextual cueing. In: Schmalhofer, F., Young, R.M., Katz, G. (eds.) Proceedings of the European Cognitive Science Conference 2003, pp. 43–47. Lawrence Erlbaum, Mahwah (2003)

Balkenius, C., Åström, K., Eriksson, A.P.: Learning in visual attention. In: ICPR '04 workshop on learning for adaptable visual systems, LAVS (2004)

Balkenius, C., Gardenfors, P.: Anticipation requires adaptation (commentary on romi nijhawan). Behavioral and Brain Sciences 31(2), 199–200 (2008)

Balkenius, C., Hulth, N.: Attention as selection-for-action: A scheme for active perception. In: Third European Workshop on Advanced Mobile Robots (Eurobot '99), pp. 113–119 (1999)

Balkenius, C., Johansson, B.: Event prediction and object motion estimation in the development of visual attention. In: Proceedings of the Fifth International Conference on Epigenetic Robotics, Nara, Japan (2005)

Balkenius, C., Johansson, B.: Anticipatory models in gaze control: A developmental model. Cognitive Processing 8, 167–174 (2007)

Balkenius, C., Morén, J.: Computational models of classical conditioning: a comparative study. In: Mayer, J.-A., Roitblat, H.L., Wilson, S.W., Blumberg, B. (eds.) From Animals to Animats 5, MIT Press, Cambridge (1998)

Balkenius, C., Morén, J.: A computational model of context processing. In: Meyer, J.-A., Berthoz, A., Floreano, D., Roitblat, H.L., Wilson, S.W. (eds.) From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behaviour, MIT Press, Cambridge (2000)

Balkenius, C., Moren, J.: From isolated components to cognitive systems. ERCIM News, No. 53 (2003)

Balkenius, C., Morén, J., Johansson, B.: System-level cognitive modeling with ikaros. In: Lund University Cognitive Studies, vol. 133 (2007)

Balkenius, C., Winberg, S.: Fast learning in an actor-critic architecture with reward and punishment. In: Proceedings of SCAI (2008)

Ballard, D.H.: Animate vision. Artificial Intelligence 48, 57–86 (1991)

Ballard, D.H., Hayhoe, M.M., Pook, P.K., Rao, R.P.: Deictic codes for the embodiment of cognition. Behav. Brain Sci. 20(4), 723–742 (1997), discussion 743–67. 0140-525X (Print) Journal Article Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S. Review

Bar, M.: Visual objects in context. Nature Reviews: Neuroscience 5, 617–629 (2004)

Bar, M.: The proactive brain: using analogies and associations to generate predictions. Trends in Cognitive Sciences 11(7), 280–289 (2007)

Baraduc, P., Guigon, E., Burnod, Y.: Where does the population vector of motor cortical cells point during reaching movements? In: Kearns, M., Solla, S., Cohn, D. (eds.) Advances in neural information processing systems, vol. 11, pp. 83–89. MIT Press, Cambridge (1999)

Baraduc, P., Guigon, E., Burnod, Y.: Recoding arm position to learn visuomotor transformation. Cerebral Cortex 11, 906–917 (2001)

Barsalou, L.W.: Perceptual symbol systems. Behavioral and Brain Sciences 22, 577–600 (1999)

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

Barto, A.G., Fagg, A.H., Sitkoff, N., Houk, J.C.: A cerebellar model of timing and prediction in the control of reaching. Neural Computation 11, 565–594 (1999)

Barto, A.G., Mahadevan, S.: Recent advances in hierarchical reinforcement learning. Discrete Event Dynamic Systems 13, 341–379 (2003)

Bates, J.: The nature of characters in interactive worlds and the Oz project. Technical Report CMU-CS-92-200, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA (1992)

Bates, J.: The role of emotion in believable agents. Communications of the ACM 37(7), 122–125 (1994)

Bechara, A., Damasio, H., Traneland, D., Damasio, A.: Deciding advantageously before knowing the advantageous strategy. Science 275, 1293–1295 (1997)

Berlyne, D.E.: Conflict, Arousal and Curiosity. McGraw-Hill, New York (1960)

Berthier, N.E., Singh, S.P., Barto, A.G., Houk, J.C.: A cortico-cerebellar model that learns to generate distributed motor commands to control a kinematic arm. Advances in Neural Information Processing Systems 4, 611–618 (1992)

Berthier, N.E., Singh, S.P., Barto, A.G., Houk, J.C.: Distributed representation of limb motor programs in arrays of adjustable pattern generators. Journal of Cognitive Neuroscience 5, 56–78 (1993)

Bickhard, M.H.: Levels of representationality. Journal of Experimental and Theoretical Artificial Intelligence 10(2), 179–215 (1998)

Bickhard, M.H., Terveen, L.: Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution. Elsevier Scientific, Amsterdam (1995)

Blakemore, S.-J., Decety, J.: From the perception of action to the understanding of intention. Nature Reviews Neuroscience 2, 561–567 (2001)

Blakemore, S.-J., Frith, C.D., Wolpert, D.M.: The cerebellum is involved in predicting the sensory consequences of action. Brain Imaging 12, 1879–1884 (2001)

Blakemore, S.-J., Wolpert, D.M., Frith, C.D.: Central cancellation of self-produced tickle sensation. Nature Neuroscience 1(7), 635–640 (1998)

Blom, H.A.P.: An efficient filter for abruptly changing systems. In: Proc. 23rd Conf. Decision Contr. (1984)

Blom, H.A.P., Bloem, E.A.: Particle filtering for stochastic hybrid systems. In: CDC: 43rd IEEE Conference on Decision and Control, vol. 3, pp. 3221–3226 (2004)

Blumberg, B.M.: D-Learning: what learning in dogs tells us about building characters that learn what they ought to learn. In: Exploring artificial intelligence in the new millennium, pp. 37–67. Morgan Kaufmann Publishers Inc., San Francisco (2003)

Bock, K., Irwin, D.E., Davidson, D.J., Levelt, W.J.M.: Minding the clock. Journal of Memory and Language 48, 653–685 (2003)

Bongard, J., Zykov, V., Lipson, H.: Resilient machines through continuous self-modeling. Science 314(5802), 1118–1121 (2006a)

Bongard, J.C., Zykov, V., Lipson, H.: Automated synthesis of body schema using multiple sensor modalities. In: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems (ALIFEX), pp. 220–226 (2006b)

Botelho, L.M., Coelho, H.: Emotion-based attention shift in autonomous agents. In: Jennings, N.R., Wooldridge, M.J., Müller, J.P. (eds.) ECAI-WS 1996 and ATAL 1996. LNCS, vol. 1193, Springer, Heidelberg (1997)

Bouguet, J.Y.: Pyramidal implementation of the lucas kanade feature tracker: Description of the algorithm (2002)

Boutilier, C.: Iterated revision and minimal change of conditional beliefs. Journal of Philosophical Logic 25(3), 262–305 (1996)

Braitenberg, V.: Vehicles: Experiments in Synthetic Psychology. MIT Press, Cambridge (1984)

Bratman, M.: Intentions, Plans, and Practical Reason. Harvard University Press, Cambridge (1987)

Braubach, L., Pokahr, A., Moldt, D., Lamersdorf, W.: Goal representation for BDI agent systems. In: Bordini, R.H., Dastani, M., Dix, J., El Fallah Seghrouchni, A. (eds.) PROMAS 2004. LNCS (LNAI), vol. 3346, pp. 44–65. Springer, Heidelberg (2005)

Breazeal, C.: Emotion and sociable humanoid robots. International Journal of Human Computer Studies 56, 119–155 (2003)

Breazeal, C., Scassellati, B.: A context-dependent attention system for a social robot. In: IJCAI '99: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, pp. 1146–1153. Morgan Kaufmann, San Francisco (1999)

Breemen, A.: Animation engine for believable interactive user-interface robots. In: IEEE/RSJ - International Conference on Intelligent Robots and Systems, vol. 3, pp. 2873–2878 (2004)

Brooks, R.A.: Intelligence without representation. Artificial Intelligence 47, 139–159 (1991)

Brown, G.W.: Iterative solution of games by fictitious play. In: Koopmans, T.C. (ed.) Activity Analysis of Production and Allocation, pp. 374–376. Wiley, New York (1951)

Buckner, R.L., Carroll, D.C.: Self-projection and the brain. Trends Cogn. Sci. 11(2), 49–57 (2007)

Bullock, D., Grossberg, S., Guenther, F.H.: A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm. Journal of Cognitive Neuroscience 5, 408–435 (1993)

Burke, R., Blumberg, B.M.: Using and ethologically-inspired model to learn apparent temporal causality for planning in synthetic creatures. In: Castelfranchi, C., Johnson, W.L. (eds.) Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems. Part 1, Bologna, Italy, pp. 326–333. ACM Press, New York (2002)

Busemeyer, J.R., Johnson, J.: Computational models of decision making. In: Koehler, D., Harvey, N. (eds.) Handbook of Judgment and Decision Making, pp. 133–154. Blackwell Publishing Co., Oxford (2004)

Busemeyer, J.R., Townsend, J.T., Stout, J.C.: Motivational underpinnings of utility in decision making: Decision field theory analysis of deprivation and satiation. In: Moore, S. (ed.) Emotional Cognition, John Benjamins, Amsterdam (2003)

Butz, M.V.: Anticipatory learning classifier systems. Kluwer Academic Publishers, Boston (2002a)

Butz, M.V.: Biasing exploration in an anticipatory learning classifier system. In: Lanzi, P.L., Stolzmann, W., Wilson, S.W. (eds.) IWLCS 2001. LNCS (LNAI), vol. 2321, pp. 3–22. Springer, Heidelberg (2002b)

Butz, M.V.: Rule-Based Evolutionary Online Learning Systems: A Principled Approach to LCS Analysis and Design. Studies in Fuzziness and Soft Computing, vol. 191. Springer, Heidelberg (2006)

Butz, M.V., Goldberg, D.E.: Generalized state values in an anticipatory learning classifier system. In: Butz, M.V., Sigaud, O., Gérard, P. (eds.) Anticipatory Behavior in Adaptive Learning Systems. LNCS (LNAI), vol. 2684, pp. 282–301. Springer, Heidelberg (2003)

Butz, M.V., Goldberg, D.E., Tharakunnel, K.: Analysis and improvement of fitness exploitation in XCS: Bounding models, tournament selection, and bilateral accuracy. Evolutionary Computation 11, 239–277 (2003a)

Butz, M.V., Herbort, O.: Context-dependent predictions and cognitive arm control with xcsf. In: GECCO 2008: Genetic and Evolutionary Computation Conference, in press (2008)

Butz, M.V., Herbort, O., Hoffmann, J.: Exploiting redundancy for flexible behavior: Unsupervised learning in a modular sensorimotor control architecture. Psychological Review 114, 1015–1046 (2007a)

Butz, M.V., Hoffmann, J.: Anticipations control behavior: Animal behavior in an anticipatory learning classifier system. Adaptive Behavior 10, 75–96 (2002)

Butz, M.V., Kovacs, T., Lanzi, P.L., Wilson, S.W.: Toward a theory of generalization and learning in XCS. IEEE Transactions on Evolutionary Computation 8, 28–46 (2004a)

Butz, M.V., Lanzi, P.L., Wilson, S.W.: Function approximation with XCS: Hyperellipsoidal conditions, recursive least squares, and compaction. IEEE Transactions on Evolutionary Computation, in press (2008a)

Butz, M.V., Reif, K., Herbort, O.: Bridging the gap: Learning sensorimotor-linked population codes for planning and motor control. In: International Conference on Cognitive Systems, CogSys 2008 (2008b)

Butz, M.V., Sigaud, O., Gérard, P.: Anticipatory behavior: Exploiting knowledge about the future to improve current behavior. In: Butz, M.V., Sigaud, O., Gérard, P. (eds.) Anticipatory Behavior in Adaptive Learning Systems. LNCS (LNAI), vol. 2684, pp. 1–10. Springer, Heidelberg (2003b)

Butz, M.V., Sigaud, O., Gérard, P.: Internal models and anticipations in adaptive learning systems. In: Butz, M.V., Sigaud, O., Gérard, P. (eds.) Anticipatory Behavior in Adaptive Learning Systems. LNCS (LNAI), vol. 2684, pp. 86–109. Springer, Heidelberg (2003c)

Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G.: Anticipations, brains, individual and social behavior: An introduction to anticipatory systems. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 1–18. Springer, Heidelberg (2007b)

Butz, M.V., Swarup, S., Goldberg, D.E.: Effective online detection of task-independent landmarks. In: Sutton, R.S., Singh, S. (eds.) Online Proceedings for the ICML'04 Workshop on Predictive Representations of World Knowledge, p. 10 (2004b)

Camacho, E.F., Bordons, C.: Model Predictive Control. Springer, Heidelberg (1999)

Camerer, C., Ho, T.-H., Chong, J.: Sophisticated ewa learning and strategic teaching in repeated games. Theory 104, 137–188 (2002)

Camerer, C.F., Ho, T.-H.: Experience-weighted attraction learning in games: Estimates from weak-link games. In: Budescu, D., Erev, I., Zwick, R. (eds.) Games and Human Behavior: Essays in Honor of Amnon, pp. 31–52. Lawrence Erlbaum, Hillsdale (1999)

Carpenter, G.A., Grossberg, S.: The art of adaptive pattern recognition by a self-organizing neural network. Computer 21(3), 77–88 (1988)

Casarotti, M., Zorzi, M., Umiltá, C.: A computational model of saccadic planning and spatial attention. In: Schmalhofer, F., Young, R.M., Katz, G. (eds.) Proceedings of the European Cognitive Science Conference 2003 (EuroCogSci03), pp. 73–78. Lawrence Erlbaum, Mahwah (2003)

Casati, R., Pasquinelli, E.: How can you be surprised? the case for volatile expectations. Phenomenology and the Cognitive Sciences 6(1-2), 171–183 (2006)

Castelfranchi, C.: Guarantees for autonomy in cognitive agent architecture. In: Wooldridge, M.J., Jennings, N.R. (eds.) ECAI 1994 and ATAL 1994. LNCS, vol. 890, pp. 56–70. Springer, Heidelberg (1995)

Castelfranchi, C.: Modelling social action for AI agents. Artificial Intelligence 103, 157–182 (1998)

Castelfranchi, C.: Mind as an anticipatory device: For a theory of expectations. In: De Gregorio, M., Di Maio, V., Frucci, M., Musio, C. (eds.) BVAI 2005. LNCS, vol. 3704, pp. 258–276. Springer, Heidelberg (2005)

Castelfranchi, C., Falcone, R., Pezzulo, G.: Belief sources for trust: Some learning mechanisms. In: Proceedings of the Sixth International Workshop on Trust, Privacy, Deception and Fraud in Agent Systems c/o AAMAS-03, Melbourne, Australia, pp. 101–106 (2003)

Castelfranchi, C., Miceli, M.: Anticipation and emotions (in press)

Castiello, U.: Mechanism of selection for the control of hand action. Trends in Cognitive Sciences 3(7), 264–271 (1999)

Castiello, U., Bennett, K.M., Paulignan, Y.: Does the type of prehension influence the kinematics of reaching? Behav. Brain Res. 50(1-2), 7–15 (1992), 0166-4328 (Print) Journal Article Research Support, Non-U.S. Gov't.

Chalmers, D., French, R., Hofstadter, D.: High-level perception, representation, and analogy. Journal of Experimental and Theoretical Artificial Intelligence 4, 185–211 (1992)

Chen, B., Kaneko, M.: Behavior selection of mobile robot based on integration of multimodal information. Electrical Engineering in Japan 158, 39–48 (2007)

Chi, M.T.H., Feltovich, P., Glaser, R.: Categorization and representation of physic problems by experts and novices. Cognitive Science 5, 121–152 (1981)

Chiel, H.J., Beer, R.D.: The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. Trends in Neurosciences 20, 553–557 (1997)

Choi, S.-B., Ban, S.-W., Lee, M.: Biologically motivated visual attention system using bottom-up saliency map and top-down inhibition. Neural Information Processing-Letters and Reviews 2(1), 19–25 (2004)

Cholvy, L., Hunter, A.: Information fusion in logic: A brief overview. In: ECSQARU-FAPR, pp. 86–95 (1997)
Chun, M.M., Jiang, Y.: Contextual cueing: Implicit learning and memory of visual context guides spatial attention. Cognitive Psychology 36(1), 28–71 (1998)
Cisek, P.: Integrated neural processes for defining potential actions and deciding between them: A computational model. Journal of Neuroscience 26, 9761–9770 (2006)
Cisek, P., Kalaska, J.F.: Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action? Neuron 45(5), 801–814 (2005)
Clark, A.: A companion to cognitive science. In: Bechtel, W., Graham, G. (eds.) Embodied, situated, and distributed cognition, pp. 506–517. Blackwell, Malden (1998)
Clark, A., Grush, R.: Towards a cognitive robotics. Adaptive Behavior 7(1), 5–16 (1999)
Colman, A.M.: Cooperation, psychological game theory, and limitations of rationality in social interaction. Behavioral Brain Science 26, 139–153 (2003)
Colombo, M., Graziano, M.S.: Effects of auditory and visual interference on auditory-visual delayed matching to sample in monkeys (maca fascicularis). Behav. Neurosci. 108, 636–639 (1994)
Colwill, R.M., Rescorla, R.A.: Postconditioning devaluation of a reinforcer affects instrumental learning. Journal of Experimental Psychology: Animal Behavior Processes 11(1), 120–132 (1985)
Colwill, R.M., Rescorla, R.A.: Evidence for the hierarchical structure of instrumental learning. Animal Learning & Behavior 18(1), 71–82 (1990)
Corchs, S., Deco, G.: Large-scale neural model for visual attention: integration of experimental single-cell and fmri data. Cereb Cortex 12(4), 339–348 (2002), 1047-3211 (Print) Journal Article Research Support, Non-U.S. Gov't
Coricelli, G., Dolan, R.J., Sirigu, A.: Brain, Emotion and Decision making: the paragdigmatic example of Regret. TRENDS in Cognitive Sciences 11(6), 258–265 (2007)
Correa, M., Coelho, H.: Collective mental states in an extended mental states framework. In: Proc. of the International Conf. on Collective Intentionality IV, Siena, Italy (2004)
Cotterill, R.: Enchanted Looms: Conscious Networks in Brains and Computers. Cambridge University Press, Cambridge (1998)
Craik, K.: The Nature of Explanation. Cambridge University Press, Cambridge (1943)
Csibra, G., Gergely, G.: 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. Acta Psychologica 124, 60–78 (2007)
Dacey, D.M.: Circuitry for color coding in the primate retina. Proceedings of the National Academy of Sciences 93, 582–588 (1996)
Damasio, A.R.: Descartes' Error: Emotion, Reason and the Human Brain. Grosset/Putnam, New York (1994)

Dautenhahn, K., Nehaniv, C.: Imitation in animals and artifacts. MIT Press, Cambridge (2002)

Davidsson, P.: Learning by linear anticipation in multi-agent systems. In: Weiss, G. (ed.) ECAI 1996 Workshops. LNCS, vol. 1221, pp. 62–72. Springer, Heidelberg (1997)

Davidsson, P.: A framework for preventive state anticipation. In: Butz, M.V., Sigaud, O., Gérard, P. (eds.) Anticipatory Behavior in Adaptive Learning Systems. LNCS (LNAI), vol. 2684, pp. 151–166. Springer, Heidelberg (2003)

Dawkins, R.: The Selfish Gene. Oxford University Press, Oxford (1989)

Dayan, P.: Motivated reinforcement learning. In: Dietterich, T.G., Becker, S., Ghahramani, Z. (eds.) Advances in Neural Information Processing Systems 14, MIT Press, Cambridge (2002)

Dayan, P., Balleine, B.W.: Reward, motivation and reinforcement learning. Neuron 36, 285–298 (2002)

Dayan, P., Hinton, G.E., Neal, R.M., Zemel, R.S.: The helmholtz machine. Neural Computation 7, 889–904 (1995)

Dayan, P., Kakade, S., Montague, P.R.: Learning and selective attention. Nature neuroscience 3(11s), 1218–1223 (2000a)

Dayan, P., Kakade, S., Montague, P.R.: Learning and selective attention. Nature Neuroscience 3, 1218–1223 (2000b)

de Brecht, M., Saiki, J.: A neural network implementation of a saliency map model. Neural Networks 19(10), 1467–1474 (2006)

De Castro, L.N., Timmis, J.: Artificial immune systems as a novel soft computing paradigm. Soft Computing 7, 526–544 (2003)

Deacon, T.W.: The Symbolic Species: The Co-evolution of Language and the Brain. W.W. Norton, New York (1997)

Decety, J., Chaminade, T.: When the self represents the other: A new cognitive neuroscience view on psychological identification. Consciousness and Cognition 12(20), 577–596 (2003)

Decety, J., Grèzes, J.: Neural mechanisms subserving the perception of human actions. Trends Cogn. Sci. 3, 172–178 (1999)

Decety, J., Grèzes, J.: The power of simulation: Imagining one's own and other's behavior. Brain Res. 1079(1), 4–14 (2006)

Decety, J., Ingvar, D.H.: Brain structures participating in mental simulation of motor behavior: a neuropsychological interpretation. Acta Psychol. 73, 13–24 (1990)

Deco, G., Rolls, E.T.: Attention, short-term memory, and action selection: a unifying theory. Progress in Neurobiology 76(4), 236–256 (2005)

Deco, G., Zihl, J.: A neurodynamical model of visual attention: feedback enhancement of spatial resolution in a hierarchical system. J. Comput. Neurosci. 10(3), 231–253 (2001)

Dehaene, S., Changeux, J.P.: A hierarchical neuronal network for planning behavior. Proc. Natl. Acad. Sci. U S A 94(24), 13293–13298 (1997)

Demiris, J., Hayes, G.: Imitative learning mechanisms in robots and humans. In: Proceedingsof the 5th European Workshop on Learning Robot, pp. 9–16 (1996)

Demiris, Y.: Prediction of intent in robotics and multi-agent systems. Cognitive Processing 8(3), 151–158 (2007)

Demiris, Y., Khadhouri, B.: Hierarchical attentive multiple models for execution and recognition (hammer). Robotics and Autonomous Systems Journal 54, 361–369 (2005)

Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM Algorithm. Journal of the Royal Statistical Society 39(1), 1–38 (1977)

Deneve, S., Pouget, A.: Bayesian multisensory integration and cross-modal spatial links. Journal of Physiology - Paris 98, 249–258 (2004)

Dennett, D.: Cognitive wheels: The frame problem in ai. In: Hookway, C. (ed.) Minds, Machines, and Evolution, pp. 128–151. Cambridge University Press, Cambridge (1984)

Dennett, D.: The intentional stance. MIT Press, Cambridge (1987)

Desimone, R.: Visual attention mediated by biased competition in extrastriate cortex. Philosophical Transactions of the Royal Society of London Series B 353, 1245–1255 (1998)

Desimone, R., Duncan, J.: Neural mechanisms of selective visual attention. Annual Review of Neuroscience 18, 193–222 (1995)

Desmurget, M., Grafton, S.: Forward modeling allows feedback control for fast reaching movements. Trends in Cognitive Sciences 4, 423–431 (2000)

DGT, Dgt electronic chessboard (2007), Internet website http://www.dgtprojects.com/eboard.htm (retrieved in 17-07-2007)

Di Ferdinando, A., Casarotti, M., Vallar, G., Zorzi, M.: Hemisperic asymmetries in the neglect syndrome: A computational study. In: Cangelosi, A., Bugmann, G., Borisyuk, R. (eds.) NCPW 9: Modelling Language, Cognition, and Action, World Scientific, London (2004)

Dias, J.: Fearnot!: Creating emotional autonomous synthetic characters for emphatic interactions. Master's thesis, Instituto Superior Tecnico, Lisboa, Portugal (2005)

Dickinson, A., Balleine, B.W.: The role of learning in motivation. In: Gallistel, C.R. (ed.) Learning, motivation and emotion, pp. 407–430. John Wiley & Sons, Chichester (2002)

Dittrich, P., Kron, T., Banzhaf, W.: On the scalability of social order: Modeling the problem of double and multi contingency following luhmann. Journal of Artificial Societies and Social Simulation 6 (2003)

Dominey, P.F., Arbib, M.A.: A cortico-subcortical model for generation of spatially accurate sequential saccades. Cerebral Cortex 2(2), 153–175 (1992)

Doya, K.: Temporal difference learning in continuous time and space. In: Touretzky, D.S., Mozer, M.C., Hasselmo, M.E. (eds.) Advances in Neural Information Processing Systems, vol. 8, pp. 1073–1079. MIT Press, Cambridge (1996)

Doya, K.: What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw. 12(7-8), 961–974 (1999)

Doya, K.: Metalearning and neuromodulation. Neural Netw. 15(4-6), 495–506 (2002)

Doyle, J.: Rationality and its roles in reasoning. Computational Intelligence 8(2), 376–409 (1992)

Drescher, G.L.: Made-up minds: A constructivist approach to artificial intelligence. MIT Press, Cambridge (1991)

D'Souza, A., Vijayakumar, S., Schaal, S.: Learning inverse kinematics. In: IEEE International Conference on Intelligent Robots and Systems (IROS 2001), vol. 1, pp. 298–303 (2001)

Dubois, D.M.: Computing anticipatory systems with incursion and hyperincursion. In: Dubois, D.M. (ed.) CASYS'97, AIP Proceedings 437. AIP Proceedings, vol. 437, pp. 3–29. American Institute of Physics, Woodbury (1998)

Duncan, J.: Selective attention and the organization of visual information. Journal of Experimental Psychology: General 113, 501–517 (1984)

Elliot, C.: The Affective reasoner: A process model for emotions in a multi-agent system. PhD thesis, Northwestern University Institute for Learning Sciences (1992)

Elliott, C., Rickel, J., Lester, J.C.: Lifelike pedagogical agents and affective computing: An exploratory synthesis. In: Veloso, M.M., Wooldridge, M.J. (eds.) Artificial Intelligence Today. LNCS (LNAI), vol. 1600, pp. 195–211. Springer, Heidelberg (1999)

Elman, J.L.: Finding structure in time. Cognitive Science 14(2), 179–211 (1990)

Elster, J.: Rationality and the emotions. The Economic Journal 106, 1386–1397 (1996)

Engel, A.K., Fries, P., Singer, W.: Dynamic predictions: Oscillations and synchrony in top-down processing. Nature Reviews Neuroscience 2(10), 704–716 (2001)

Engelbrecht, S.E.: Minimum principles in motor control. Journal of Mathematical Psychology 45, 497–542 (2001)

Epstein, J.M., Axtell, R.: Growing Artificial Societies: Social Science from the Bottom Up. MIT Press, Cambridge (1996)

Erlhagen, W., Schöner, G.: Dynamic field theory of movement preparation. Psychological Review 109(3), 545–572 (2002)

Falkenhainer, B., Forbus, K., Gentner, D.: The structure mapping engine: Algorithm and examples. Artificial Intelligence 41(1), 1–63 (1989)

Fischer, M.H., Rosenbau, D.A., Vaughan, J.: Speed and sequential effects in reaching. Journal of Experimental Psychology: Human Perception and Performance 23(2), 404–428 (1997)

Fitts, P.M., Posner, M.I.: Human Performance. Brooks/Cole, Belmont (1967)

Fleischer, J.G.: Neural correlates of anticipation in cerebellum, basal ganglia, and hippocampus. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 19–34. Springer, Heidelberg (2007)

Fogassi, L., Ferrari, P.F., Chersi, F., Gesierich, B., Rozzi, S., Rizzolatti, G.: Parietal lobe: from action organization to intention understanding. Science 308, 662–667 (2005)

Forbus, K., Gentner, D., Law, K.: Mac/fac: A model of similarity-based retrieval. Cognitive Science 19(2), 141–205 (1995)

French, R.: The subtlety of sameness: A theory and computer model of analogy making. MIT Press, Cambridge (1995)

Fridjda, N.H.: The Emotions. Cambridge University Press, Cambridge (1986)

Fries, P., Nikolic, D., Singer, W.: The gamma cycle. Trends in Neurosciences 30, 309–316 (2007)

Frijda, N.: Emotion, cognitive structure and action tendency. Cognition and Emotion 1, 115–143 (1987)

Friston, K.: Learning and inference in the brain. Neural Netw. 16(9), 1325–1352 (2003)

Friston, K.: A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360(1456), 815–836 (2005)

Frith, C.: Making up the Mind. How the Brain Creates our Mental World. Blackwell, Malden (2007)

Frith, C.D., Blakemore, S.J., Wolpert, D.M.: Abnormalities in the awareness and control of action. Philos. Trans. R. Soc. Lond. B Biol. Sci. 355(1404), 1771–1788 (2000)

Frith, C.D., Frith, U.: How we predict what other people are going to do. Multiple Perspectives on the Psychological and Neural Bases of Understanding Other People's Behavior 1079(1), 36–46 (2006)

Gallese, V.: The inner sense of action: agency and motor representations. Journal of Consciousness Studies 7, 23–40 (2000)

Gallese, V.: The 'shared manifold' hypothesis: From mirror neurons to empathy. Journal of Consciousness Studies: Between Ourselves - Second-Person Issues in the Study of Consciousness 8(5-7), 33–50 (2001)

Gallese, V., Goldman, A.: Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Sciences 2(12), 493–501 (1998)

Gallese, V., Keysers, C., Rizzolatti, G.: A unifying view of the basis of social cognition. Trends Cogn. Sci. 8(9), 396–403 (2004)

Gallese, V., Metzinger, T.: Motor ontology: The representational reality of goals, actions, and selves. Philosophical Psychology 13(3), 365–388 (2003)

Gallistel, C.R., Gibbon, J.: Time, rate and conditioning. Psychological Review 107, 289–344 (2000)

Gardenfors, P.: How Homo Became Sapiens: On the Evolution of Thinking. Oxford University Press, Oxford (2003)

Gardenfors, P.: Cooperation and the evolution of symbolic communication. In: Oller, K., Griebel, U. (eds.) Evolution of Communication Systems: A Comparative Approach, pp. 237–256. MIT Press, Cambridge (2004)

Gardenfors, P., Orvath, M.: The evolution of anticipatory cognition as a precursor to symbolic communication. In: Proceedings of the Morris Symposium on the Evolution of Language, Stony Brook, NY, USA (2005)

Gentner, D.: Structure-mapping: A theoretical framework for analogy. Cognitive Science 7, 155–170 (1983)

Georgeff, M., Pell, B., Pollack, M.E., Tambe, M., Wooldridge, M.J.: The Belief-Desire-Intention Model of Agency. In: Rao, A.S., Singh, M.P., Müller, J.P. (eds.) ATAL 1998. LNCS (LNAI), vol. 1555, pp. 1–10. Springer, Heidelberg (1999)

Gérard, P., Meyer, J.-A., Sigaud, O.: Combining latent learning and dynamic programming in MACS. European Journal of Operational Research 160, 614–637 (2005)

Gérard, P., Sigaud, O.: YACS: Combining dynamic programming with generalization in classifier systems. In: Lanzi, P.L., Stolzmann, W., Wilson, S.W. (eds.) IWLCS 2000. LNCS (LNAI), vol. 1996, pp. 52–69. Springer, Heidelberg (2001)

Gers, F.A., Schmidhuber, J., Cummins, F.: Learning to forget: Continual prediction with LSTM. Neural Computation 12, 2451–2471 (2000)

Gers, F.A., Schraudolph, N.N., Schmidhuber, J.: Learning precise timing with lstm recurrent networks. J. Mach. Learn. Res. 3, 115–143 (2003)

Gibson, J.J.: The Senses Considered as Perceptual Systems. Houghton Mifflin, Boston (1966)

Giese, M.A., Poggio, T.: Neural mechanisms for the recogniton of biological movements. Nature Reviews Neuroscience 4, 179–192 (2003)

Gilovich, T., Medvec, V.H.: The temporal pattern to the experience of regret. Journal of Personality and Social Psychology 67(3), 357–365 (1994)

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

Gloye, A., Goektekin, C., Egorova, A., Tenchio, O., Rojas, R.: Fu-fighters small size. In: RoboCup-2004: Robot Soccer World Cup VIII (2004)

Gmytrasiewicz, P.J., Lisetti, C.L.: Using decision theory to formalize emotions for multi-agent systems. In: 2nd Workshop on Game Theoretic and Decision Theoretic Agents, ICMAS-2000 (2000)

Gnadt, J.W., Andersen, R.A.: Memory related motor planning activity in posterior parietal cortex of macaque. Experimental Brain Research 70(1), 216–220 (1988)

Goldman, A.: Imitation, mind reading, and simulation. In: Hurley, S., Chater, N. (eds.) Perspectives on Imitation II, pp. 80–81. MIT Press, Cambridge (2005)

Gómez, M., Carbó, J., Benac-Earle, C.: An anticipatory trust model for open distributed systems. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 307–324. Springer, Heidelberg (2007)

Goncalves, L.M.G., Giraldi, G.A., Antonio, A.F.O., Grupen, R.A.: Learning policies for attentional control. In: Proceedings of the 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation, CIRA '99 (1999)

Goswami, U.: Analogical reasoning in children. In: Holyoak, K., Gentner, D., Kokinov, B. (eds.) The Analogical Mind: Perspectives from Cognitive Science, MIT Press, Cambridge (2001)

Gratch, J., Marsella, S.: A domain-independent framework for modeling emotion. Cognitive Systems Research 5, 269–306 (2004)

Gratch, J., Marsella, S.: The architectural role of emotions in cognitive systems. In: Gray, W. (ed.) Integrated Models of Cognitive Systems, Oxford University Press, Oxford (2006)

Griffin, Z.M., Bock, J.K.: What the eyes say about speaking. Psychological Science 11, 274–279 (2000)

Gross, H.-M., Volker, S., Torsten, S.: A neural architecture for sensorimotor anticipation. Neural Networks 12, 1101–1129 (1999)

Grossberg, S.: The complementary brain: Unifying brain dynamics and modularity. Trends in Cognitive Sciences 4, 233–246 (2000)

Grush, R.: The emulation theory of representation: Motor control, imagery, and perception. Behavioral and Brain Sciences 27, 377–396 (2004)

Guttenplan, S.: A companion to the philosophy of mind. Blackwell Publishers Ltd., Oxford (1994)

Hamilton, A., Grafton, S.T.: The motor hierarchy: from kinematics to goals and intentions. In: Haggard, P., Rossetti, Y., Kawato, M. (eds.) Sensorimotor Foundations of Higher Cognition, Oxford University Press, Oxford (2007)

Hammond, L.: Conditioned emotional state. In: Black, P. (ed.) Physiological Correlates of Emotion, pp. 245–259. Academic Press, New York (1970)

Harlow, H., Stagner, R.: Psychology of feelings and emotions. ii. theory of emotions. Psychological Review 40(2), 84–194 (1933)

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

Harris, C., Stephens, M.: A combined corner and edge detector. In: 4th Alvey Vision Conference, pp. 147–151 (1988)

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

Haruno, M., Wolpert, D.M., Kawato, M.: Hierarchical mosaic for movement generation. In: Ono, T., Matsumoto, G., Llinas, R.R., Berthoz, A., Norgren, H., Tamura, R. (eds.) Excepta Medica International Coungress Series, Elsevier Science, Amsterdam (2003)

Hawkins, J., Blakeslee, S.: On intelligence. Henry Holt and Company, New York (2004)

Held, R., Hein, A.: Movement-produced stimulation in the development of visually guided behavior. Journal of Comparative and Physiological Psychology 56(5), 872–876 (1963)

Herbart, J.F.: Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter, analytischer Teil. August Wilhem Unzer, Königsberg (1825)

Herbort, O., Butz, M.V.: Encoding complete body modely enables task dependent optimal behavior. In: Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007, pp. 1424–1429 (2007)

Herbort, O., Ognibene, D., Butz, M.V., Baldassarre, G.: Learning to select targets within targets in reaching tasks. In: 6th IEEE International Conference on Development and Learning, ICDL 2007, pp. 7–12 (2007)

Herwig, U., Kaffenberger, T., Baumgartner, T., Jancke, L.: Neural correlates of a 'pessimistic' attitude when anticipating events of unknown emotional valence. NeuroImage 34(2), 848–858 (2007)

Hesslow, G.: Conscious thought as simulation of behaviour and perception. Trends in Cognitive Sciences 6, 242–247 (2002)

Hinton, G., Dayan, P.: Varieties of helmholtz machine. Neural Netw. 9(8), 1385–1403 (1996)

Hinton, G., McClelland, J.: Learning representations by recirculation. In: Anderson, D.Z. (ed.) Neural Information Processing Systems, pp. 358–366. American Institute of Physics, New York (1988)

Hinton, G.E.: Learning multiple layers of representation. Trends in Cognitive Sciences 11, 428–434 (2007)
Hinton, G.E., Dayan, P., Frey, B.J., Neal, R.M.: The wake-sleep algorithm for unsupervised neural networks. Science 268, 1158–1161 (1995)
Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation 9, 1735–1780 (1997)
Hoffman, M.W., Grimes, D.B., Shon, A.P., Rao, R.P.N.: A probabilistic model of gaze imitation and shared attention. Neural Networks 19(3), 299–310 (2006)
Hoffmann, H.: Perception through visuomotor anticipation in a mobile robot. Neural Networks 20, 22–33 (2007)
Hoffmann, H., Möller, R.: Action selection and mental transformation based on a chain of forward models. In: Schaal, S., Ijspeert, A., Billard, A., Vijayakumar, S., Hallam, J., Meyer, J.-A. (eds.) From Animals to Animats 8, Proceedings of the Eighth International Conference on the Simulation of Adaptive Behavior, Los Angeles, CA, pp. 213–222. MIT Press, Cambridge (2004)
Hoffmann, J.: Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung und Wahrnehmung (Anticipation and cognition: The function of anticipations in human behavioral control and perception). Hogrefe, Göttingen (1993)
Hoffmann, J.: Anticipatory behavioral control. In: Butz, M.V., Sigaud, O., Gérard, P. (eds.) Anticipatory Behavior in Adaptive Learning Systems. LNCS (LNAI), vol. 2684, pp. 44–65. Springer, Heidelberg (2003)
Hoffmann, J., Berner, M., Butz, M.V., Herbort, O., Kiesel, A., Kunde, W., Lenhard, A.: Explorations of anticipatory behavioral control (abc): a report from the cognitive psychology unit of the University of Würzburg. Cogn. Process. 8(2), 133–142 (2007)
Hoffmann, J., Stöcker, C., Kunde, W.: Anticipatory control of actions. International Journal of Sport and Exercise Psychology 2, 346–361 (2004)
Hofstadter, D., the Fluid Analogies Research Group: Fluid concepts and creative analogies. Basic Books, New York (1995)
Hofstadter, D.R.: Analogy as the core of cognition. In: Gentner, D., Holyoak, K.J., Kokinov, B.N. (eds.) The Analogical Mind: Perspectives from Cognitive Science, pp. 499–538. The MIT Press/Bradford Book, Cambridge (2001)
Holmes, M., Isbell, C.: Schema learning: Experience-based construction of predictive action models. In: Advances in Neural Information Processing Systems, vol. 17, pp. 585–592 (2005)
Holyoak, K., Gentner, D., Kokinov, B.: The place of analogy in cognition. In: Holyoak, K., Gentner, D., Kokinov, B. (eds.) The Analogical Mind: Perspectives from Cognitive Science, MIT Press, Cambridge (2001)
Holyoak, K., Thagard, P.: Analogical mapping by constraint satisfaction. Cognitive Science 13, 295–355 (1989)
Hommel, B.: How we do what we want: a neuro-cognitive perspective on human action planning. In: van Wezel, W., Jorna, R.J. (eds.) Planning: Its Aspects, Motivations, and Methods, John Wiley & Sons, New York (2003)

Hommel, B.: Event files: feature binding in and across perception and action. Trends Cogn. Sci. 8, 494–500 (2004)

Hommel, B., Musseler, J., Aschersleben, G., Prinz, W.: The theory of event coding (tec): a framework for perception and action planning. Behavioral and Brain Science 24(5), 849–878 (2001)

Hristova, E., Grinberg, M.: Context effects on judgment scales in the prisoner's dilemma game. In: Gif-sur-Yvette (ed.) Proceedings of the 1st European Conference on Cognitive Economics (2004)

Hristova, E., Grinberg, M.: Information acquisition in the iterated prisoner's dilemma game: An eye-tracking study. In: Proceedings of the 27th Annual Conference of the Cognitive Science Society (2005a)

Hristova, E., Grinberg, M.: Investigation of context effects in iterated prisoner's dilemma game. In: Dey, A.K., Kokinov, B., Leake, D.B., Turner, R. (eds.) CONTEXT 2005. LNCS (LNAI), vol. 3554, pp. 183–196. Springer, Heidelberg (2005b)

Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. J. Physiol. 195(1), 215–243 (1968)

Hume, D.: Philosophical essays concerning human understanding. Millar, London (1748)

Hummel, J., Holyoak, K.: Distributed representation of structure: A theory of analogical access and mapping. Psychological Review 104, 427–466 (1997)

Hurley, S.L.: Active perception and perceiving action: the shared circuits hypothesis. In: Gendler, T., Hawthorne, J. (eds.) Perceptual Experience, Oxford University Press, Oxford (2005)

Iacoboni, M.: Modulation of motor and premotor activity during imitation of target-directed actions. Cereb Cortex 12, 847–855 (2002)

Iacoboni, M.: Understanding others: Imitation, language, empathy. In: Hurley, S., Chater, N. (eds.) Perspectives on Imitation: From Cognitive Neuroscience to Social Science, MIT Press, Cambridge (2003)

Inagaki, K., Hatano, G.: Young children's spontaneous personification as analogy. Child Development 58, 1013–1020 (1987)

Ingvar, D.H.: Memory of the future: an essay on the temporal organization of conscious awareness. Hum. Neurobiol. 4, 127–136 (1985)

Isaac, R.M., Walker, J.M., Williams, A.W.: Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups. Journal of Public Economics 54, 1–36 (1994)

Isard, M., Blake, A.: Condensation – conditional density propagation for visual tracking. International Journal of Computer Vision 29(1), 5–28 (1998a)

Isard, M., Blake, A.: A mixed-state condensation tracker with automatic model-switching. In: Proceedings of the 6th International Conference of Computer Vision (ICCV), Bombay, India (1998b)

Isla, D.A., Blumberg, B.M.: Object persistence for synthetic characters. In: Castelfranchi, C., Johnson, W.L. (eds.) Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems. Part 1, Bologna, Italy, pp. 1356–1363. ACM Press, New York (2002)

Itti, L.: Models of Bottom-Up and Top-Down Visual Attention. PhD thesis, California Institute of Technology, Pasadena, California (2000)

Itti, L., Baldi, P.: Bayesian surprise attracts human attention. In: Advances in Neural Information Processing Systems, pp. 1–8. MIT Press, Cambridge (2006)

Itti, L., Koch, C.: Computational modeling of visual attention. Nature Reviews Neuroscience 2, 194–203 (2001a)

Itti, L., Koch, C.: Feature combination strategies for saliency-based visual attention systems. Journal of Electronic Imaging 10(1), 161–169 (2001b)

Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(11), 1254–1259 (1998)

Jacob, P., Jeannerod, M.: The motor theory of social cognition: a critique. Trends Cogn. Sci. 9(1), 21–25 (2005)

Jacobs, R.A., Jordan, M.I., Nowlan, S.J., Hinton, G.E.: Adaptive mixtures of local experts. Neural Computation 3, 79–87 (1991)

Jaeger, H.: Observable operator models for discrete stochastic time series. Neural Computation 12, 1371–1398 (2000)

Jaeger, H.: Online learning algorithms for observable operator models (2004)

James, M.R., Singh, S., Littman, M.L.: Planning with predictive state representations. In: Proceedings of the International Conference on Machine Learning and Applications, pp. 304–311 (2004)

James, W.: The Principles of Psychology. Dover Publications, New York (1890)

Jeannerod, M.: The representing brain: Neural correlates of motor intention and imagery. Behav. Brain Sci. 17, 187–245 (1994)

Jeannerod, M.: To act or not to act: Perspectives on the representation of actions. Quarterly Journal of Experimental Psychology 52A(1), 1–29 (1999)

Jeannerod, M.: Neural simulation of action: A unifying mechanism for motor cognition. NeuroImage 14, 103 (2001)

Jerne, N.K.: Towards a network theory of the immune system. Annals of Immunology 125, 373–389 (1974)

Johansson, B., Balkenius, C.: An experimental study of anticipation in simple robot navigation. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 365–378. Springer, Heidelberg (2007)

Johnson, M., Demiris, Y.: Hierarchies of coupled inverse and forward models for abstraction in robot action planning, recognition and imitation. In: Proceedings of the AISB 2005 Symposium on Imitation in Animals and Artifacts, Hertfordshire, UK, pp. 69–76 (2005a)

Johnson, M., Demiris, Y.: Perceptual perspective taking and action recognition. International Journal of Advanced Robotic Systems 2(4), 301–309 (2005b)

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

Kahneman, D., Miller, D.T.: Norm theory: comparing reality to its alternatives. Psychological Review 93, 136–153 (1986)

Kahneman, D., Tversky, A.: Variants of uncertainty. Cognition 11, 143–157 (1982)

Kakade, S., Dayan, P.: Dopamine: Generalization and bonuses. Neural Networks 15, 549–559 (2002)

Kalman, R.E.: A new approach to linear filtering and prediction problems. Transactions of the ASME-Journal of Basic Engineering 82(Series D), 35–45 (1960)

Kant, I.: Critique of Pure Reason (1781). Cambridge University Press, Cambridge (1998)

Kaplan, F., Hafner, V.: The challenges of joint attention. Interaction Studies 7(2) (2006)

Karniel, A., Inbar, G.F.: A model for learning human reaching movements. Biological Cybernetics 77, 173–183 (1997)

Kawato, M.: Internal models for motor control and trajectory planning. Current Opinion in Neurobiology 9, 718–727 (1999)

Kawato, M., Furukawa, K., Suzuki, R.: A hierarchical neural-network model for control and learning of voluntary movement. Biological Cybernetics 57, 169–185 (1987)

Kawato, M., Gomi, H.: A computational model of four regions of the cerebellum based on feedback-error learning. Biological Cybernetics 68(2), 95–103 (1992)

Keijzer, F.: Representation and behavior. MIT Press, Cambridge (2001)

Kennedy, J., Eberhart, R.C.: Swarm intelligence. Morgan Kaufmann Publishers Inc., San Francisco (2001)

Kerrigan, T.: Tom Kerrigan's simple chess program (2007)

Kersting, K., Van Otterlo, M., De Raedt, L.: Bellman goes relational. In: Proceedings of the Twenty-First International Conference on Machine Learning (ICML-2004), pp. 465–472 (2004)

Kilner, J.M., Friston, K.J., Frith, C.D.: Predictive coding: An account of the mirror neuron system. Cognitive Processing 8(3) (2007)

Kirsh, D., Maglio, P.: On distinguishing epistemic from pragmatic action. Cognitive Science 18, 513–549 (1994)

Kiryazov, K., Petkov, G., Grinberg, M., Kokinov, B., Balkenius, C.: The interplay of analogy-making with active vision and motor control in anticipatory robots. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 233–253. Springer, Heidelberg (2007)

Klein, R.M.: Inhibition of return. Trends in Cognitive Sciences 4(4), 138–147 (2000)

Kline, C.: Observation-based expectation generation and response for behaviour-based artificial creatures. Master's thesis, MIT Media Lab, Cambridge, MA, USA (1999)

Knill, D.C., Pouget, A.: The bayesian brain: The role of uncertainty in neural coding and computation. Trends in Neurosciences 27(12), 712–719 (2004)

Knoblich, G., Thornton, I.M., Grosjean, M., Shiffrar, M. (eds.): Human body perception from the inside out. Oxford University Press, New York (2005)

Koch, C., Ullman, S.: Shifts in selective visual attention: towards the underlying neural circuitry. Human Neurobiology 4(4), 219–227 (1985)

Koechlin, E., Summerfield, C.: An information theoretical approach to prefrontal executive function. Trends in Cognitive Sciences (2007)

Kokinov, B.: A hybrid model of reasoning by analogy. In: Holyoak, K., Barnden, J. (eds.) Advances in connectionist and neural computation theory: Vol. 2. Analogical connections, pp. 247–318. Ablex, Norwood (1994a)

Kokinov, B., Petrov, A.: Dynamic extension of episode representation in analogy-making in ambr. In: Proceedings of the 22nd Annual Conference of the Cognitive Science Society, Erlbaum, Hillsdale (2000)

Kokinov, B., Petrov, A.: Integration of memory and reasoning in analogy-making: The ambr model. In: Gentner, D., Holyoak, K., Kokinov, B. (eds.) The Analogical Mind: Perspectives from Cognitive Science, MIT Press, Cambridge (2001)

Kokinov, B.N.: The context-sensitive cognitive architecture dual. In: Proceedings of Cogsci XVI, Lawrence Erlbaum, Mahwah (1994b)

König, P., Krüger, N.: Symbols as self-emergent entities in an optimization process of feature extraction and predictions. Biological Cybernetics 94(4), 325–334 (2006)

Kopp, L.: Natural Vision for Artificial Systems: Active Vision and Thought. Lund University Cognitive Studies, vol. 100 (2003)

Kording, K.P., Wolpert, D.M.: Bayesian decision theory in sensorimotor control. Trends Cogn. Sci. 10, 319–326 (2006)

Kunde, W., Elsner, K., Kiesel, A.: No anticipation-no action: the role of anticipation in action and perception. Cogn. Process. 8(2), 71–78 (2007)

Kunde, W., Koch, I., Hoffmann, J.: Anticipated action effects affect the selection, initiation, and execution of actions. The Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology 57, 87–106 (2004)

Kuperstein, M.: Neural model of adaptive hand-eye coordination for single postures. Science 239, 1308–1311 (1988)

Kuperstein, M.: Infant neural controller for adaptive sensory-motor coordination. Neural Netw. 4(2), 131–145 (1991)

Kwok, C., Fox, D.: Map-based multiple model tracking of a moving object. In: Nardi, D., Riedmiller, M., Sammut, C., Santos-Victor, J. (eds.) RoboCup 2004. LNCS (LNAI), vol. 3276, pp. 18–33. Springer, Heidelberg (2005)

Lalev, E., Grinberg, M.: Backward vs. forward-oriented decision making in the iterated prisoner's dilemma: A comparison between two connectionist models. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 345–364. Springer, Heidelberg (2007)

Lamme, V.A.F., Roelfsema, P.R.: The distinct modes of vision offered by feedforward and recurrent processing. Trends in Neurosciences 23(11), 571–579 (2000)

Landau, S., Picault, S., Sigaud, O., Gérard, P.: Further comparison between ATNoSFERES and XCSM. In: Lanzi, P.L., Stolzmann, W., Wilson, S.W. (eds.) IWLCS 2003. LNCS (LNAI), vol. 2661, pp. 99–117. Springer, Heidelberg (2003)

Lanzi, P.L., Loiacono, D., Wilson, S.W., Goldberg, D.E.: Classifier prediction based on tile coding. In: GECCO 2006: Genetic and Evolutionary Computation Conference, pp. 1497–1504 (2006)

Lashley, K.S.: The problem of serial order in behavior. In: Jeffress, L.A. (ed.) Cerebral mechanisms in behavior, Wiley, New York (1951)

Lasseter, J.: Principles of traditional animation applied to 3d computer animation. ACM SIGGRAPH Computer Graphics (Proceedings of the 14th annual conference on Computer Graphics and Interactive Techniques) 21(4), 35–44 (1987)
Lazarus, R.S.: Emotion and adaptation. Oxford University Press, New York (1991)
Lee, D.K., Itti, L., Koch, C., Braun, J.: Attention activates winner-take-all competition among visual filters. Nature neuroscience 2(4), 375–381 (1999)
Levesque, H.J.: A logic of implicit and explicit belief. In: Proceedings of the Fourth National Conference on Artificial Intelligence (AAAI84), pp. 198–202. Morgan Kaufmann, Los Altos (1984)
Lewandowski, A.: Learning to anticipate a temporarily hidden moving object. Technical Report TR-2007-08, Österreichisches Forschungsinstitut für Artifiical Intelligence, Wien (2007)
Leydesdorff, L.: Anticipatory systems and the processing of meaning: a simulation study inspired by luhmann's theory of social systems. Journal of Artificial Societies and Social Simulation 10 (2005)
Leydesdorff, L., Dubois, D.: Anticipation in social systems. International Journal of Computing Anticipatory Systems 15, 203–216 (2004)
Lipson, A., Leyton-Brown, K.: Empirically evaluating multiagent reinforcement learning algorithms (Working Paper). Machine Learning (2006)
Littman, M., Sutton, R., Singh, S.: Predictive representations of state. In: Proc. NIPS-02, Vancouver (2001)
Loeb, G.E., Brown, I.E., Cheng, E.J.: A hierarchical foundation for models of sensorimotor control. Experimental Brain Research 126, 1–18 (1999)
Lorini, E., Castelfranchi, C.: The unexpected aspects of surprise. International Journal of Pattern Recognition and Artificial Intelligence 20(6), 817–833 (2006)
Lorini, E., Castelfranchi, C.: The cognitive structure of surprise: Looking for basic principles. Topoi: an International Review of Philosophy 26(1), 133–149 (2007)
Lorini, E., Piunti, M.: The benefits of surprise in dynamic environments: From theory to practice. In: Paiva, A., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 362–373. Springer, Heidelberg (2007)
Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)
Lowenstein, G., Lerner, J.S.: The role of affect in decision making. In: Davidson, R., Goldsmith, H., Scherer, K. (eds.) Handbook of Affective Science, pp. 619–642. Oxford University Press, Oxford (2003)
Loyall, B.: Believable agents: building interactive personalities. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA (1997)
Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Proceedings of Imaging Understanding Workshop, pp. 121–130 (1981)
Luhmann, N.: Soziale Systeme, pp. 148–179. Suhrkamp, Frankfurt/Main (1984)
Luria, A.R.: Higher Cortical Functions in Man. Basic Books, New York (1966)
Lyons, D.M., Arbib, M.A.: A formal model of computation for sensory-based robotics. IEEE Journal of Robotics and Automation 5(3), 280–293 (1989)

Macedo, L., Cardoso, A.: Modelling forms of surprise in an artificial agent. In: Proceedings of the 23rd conf. of the Cognitive Science Society, Edinburgh, UK (2001)

Macedo, L., Cardoso, A.: Exploration of unknown environments with motivational agents. In: 3rd Int. Conf. on Autonomous Agents and Multi Agent System(AAMAS04), NYC, USA (2004)

MacKay, D.G.: The organization of perception and action: a theory for language and other cognitive skills. Springer, New York (1987)

Maes, P.: Situated agents can have goals. In: Maes, P. (ed.) Designing Autonomous Agents, pp. 49–70. MIT Press, Cambridge (1990)

Maravita, A., Spence, C., Driver, J.: Multisensory integration and the body schema: Close to hand and within reach. Current Biology 13, 531–539 (2003)

Marsella, S., Gratch, J.: Modeling coping behavior in virtual humans: Don't worry, be happy. In: Proceedings of the second International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 313–320. ACM Press, New York (2003)

Martinez-Trujillo, J.C., Treue, S.: Feature-based attention increases the selectivity of population responses in primate visual cortex. Curr. Biol. 14, 744–751 (2004)

Martinho, C.: Emotivector: Affective Anticipatory Mechanism for Synthetic Characters. PhD thesis, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal (2007)

Martinho, C., Paiva, A.: Synthetic emotivectors. In: Proc. of Social Intelligence and Interaction in Animals, Robots and Agents, AISB (2005)

Martinho, C., Paiva, A.: Aini: achieving believability through anticipation. In: Proceedings of Nineteenth Conference on Computer Animation and Social Agents (2006a)

Martinho, C., Paiva, A.: Using anticipation to create believable behaviour. In: Proceedings of the 21st National Conference on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence Conference, vol. 1, pp. 175–180. AAAI Press, Menlo Park (2006b)

Martinho, C., Paiva, A., Gomes, M.: Emotions for a motion: Rapid development of believable pathematic agents in intelligent virtual environments. Applied Artificial Intelligence Journal 14(1), 33–68 (2000)

Mazor, E., Averbuch, A., Bar-Shalom, Y., Dayan, J.: Interacting multiple model methods in target tracking: a survey. IEEE Transactions on Aerospace and Electronic Systems 34(1), 103–123 (1998)

McCarthy, J., Hayes, P.J.: Some philosophical problems from the standpoint of artificial intelligence. In: Meltzer, B., Michie, D. (eds.) Machine Intelligence, vol. 4, pp. 463–502. Edinburgh University Press, Edinburgh (1969)

Mccauley, L.: Neural schemas: A mechanism for autonomous action selection and dynamic motivation. In: the 3rd WSES Neural Networks and Applications Conference (2002)

Mehta, B., Schaal, S.: Forward models in visuomotor control. Journal of Neurophysiology 88, 942–953 (2002)

Mel, B.W.: Vision-based robot motion planning. In: Neural networks for control, pp. 229–253 (1990)
Mel, B.W.: A connectionist model may shed light on neural mechanisms for visually guided reaching. Journal of cognitive neuroscience 3, 273–292 (1991)
Meller, B.: Anticipated emotions as guides to choice. Current Directions in Psychological Science 10(6), 210–214 (2001)
Meltzoff, A.N., Decety, J.: What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358(1431), 491–500 (2003)
Meltzoff, A.N., Moore, M.K.: Explaining facial imitation: A theoretical model. Early Development and Parenting 6, 179–192 (1997)
Métivier, M., Lattaud, C.: Anticipatory classifier system using behavioral sequences in non-markov environments. In: Lanzi, P.L., Stolzmann, W., Wilson, S.W. (eds.) IWLCS 2003. LNCS (LNAI), vol. 2661, pp. 143–162. Springer, Heidelberg (2003)
Meyer, W.U., Reisenzein, R., Schützwohl, A.: Towards a process analysis of emotions: The case of surprise. Motivation and Emotion 21, 251–274 (1997)
Miall, R.C., Wolpert, D.M.: Forward models for physiological motor control. Neural Networks 9(8), 1265–1279 (1996)
Miceli, M., Castelfranchi, C.: The role of evaluation in cognition and social interaction. In: Dautenhahn, K. (ed.) Human cognition and agent technology, Benjamins, Amsterdam (2000)
Miceli, M., Castelfranchi, C.: The mind and the future. the (negative) power of expectations. Theory & Psychology 12, 335–366 (2002)
Middleton, F.A., Strick, P.L.: Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. Brain Cogn. 42(2), 183–200 (2000)
Milinski, M., Semmann, D., Krambeck, H.-J.: Reputation helps solve the 'tragedy of the commons'. Nature 415, 424–426 (2002)
Miller, G.A., Galanter, E., Pribram, K.H.: Plans and the Structure of Behavior. Holt, Rinehart and Winston, New York (1960)
Millikan, R.G.: Varieties Of Meaning. MIT Press, Cambridge (2004)
Minsky, M.: The Society of Mind. Simon & Schuster, New York (1988)
Minut, S., Mahadevan, S.: A reinforcement learning model of selective visual attention. In: Müller, J.P., Andre, E., Sen, S., Frasson, C. (eds.) Proceedings of the Fifth International Conference on Autonomous Agents, Montreal, Canada, pp. 457–464. ACM Press, New York (2001)
Mitchell, M.: Analogy-making as perception: A computer model. MIT Press, Cambridge (1993)
Mitchinson, B., Pearson, M., Melhuish, C., Prescott, T.J.: A model of sensorimotor coordination in the rat whisker system. In: Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J.C.T., Marocco, D., Meyer, J.-A., Miglino, O., Parisi, D. (eds.) SAB 2006. LNCS (LNAI), vol. 4095, pp. 77–88. Springer, Heidelberg (2006)
Moffat, D., Frijda, N.: Where there's a will there's am agent. In: Wooldridge, M.J., Jennings, N.R. (eds.) ECAI 1994 and ATAL 1994. LNCS, vol. 890, pp. 245–260. Springer, Heidelberg (1995)

Mohan, V., Morasso, P.: A forward / inverse motor controller for cognitive robotics. In: Kollias, S., Stafylopatis, A., Duch, W., Oja, E. (eds.) ICANN 2006. LNCS, vol. 4131, pp. 602–611. Springer, Heidelberg (2006)

Moore, A.W., Atkeson, C.G.: Prioritized sweeping: Reinforcement learning with less data and less real time. Machine Learning 13, 103–130 (1993)

Morrison, C.T., Oates, T., King, G.W.: Grounding the unobservable in the observable: The role and representation of hidden state in concept formation and refinement. In: Working Notes of AAAI Spring Symposium Workshop: Learning Grounded Representations (2001)

Müller, H.J., Rabbit, P.M.A.: Reflexive orienting of visual attention: time course of activation and resistance to interruption. Journal of Experimental Psychology: Human Perception and Performance 15(2), 315–330 (1989)

Navalpakkam, V., Itti, L.: Towards a unified model for attention and recognition. In: Proc. Society for Neuroscience Annual Meeting, SFN'03 (Nov. 2003)

Navalpakkam, V., Itti, L.: Modeling the influence of task on attention. Vision Research 45(2), 205–231 (2005)

Navalpakkam, V., Itti, L.: Search goal tunes visual features optimally. Neuron 53(4), 605–617 (2007), Also see commentary / preview entitled "Paying Attention to Neurons with Discriminating Taste" by Pouget, A. and Bavelier, D., Neuron 53(4), 473–475 (2007)

Neisser, U.: Cognition and reality. W.H. Freeman, New York (1976)

Nolfi, S., Floreano, D.: Evolutionary Robotics. MIT Press, Cambridge (2000)

Nolfi, S., Tani, J.: Extracting regularities in space and time through a cascade of prediction networks: The case of a mobile robot navigating in a structured environment. Connection Science 11(2), 125–148 (1999)

Norman, D.A., Shallice, T.: Attention to action: Willed and automatic control of behaviour. In: Davidson, R.J., Schwartz, G.E., Shapiro, D. (eds.) Consciousness and Self-Regulation: Advances in Research and Theory, Plenum Press, New York (1986)

Ognibene, D., Rega, A., Baldassarre, G.: A model of reaching integrating continuous reinforcement learning, accumulator models, and direct inverse modeling. In: From Animals to Animats, vol. 9, pp. 381–393 (2006)

O'Regan, J.K., Noë, A.: A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24, 939–1031 (2001)

Ortony, A., Clore, G., Collins, A.: The cognitive structure of emotions. Cambridge University Press, New York (1988)

Ortony, A., Norman, D.A., Revelle, W.: Effective functioning: A three level model of affect, motivation, cognition, and behavior. In: Fellous, J.M., Arbib, A. (eds.) Who Needs Emotions? The Brain Meets the Machine, Oxford University Press, New York (2005)

Ortony, A., Partridge, D.: Surprisingness and expectation failure: What's the difference? In: Proceedings of the tenth International Joint Conference on Artificial Intelligence (IJCAI), pp. 106–108. Morgan Kaufmann, Los Altos (1987)

Oudeyer, P.-Y., Kaplan, F., Hafner, V.: Intrinsic motivation systems for autonomous mental development. IEEE Transactions on Evolutionary Computation 11(2), 265–286 (2007)

Oudeyer, P.-Y., Kaplan, F., Hafner, V.V., Whyte, A.: The playground experiment: Task-independent development of a curious robot. In: Bank, D., Meeden, L. (eds.) Proceedings of the AAAI Spring Symposium on Developmental Robotics, Stanford, California, pp. 42–47 (2005)

Oztop, E., Kawato, M., Arbib, M.: Mirror neurons and imitation: a computationally guided review. Neural Netw. 19(3), 254–271 (2006)

Oztop, E., Wolpert, D., Kawato, M.: Mental state inference using visual control parameters. Cognitive Brain Research 22, 129–151 (2005)

Pacherie, E.: The content of intentions. Mind and Language 15, 400–432 (2000)

Pacherie, E.: The sense of control and the sense of agency. Psyche 13, 1–30 (2007)

Paine, R.W., Tani, J.: How hierarchical control self-organizes in artificial adaptive systems. Adaptive Behavior 13(3), 211–225 (2005)

Paradiso, M.A., Nakayama, K.: Brightness perception and filling-in. Vision Research 31, 1221–1236 (1991)

Parrott, W.G., Schulkin, J.: Neuropsychology and the cognitive nature of the emotions. Cognition and Emotion 7, 43–59 (1993)

Pashler, H., Johnston, J.C., Ruthruff, E.: Attention and performance. Annual Review of Psychology 52, 629–651 (2001)

Pashler, H.E.: The Psychology of Attention. MIT Press, Cambridge (1998)

Pérez-Ortiz, J.A., Gers, F.A., Eck, D., Schmidhuber, J.: Kalman filters improve LSTM network performance in problems unsolvable by traditional recurrent nets. Neural Networks 16, 241–250 (2003)

Petkov, G., Kiryazov, K., Grinberg, M., Kokinov, B.: Modelling top-down perception and analogical transfer with single anticipatory mechanism. In: Proceedings of the Second European Cognitive Science Conference, Greece (2007)

Petta, P., Trappl, R. (eds.): Creating Personalities for Synthetic Actors. LNCS, vol. 1195. Springer, Heidelberg (1997)

Pezzulo, G.: Anticipation and future-oriented capabilites in natural and artificial cognition. In: Lungarella, M., Iida, F., Bongard, J.C., Pfeifer, R. (eds.) 50 Years of Aritficial Intelligence. LNCS (LNAI), vol. 4850, pp. 258–271. Springer, Heidelberg (2007)

Pezzulo, G.: Coordinating with the future: the anticipatory nature of representation. Minds and Machines 18(2), 179–225 (2008a)

Pezzulo, G.: A study of off-line uses of anticipation. In: Asada, M., Tani, J., Hallam, J., Meyer, J.-A. (eds.) Proceedings of SAB 2008. LNCS (LNAI), vol. 5040, pp. 372–382. Springer, Heidelberg (2008b)

Pezzulo, G., Baldassarre, G., Butz, M.V., Castelfranchi, C., Hoffmann2, J.: From actions to goals and vice-versa: Theoretical analysis and models of the ideomotor principle and TOTE. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 73–93. Springer, Heidelberg (2007)

Pezzulo, G., Calvi, G.: Dynamic computation and context effects in the hybrid architecture akira. In: Dey, A.K., Kokinov, B., Leake, D.B., Turner, R. (eds.) CONTEXT 2005. LNCS (LNAI), vol. 3554, pp. 368–381. Springer, Heidelberg (2005)

Pezzulo, G., Calvi, G.: A schema based model of the praying mantis. In: Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J.C.T., Marocco, D., Meyer, J.-A., Miglino, O., Parisi, D. (eds.) SAB 2006. LNCS (LNAI), vol. 4095, pp. 211–223. Springer, Heidelberg (2006a)

Pezzulo, G., Calvi, G.: Toward a perceptual symbol system. In: Proceedings of the Sixth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems. Lund University Cognitive Science Studies, vol. 118 (2006b)

Pezzulo, G., Calvi, G.: Designing modular architectures in the framework akira. Multiagent and Grid Systems 3(1), 65–86 (2007a)

Pezzulo, G., Calvi, G.: Schema-based design and the AKIRA schema language: An overview. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 128–152. Springer, Heidelberg (2007b)

Pezzulo, G., Calvi, G., Ognibene, D., Lalia, D.: Fuzzy-based schema mechanisms in akira. In: CIMCA '05: Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce Vol.-2, Washington, DC, USA, pp. 146–152. IEEE Computer Society Press, Los Alamitos (2005)

Pezzulo, G., Castelfranchi, C.: The symbol detachment problem. Cognitive Processing 8(2), 115–131 (2007)

Pfeifer, R., Gomez, G.: Interacting with the real world: Design principles for intelligent systems. Artificial Life and Robotics 9, 1–6 (2004)

Pfeifer, R., Scheier, C.: Understanding Intelligence. MIT Press, Cambridge (1999)

Piaget, J.: The Construction of Reality in the Child. Ballentine, New York (1954)

Piaget, J.: Biology and Knowledge. Edinburgh University Press, Edinburgh (1971)

Picard, R.: Affective Computing. MIT Press, Cambridge (1997)

Pierce, D.M., Kuipers, B.J.: Map learning with uninterpreted sensors and effectors. Artificial Intelligence 92, 169–227 (1997)

Piunti, M., Castelfranchi, C., Falcone, R.: Anticipatory coordination through action observation and behavior adaptation. In: Proceedings of AISB'07 - Artificial and Ambient Intelligence - Session on Mindful Environments (2007a)

Piunti, M., Castelfranchi, C., Falcone, R.: Expectations driven approach for situated, goal-directed agents. In: Dagli Oggetti agli Agenti: AI*IA/TABOO Joint Workshop (WOA 2007), Genova, Italy (2007b)

Piunti, M., Castelfranchi, C., Falcone, R.: Surprise as shortcut for Anticipation: clustering Mental States in reasoning. In: Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-07), Hyberadad, India (2007c)

Piunti, M., Gonçalves, J., Martinho, C.: Modeling expectations for affective agents. In: Cowie, R., de Rossis, F. (eds.) Proceedings of the Doctoral Consortium. Second International Conference on Affective Computing and Intelligent Interaction (ACII 07), pp. 86–93 (2007d)

Plutchik, R.: Emotion and evolution. In: Strongman, K.T. (ed.) International review of studies of emotions, vol. 1, John Wiley and Sons, Chichester (1991)

Poggio, T., Bizzi, E.: Generalization in vision and motor control. Nature 431, 768–774 (2004)

Pollack, J.B.: The induction of dynamical recognizers. Machine Learning 7, 227–252 (1990)

Pollock, J.L.: Taking perception seriously. In: Johnson, W.L., Hayes-Roth, B. (eds.) Proceedings of the 1st International Conference on Autonomous Agents (Agents'97), pp. 526–527. ACM Press, New York (1997)

Pomerleau, D.: Alvinn: An autonomous land vehicle in a neural network. In: Advances in Neural Information Processing Systems, vol. 1 (1989)

Pomerleau, D.: Reliability estimation for neural network based autonomous driving. Robotics and Autonomous Systems 12, 113 (1994)

Posner, M.I.: Orienting of attention. Quaterly Journal of Experimental Psychology 32, 3–25 (1980)

Pouget, A., Dayan, P., Zemel, R.: Information processing with population codes. Nature Reviews Neuroscience 1(2), 125–132 (2000)

Powers, W.T.: Behavior: The Control of Perception. Aldine de Gruyter, New York (1973)

Prinz, W.: A common coding approach to perception and action. In: Neumann, O., Prinz, W. (eds.) Relationships between perception and action, pp. 167–201. Springer, Berlin (1990)

Prinz, W.: Experimental approaches to action. In: Roessler, J., Eilan, N. (eds.) Agency and Self-awareness, pp. 175–187. Oxford Univ. Press, Oxford (2003)

Prinz, W.: An ideomotor approach to imitation. In: Hurley, S., Chater, N. (eds.) Perspectives on imitation: From neuroscience to social science, vol. 1, pp. 141–156. MIT Press, Cambridge (2005)

Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley and Sons, New York (1994)

Rainer, G., Rao, S.C., Miller, E.K.: Prospective coding for objects in primate prefrontal cortex. The Journal of Neuroscience 19(13), 5493–5505 (1999)

Rao, A.S., Georgeff, M.: An abstract architecture for rational agents. In: Rich, C., Swartout, W., Nebel, B. (eds.) Proceedings of the 3th International Conference on Principles of Knowledge Representation and Reasoning (KR'92), pp. 439–449. Morgan Kaufmann, San Francisco (1992)

Rao, A.S., Georgeff, M.P.: BDI agents: From theory to practice. In: Proc. of the 1st conf. on MAS, ICMAS95 (1995)

Rao, R.P., Ballard, D.H.: Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2(1), 79–87 (1999)

Rao, R.P., Zelinsky, G.J., Hayhoe, M.M., Ballard, D.H.: Eye movements in iconic visual search. Vision Research 42(11), 1447–1463 (2002)

Rao, R.P.N., Ballard, D.H.: An active vision architecture based on iconic representations. Artificial Intelligence 78(1-2), 461–505 (1995)

Rao, R.P.N., Ballard, D.H.: Dynamic model of visual recognition predicts neural response properties in the visual cortex. Neural Computation 9, 721–763 (1997)

Rapoport, A., Chammah, A.: Prisoner's Dilemma: A Study in Conflict and Cooperation. University of Michigan Press, Ann Arbor (1965)

Redgrave, P., Gurney, K.: The short-latency dopamine signal: a role in discovering novel actions? Nature Reviews Neuroscience 7, 967–975 (2006)

Redgrave, P., Prescott, T.J., Gurney, K.: The basal ganglia: a vertebrate solution to the selection problem? Neuroscience 89, 1009–1023 (1999)

Reilly, S., Bates, J.: Building emotional agents. Technical Report CMU-CS-92-143, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA (1992)

Reisenzein, R.: The subjective experience of surprise. In: Bless, H., Forgas, J. (eds.) The message within: The role of subjective experience in social cognition and behavior, Psychology Press, Philadelphia (2000)

Reisfeld, D., Wolfson, H., Yeshurun, Y.: Context-free attentional operators: The generalized symmetry transform. International Journal of Computer Vision 14(2), 119–130 (1995)

Rescorla, R.A., Wagner, A.R.: A theory of Pavlovian conditioning: Variation in the effectiveness of reinforcement and non reinforcement. In: Black, A.H., Prokasy, W.F. (eds.) Classical Conditioning II: Current Research and Theory, pp. 64–99. Appleton Century Crofts, New York (1972)

Richard, R., van der Pligt, J., de Varies, N.: Anticipated regret and time perspective: changing sexual risk taking behaviour. Journal of Behavioural Decision Making 9, 185–199 (1996)

Riesenhuber, M., Poggio, T.: Hierarchical models of object recognition in cortex. Nature Neuroscience 2(11), 1019–1025 (1999)

Rizzolatti, G., Arbib, M.A.: Language within our grasp. Trends in Neurosciences 21(5), 188–194 (1998)

Rizzolatti, G., Craighero, L.: The mirror-neuron system. Annual Review of Neuroscience 27, 169–192 (2004)

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L.: Premotor cortex and the recognition of motor actions. Cognitive Brain Research 3 (1996)

Rizzolatti, G., Fogassi, L., Gallese, V.: Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews Neuroscience 2, 661–670 (2001)

Roitblat, H.L.: Codes and coding processes in pigeon short-term memory. Anim. Learn. Behav. 8, 341–351 (1980)

Rojas, R., Förster, A.G.: Holonomic control of a robot with an omnidirectional drive. KI - Künstliche Intelligenz 20(2) (2006)

Rosen, R.: Anticipatory Systems. Pergamon Press, Oxford (1985)

Rosenbaum, D.A., Loukopoulos, L.D., Meulenbroek, R.G.J., Vaughan, J., Engelbrecht, S.E.: Planning reaches by evaluating stored postures. Psychological Review 102(1), 28–67 (1995)

Rosenblueth, A., Wiener, N., Bigelow, J.: Behavior, purpose and teleology. Philosophy of Science 10(1), 18–24 (1943)

Ross, D.A., Osindero, S., Zemel, R.S.: Combining discriminative features to infer complex trajectories. In: Proceedings of the 23rd international conference on Machine learning (ICML), Pittsburgh, Pennsylvania, June 25-29 (2006)

Rott, H.: Conditionals and theory change: Revisions, expansions and additions. Synthese 81, 91–113 (1989)

Roy, D.: Semiotic schemas: a framework for grounding language in action and perception. Artificial Intelligence 167(1-2), 170–205 (2005)

Roy, D., Hsiao, K.-y., Mavridis, N., Gorniak, P.: Ripley, hand me the cup: Sensorimotor representations for grounding word meaning. In: Int. Conf. of Automatic Speech Recognition and Understanding (2006)

Russell, S., Norvig, P.: Artificial Intelligence - A Modern Approach. Prentice-Hall International Limited, London, UK (1995)

Savage, L.J.: The Foundations of Statistics. Dover, Mineola (1954)

Scassellati, B.: Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot. In: Nehaniv, C.L. (ed.) CMAA 1998. LNCS (LNAI), vol. 1562, pp. 176–195. Springer, Heidelberg (1999)

Schacter, D.L., Addis, D.R., Buckner, R.L.: Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. 8(9), 657–661 (2007)

Schank, R.C., Abelson, R.P.: Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures. L. Erlbaum, Hillsdale (1977)

Schenck, W., Möller, R.: Training and application of a visual forward model for a robot camera head. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 153–169. Springer, Heidelberg (2007)

Scherer, K.: Psychological models of emotion. In: Borod, J. (ed.) The neuropsychology of emotion, pp. 137–162. Oxford University Press, New York (2000)

Scheutz, M.: Agents with or without emotions? In: Proceedings of FLAIRS 02, pp. 89–94. AAAI Press, Menlo Park (2002)

Scheutz, M.: An artificial life approach to the study of basic emotions. In: Proceedings of CogSci 2004 (2004a)

Scheutz, M.: How to determine the utility of emotions. In: Proc. of AAAI Spring Symposium 2004 (2004b)

Scheutz, M., Sloman, A.: Affect and agent control: Experiments with simple affective states. In: Proceedings of IAT-01 (2001)

Schlesinger, M., Barto, A.: Optimal control methods for simulating the perception of causality in young infants. In: Proceedings of the Twenty First Annual Conference of the Cognitive Science Society, pp. 625–630. Lawrence Erlbaum, Mahwah (1999)

Schmidhuber, J.: Making the world differentiable: On using fully recurrent self-supervised neural networks for dynamic reinforcement learning and planning in non-stationary environments. Technical Report FKI-126-90, Institut fuer Informatik, Technische Universitaet Muenchen (1990a)

Schmidhuber, J.: An on-line algorithm for dynamic reinforcement learning and planning in reactive environments. In: Proc. IEEE/INNS International Joint Conference on Neural Networks, vol. 2, pp. 253–258 (1990b)

Schmidhuber, J.: Curious model-building control systems. In: Proc. International Joint Conference on Neural Networks, vol. 2, pp. 1458–1463 (1991a)

Schmidhuber, J.: Learning to generate sub-goals for action sequences. In: Kohonen, T., Makisara, K., Simula, O., Kangas, J. (eds.) Artificial Neural Networks, pp. 967–972. North-Holland, Amsterdam (1991b)

Schmidhuber, J.: A possibility for implementing curiosity and boredom in model-building neural controllers. In: Meyer, J.A., Wilson, S.W. (eds.) Proceedings of the International Conference on Simulation of Adaptive Behavior, pp. 222–227. MIT Press, Cambridge (1991c)

Schmidhuber, J.: Learning complex extended sequences using the principle of history compression. Neural Computation 4(2), 234–242 (1992a)

Schmidhuber, J.: Learning unambiguous reduced sequence descriptions. In: Moody, J.E., Hanson, S.J., Lippman, R.P. (eds.) Neural Information Processing Systems, vol. 4, pp. 291–298. Morgan Kaufmann, San Francisco (1992b)

Schmidhuber, J.: Exploring the predictable. In: Ghosh, S., Tsutsui, S. (eds.) Advances in Evolutionary Computing, pp. 579–612. Springer, Heidelberg (2002)

Schmidhuber, J.: Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity & creativity. In: Corruble, V., Takeda, M., Suzuki, E. (eds.) DS 2007. LNCS (LNAI), vol. 4755, pp. 26–38. Springer, Heidelberg (2007)

Schmidhuber, J., Huber, R.: Learning to generate artificial fovea trajectories for target detection. International Journal of Neural Systems 2, 135–141 (1991)

Schmidt, R.A.: A schema theory of discrete motor skill-learning. Psychological Review 82(4), 225–260 (1975)

Schomaker, L.: Anticipation in cybernetic systems: A case against mindless anti-representationalism. In: IEEE lnternational Conference on Systems, Man and Cybernetics (2004)

Schubotz, R.I.: Prediction of external events with our motor system: towards a new framework. Trends in Cognitive Sciences 11(5), 211–218 (2007)

Schultz, W.: Predictive reward signal of dopamine neurons. Journal of Neurophysiology 80, 1–27 (1998)

Schultz, W., Dayan, P., Montague, P.R.: A neural substrate of prediction and reward. Science 275, 1593–1599 (1997)

Schwartz, A.B., Moran, D.W., Reina, G.A.: Differential representation of perception and action in the frontal cortex. Science 303, 380–383 (2004)

Schweighofer, N., Arbib, M.A., Kawato, M.: Role of the cerebellum in reaching movements in humans. I. Distributed inverse dynamics control. European Journal of Neuroscience 10, 86–94 (1998a)

Schweighofer, N., Spoelstra, J., Arbib, M.A., Kawato, M.: Role of the cerebellum in reaching movements in humans. II. A neural model of the intermediate cerebellum. European Journal of Neuroscience 10(1), 95–105 (1998b)

Shadmehr, R., Wise, S.P. (eds.): The Computational Neurobiology of Reaching and Pointing. The MIT Press, Cambridge (2005)

Shanahan, M.P.: Cognition, action selection, and inner rehearsal. In: Proceedings IJCAI 2005 Workshop on Modelling Natural Action Selection, pp. 92–99 (2005)

Shapiro, D.C., Schmidt, R.A.: The schema theory: Recent evidence & developmental implications. In: Kelso, J.A.S., Clark, J.E. (eds.) The development of movement control and coordination, Wiley, New York (1982)

Shi, H., Yang, Y.: A computational model of visual attention based on saliency maps. Applied Mathematics and Computation 188, 1671–1677 (2007)

Shibata, K., Nishino, T., Okabe, Y.: Active perception based on reinforcement learning. In: Proc. of WCNN'95, Washington DC, pp. 170–173 (1995)

Shibata, T., Schaal, S.: Biomimetic gaze stabilization based on feedback-error-learning with nonparametric regression networks. Neural Networks 14, 201–216 (2001)

Shibata, T., Vijayakumar, S., Conradt, J., Schaal, S.: Humanoid oculomotor control based on concepts of computational neuroscience. In: Proceedings of Humanoids 2001, Second IEEE-RAS Intl. Conf. on Humanoid Robots (2001)

Shidara, M., Kawano, K., Gomi, H., Kawato, M.: Inverse-dynamics model eye movement control by purkinje cells in the cerebellum. Nature 365(6441), 50–52 (1993)

Siagian, C., Itti, L.: Rapid biologically-inspired scene classification using features shared with visual attention. IEEE Trans Pattern Anal Mach Intell 29(2), 300–312 (2007)

Sigaud, O., Wilson, S.W.: Learning classifier systems: a survey. Soft Computing – A Fusion of Foundations, Methodologies and Applications 11(11), 1065–1078 (2007)

Simon, H.A.: Motivational and emotional controls of cognition. Psychological Review 74, 29–39 (1967)

Simons, D.J., Chabris, C.F.: Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception 28, 1059–1074 (1999)

Simsek, Ö., Barto, A.G.: Using relative novelty to identify useful temporal abstractions in reinforcement learning. In: Proceedings of the Twenty-First International Conference on Machine Learning (ICML-2004), pp. 751–758 (2004)

Singh, S., Barto, A.G., Chentanez, N.: Intrinsically motivated reinforcement learning. In: Saul, L.K., Weiss, Y., Bottou, L. (eds.) Advances in Neural Information Processing Systems, vol. 17, pp. 1281–1288 (2005)

Singh, V.K., Maji, S., Mukerjee, A.: Confidence based updation of motion conspicuity in dynamic scenes. In: The 3rd Canadian Conference on Computer and Robot Vision, pp. 13–13 (2006)

Sloman, A.: Orthogonal recombinable competences acquired by altricial species. Technical report, University of Birmingham (2006)

Smith, B.C.: On the Origin of Objects. MIT Press, Cambridge (1996)

Solomon, J.A.: The effect of spatial cues on visual sensitivity. Vision Research 44(12), 1209–1216 (2004)

Sommer, W., Leuthold, H., Matt, J.: The expectancies that govern the P300 amplitude are mostly automatic and unconscious. Behavioral and Brain Sciences 21, 149–150 (1998)

Spratling, M.W., Johnson, M.H.: A Feedback Model of Visual Attention. Journal of Cognitive Neuroscience 16(2), 219–237 (2004)

Stalnaker, R.: A theory of conditionals. In: Harper, W.L., Stalnaker, R., Pearce, G. (eds.) IFs: Conditionals, Belief, Decision, Chance, and Time, pp. 98–122. D. Reidel, Dordrecht (1981)

Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Fort Colins, CO, vol. 2 (1999)

Stephan, V., Gross, H.-M.: Visuomotor anticipation - a powerful approach to behavior-driven perception. Künstliche Intelligenz 17(2), 12–17 (2003)

Sternberg, S., Monsell, S., Knoll, R.L., Wright, C.E.: The latency and duration of rapid movement sequences: comparisons of speech and typewriting. In: Stelmach, G.E. (ed.) Information Processing in Motor Control and Learning, pp. 117–152. Academic Press, New York (1978)

Stolzmann, W.: Anticipatory classifier systems. In: Genetic Programming 1998: Proceedings of the Third Annual Conference, pp. 658–664 (1998)

Stolzmann, W.: An introduction to anticipatory classifier systems. In: Lanzi, P.L., Stolzmann, W., Wilson, S.W. (eds.) IWLCS 1999. LNCS (LNAI), vol. 1813, pp. 175–194. Springer, Heidelberg (2000)

Stringer, S.M., Rolls, E.T.: Hierarchical dynamical models of motor function. Neurocomput. 70(4-6), 975–990 (2007)

Strongman, K.: The Psychology of Emotion, 4th edn. John Wiley and Sons Ltd., New York (1996)

Sutton, R.S.: Learning to predict by the methods of temporal differences. Machine Learning 3, 9–44 (1988)

Sutton, R.S.: Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In: Proceedings of the Seventh International Conference on Machine Learning, pp. 216–224 (1990)

Sutton, R.S.: Integrated modeling and control based on reinforcement learning and dynamic programming. In: Touretzky, D.S. (ed.) Advances in Neural Information Processing Systems, vol. 3, pp. 471–478 (1991)

Sutton, R.S., Barto, A.G.: An adaptive network that constructs and uses an internal model of its world. Cognition and Brain Theory 4, 217–246 (1981a)

Sutton, R.S., Barto, A.G.: Toward a modern theory of adaptive networks: Expectation and prediction. Psychological Review 88, 135–170 (1981b)

Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge (1998)

Swarup, S., Gasser, L.: The role of anticipation in the emergence of language. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 35–56. Springer, Heidelberg (2007)

Taiji, M., Ikegami, T.: Dynamics of internal models in game players. Physica D 134, 253–266 (1999)

Tani, J.: Model-based learning for mobile robot navigation from the dynamical systems perspective. IEEE Transactions on Systems, Man, and Cybernetics 26, 421–436 (1996)

Tani, J.: Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. Neural Netw. 16(1), 11–23 (2003)

Tani, J., Ito, M., Sugita, Y.: Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using rnnpb. Neural Netw. 17(8-9), 1273–1289 (2004)

Tani, J., Nolfi, S.: Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. Neural Netw. 12(7-8), 1131–1141 (1999)

Thangarajah, J., Padgham, L., Harland, J.: Representation and reasoning for goals in BDI agents. In: Proc. of the 25th Australian Computer Science Conf., ACS2002(2002)

Thomas, F., Johnston, O.: The Illusion of Life. Hyperion Press, New York (1994)

Thomas, N.J.T.: Are theories of imagery theories of imagination? Cognitive Science 23(2), 207–245 (1999)

Toby, J., Cosmides, L.: The past explains the present: Emotional adaptations and the structure of ancestral environment. Ethology and Sociobiology 11, 375–424 (1990)

Todorov, E.: Optimality principles in sensorimotor control. Nature Review Neuroscience 7(9), 907–915 (2004)

Tolman, E.C.: Purposive Behavior in Animals and Men. Appleton-Century-Crofts, New York (1932)

Torralba, A.: Contextual priming for object detection. International Journal of Computer Vision 53(2), 169–191 (2003)

Treptow, A., Masselli, A., Zell, A.: Real-time object tracking for soccer-robots without color information. In: European Conference on Mobile Robots (ECMR), Radziejowice, Poland (2003)

Treue, S., Martinez-Trujillo, J.C.: Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399(6736), 575–579 (1999)

Tsoukalas, L.H.: Neurofuzzy approaches to anticipation: a new paradigm for intelligent systems. IEEE Transactions on Systems, Man, and Cybernetics, Part B 28(4), 573–582 (1998)

Tulving, E.: Elements of episodic memory. Clarendon Press, Oxford (1983)

Tulving, E.: Episodic memory and autonoesis: Uniquely human? In: Terrace, H.S., Metcalfe, J. (eds.) The Missing Link in Cognition: Origins of Self-Reflective Consciousness, Oxford University Press, Oxford (2005)

Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keyers, C., Rizzolatti, G.: I know what you are doing. a neurophysiological study. Neuron 31(1), 155–165 (2001)

van Dartel, M.F.: Situated Representation. PhD thesis, Maastricht University Press, Maastricht, The Netherlands (2005)

Vaughan, R.T., Zuluaga, M.: Use your illusion: Sensorimotor self-simulation allows complex agents to plan with incomplete self-knowledge. In: Proceedings of the International Conference on Simulation of Adaptive Behaviour (SAB), Rome, Italy (2006)

Vickery, T.J., King, L.-W., Jiang, Y.: Setting up the target template in visual search. Journal of Vision 5, 81–92 (2005)

Vijayakumar, S., Conradt, J., Shibata, T., Schaal, S.: Overt visual attention for a humanoid robot. In: Proceedings of the International Conference on Intelligent Robots and Systems (IROS-01), vol. 4, pp. 2332–2337 (2001)

Vijayakumar, S., D'souza, A., Schaal, S.: Incremental online learning in high dimensions. Neural Comput. 17(12), 2602–2634 (2005)

von Hofsten, C., Rosander, K.: Development of smooth pursuit tracking in young infants. Vision Research 37(13), 1799–1810 (1997)

Walther, D., Koch, C.: Modeling attention to salient proto-objects. Neural Networks 19, 1395–1407 (2006)

Weber, C., Wermter, S., Elshaw, M.: A hybrid generative and predictive model of the motor cortex. Neural Networks 19(4), 339–353 (2006)

Weitzenfeld, A., Peguero, O., Gutiérrez, S.: NSL/ASL: Distributed simulation of modular neural networks. In: Cairó, O., Cantú, F.J. (eds.) MICAI 2000. LNCS, vol. 1793, pp. 326–337. Springer, Heidelberg (2000)

Welch, G., Bishop, G.: An introduction to the Kalman filter. Technical Report TR 95-041, UNC-Chapel Hill (2006)

Westin, C.F., Westelius, C.J., Knutsson, H., Granlund, G.: Attention control for robot vision. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR '96, pp. 726–733 (1996)

Whitehead, S., Lin, L.: Reinforcement learning of non-markov decision processes. Artificial Intelligence 7(1-2), 271–306 (1995)

Whitney, D.E.: Resolved motion rate control of manipulators and human prostheses. IEEE Transactions on Man-Machine Systems 10, 47–53 (1969)

Wiener, S.I., Berthoz, A., Zugaro, M.B.: Multisensory processing in the elaboration of place and head direction responses by limbic system neurons. Cognitive Brain Research 14, 75–90 (2002)

Wiering, M., Schmidhuber, J.: Hq-learning. Adaptive Behavior 6(2), 219–246 (1997)

Wierstra, D., Foerster, A., Peters, J., Schmidhuber, J.: Solving deep memory pOMDPs with recurrent policy gradients. In: de Sá, J.M., Alexandre, L.A., Duch, W., Mandic, D. (eds.) ICANN 2007. LNCS, vol. 4668, pp. 697–706. Springer, Heidelberg (2007)

Wilson, S.W.: Classifier fitness based on accuracy. Evolutionary Computation 3(2), 149–175 (1995)

Wilson, S.W.: Classifiers that approximate functions. Natural Computing 1, 211–234 (2002)

Witkowski, M.: Towards a four factor theory of anticipatory learning. In: Butz, M.V., Sigaud, O., Gérard, P. (eds.) Anticipatory Behavior in Adaptive Learning Systems. LNCS (LNAI), vol. 2684, pp. 66–85. Springer, Heidelberg (2003)

Wolfe, B., James, M.R., Singh, S.: Learning predictive state representations in dynamical systems without reset. In: ICML '05: Proceedings of the 22nd international conference on Machine learning, Bonn, Germany, pp. 980–987. ACM Press, New York (2005)

Wolfe, B., Singh, S.: Predictive state representations with options. In: ICML '06: Proceedings of the 23rd international conference on Machine learning, pp. 1025–1032. ACM Press, New York (2006)

Wolpert, D.M., Doya, K., Kawato, M.: A unifying computational framework for motor control and social interaction. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358(1431), 593–602 (2003)

Wolpert, D.M., Flanagan, J.R.: Motor prediction. Current Biology 11, R729–732 (2001)

Wolpert, D.M., Flanagan, J.R.: Sensorimotor learning. In: Arbib, M. (ed.) The Handbook of Brain Theory and Neural Networks, 2nd edn., pp. 1020–1023. MIT Press, Cambridge (2003)

Wolpert, D.M., Ghahramani, Z., Flanagan, J.R.: Perspectives and problems in motor learning. Trends in Cognitive Sciences 5(11), 487–494 (2001)

Wolpert, D.M., Gharamani, Z., Jordan, M.I.: An internal model for sensorimotor integration. Science 269, 1179–1182 (1995)

Wolpert, D.M., Kawato, M.: Multiple paired forward and inverse models for motor control. Neural Networks 11, 1317–1329 (1998)

Wolpert, D.M., Miall, R.C., Kawato, M.: Internal models in the cerebellum. Trends in Cognitive Science 2 (1998)

Wooldridge, M.: An Introduction to Multiagent Systems. John Wiley & Sons, Chichester (2002)

Yan, F., Christmas, W., Kittler, J.: A tennis ball tracking algorithm for automatic annotation of tennis match. In: 16th British Machine Vision Conference, Oxford Brookes University, Oxford (2005)

Yarbus, A.: Eye Movements and Vision. Plenum Press, New York (1967)

Young, P.: Motivation and Emotion. John Wiley and Sons Ltd., New York (1961)

Yuille, A., Kersten, D.: Vision as bayesian inference: analysis by synthesis? Probabilistic models of cognition 10(7), 301–308 (2006)

Zappacosta, S., Nolfi, S., Baldassarre, G.: A testbed for neural-network models capable of integrating information in time. In: Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (eds.) ABiALS 2006. LNCS (LNAI), vol. 4520, pp. 189–217. Springer, Heidelberg (2007)

Zhang, K.: Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. Journal of Neuroscience 16(6), 2112–2126 (1996)

Zhumatiy, V., Gomez, F., Hutter, M., Schmidhuber, J.: Metric state space reinforcement learning for a vision-capable mobile robot. In: Proc. of the Int'l Conf. on Intelligent Autonomous Systems, IAS-06 (2006)

Ziemke, T., Jirenhed, D.-A., Hesslow, G.: Internal simulation of perception: a minimal neuro-robotic model. Neurocomputing 68, 85–104 (2005)