UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Emotion-based Mechanisms for Decision Making in Autonomous Agents

Rodrigo Martins de Matos Ventura
(Mestre)

Dissertação para obtenção do Grau de Doutor em
Engenharia Electrotécnica e de Computadores

Orientador: Doutor Carlos Alberto Pinto-Ferreira

Júri

Presidente: Reitor da Universidade Técnica de Lisboa

Vogais: Doutor Helder Manuel Ferreira Coelho
Doutor Carlos Alberto Pinto-Ferreira
Doutora Ana Maria Severino de Almeida e Paiva
Doutor Paolo Petta
Doutor Luís Manuel Marques Custódio

Setembro de 2007

Título: Mecanismos Baseados em Emoções para a Tomada de Decisão em Agentes Autónomos

Nome: Rodrigo Martins de Matos Ventura

Curso de Doutoramento em: Engenharia Electrotécnica e de Computadores

Orientador: Prof. Carlos Alberto Pinto-Ferreira

Provas concluídas em:

Resumo: Tomando como inspiração biológica a proposta de António Damásio de que os mecanismos das emoções no cérebro são essenciais para uma apropriada tomada de decisão, esta tese apresenta um modelo conceptual para um agente autónomo baseado num paradigma de dupla representação. Estímulos são representados sob duas perspectivas distintas, induzindo dois esquemas de representação com propriedades diferentes. As consequências deste modelo são exploradas de várias formas. Primeiro, a aplicabilidade deste modelo à antecipação, e à formulação de modelos causais sobre o mundo é explorada. De seguida é apresentado um modelo formal, onde consequências teóricas são derivadas, inicialmente de um ponto de vista probabilístico, seguido por uma abordagem baseada na assunção de que as representações acima mencionadas vivem em espaços métricos. Seguindo esta última abordagem, é proposto um algoritmo para adaptar a métrica de um desses espaços, tal como fornecer indicações para a melhoria dessa representação, com vista à criação de novas características. A formulação deste algoritmo é baseado em técnicas de *Multidimensional Scaling*. Resultados utilizando um mundo sintético corroboram as hipóteses levantadas na proposta deste algoritmo.

Palavras-chave: Emoções, Agentes, Decisão, Inteligência Artificial, Neurociência.

Title: Emotion-based Mechanisms for Decision Making in Autonomous Agents

Abstract: Taking as biological inspiration the António Damásio proposal that the brain emotion mechanisms are essential for appropriate decision-making, this thesis presents a conceptual model for an autonomous agent based on a double-representation paradigm. Stimuli is represented under two distinct perspectives, thus inducing two representation schemata with different properties. The consequences of this model are explored in various forms. First, the applicability of the model to anticipation, and to the formulation of causal models about the world are explored. And second, a formal approach is presented, where theoretical consequences are derived, first from a probabilistic standpoint, followed by an approach based on the assumption that the above-mentioned representations live in metric spaces. Following this latter approach, an algorithm is proposed to adapt the metric for one of the spaces, as well as to provide a guidance for the improvement of that representation, aiming at the creation of new features. The formulation of this algorithm is based on Multidimensional Scaling techniques. Results employing a synthetic world corroborate the hypotheses raised by the proposal of the algorithm.

Key-words: Emotions, Agents, Decision, Artificial Intelligence, Neuroscience.

— Posible, pero no interesante — respondió Lönnrot —. Usted replicará que la realidad no tiene la menor obligación de ser interesante. Yo le replicaré que la realidad puede prescindir de esa obligación, pero no las hipótesis.

Jorge Luis Borges

Acknowledgments

I would like to express my warmest gratitude to my supervisor Prof. Carlos Pinto-Ferreira for many enlightening conversations that helped me walking my path, as well as removing many stones I found on the way. Many thanks to Prof. Luís Custódio and Dr. Sandra Gadanho for the discussions we had, giving me irreplaceable feedback regarding this research. A special word to Prof. Paolo Petta for his persistent encouragement and warm friendship. Finally, this work would not be possible without the scientific environment of excellence at the Institute for Systems and Robotics (ISR).

I dedicate this thesis to my loving wife, Dina Silva, for all the encouragement during this endeavor.

This work was partially supported by Fundação para a Ciência e a Tecnologia (ISR/IST plurianual funding) through the POS_Conhecimento Program that includes FEDER funds.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

People usually say "don't get emotional over this matter" in a manner of warning that emotions threaten to get in the way of the sound analysis of a situation. In fact, Western culture has been dominated by a Cartesian view of intelligence as dispassionate reasoning, happening in the realm of a disembodied mind. And on the contrary, emotions are viewed as something pertaining to the body, hence outside of the realm of reason. Intelligence and emotions are thus two things living in different, contradictory worlds.

This perspective has dominated the way Artificial Intelligence (A.I.) has been progressing towards the goal of machine intelligence. This point of view became more evident when the field became dominated by research on logic reasoning. One notable exception can be found in the writings of the Nobel laureate Herbert Simon, one of the founders of the field. He was perhaps the first one to propose in 1967 that emotions are an essential mechanism of intelligent machines [172]. According to Simon, emotions provide an interrupt mechanism, triggered by relevant real-time events, that among other things, regulates the prioritization of goals on a multiple-goal systems. However, most of the later research in A.I. disregarded any similar modeling of emotional phenomena.

This was the state of affairs in A.I. until the neuroscientist António Damásio brought forward a controversial proposition, in which emotion mechanisms play a crucial role in many apparently dispassionate human behaviors [55]. More than in the case of reasoning, where the brain can be loosely seen as the "hardware" on which the mind runs, the body proper holds a closer relationship with emotions. Elicitation of emotions provokes actual physiological changes, often externally visible (and measurable). Hence, reclaiming the role of emotions is also a way of reclaiming the role of the body in mental activity. In fact, Damásio claims that the body proper is a decisive actor in most of our intelligent activity. The reason for this is a closer relationship among the brain regions responsible for reasoning and emoting than previously thought. Its effects are exerted mostly in a covert, unconscious fashion, and therefore unnoticeable through introspection. This proposition thus shakes the foundations of the Cartesian view of a rational mind apart from a physiological body. Extensive experimental research, conducted by Damásio and colleagues, has been corroborating their theses. For instance, patients with lesions on the pathways connecting the "reasoning" and the "emotional" brains, exhibit severe impairments in their ability to perform simple common real-life decisions. Although I.Q. tests have shown no measurable deficits, they became unable to deal appropriately with many daily life matters.

This thesis takes this proposal as a biological inspiration for the formulation of a model for autonomous agents. Then, this model is explored from various standpoints, including theoretical analysis as well as experimentation in simple environments.

1.2 Scope

The subject of emotions in the field of A.I. can be broadly divided into two areas. One area concerns emotions from the standpoint of interaction between humans and machines. This includes on the one hand, the problem of recognizing emotions in humans, and how to respond appropriately to them, and on the other, the problem of machines believably expressing emotions under appropriate circumstances. The other area concerning emotions in A.I. acknowledges the internal implications of emotional phenomena for the internal workings of agents, devoting less attention to its role in interaction. Of course, this does not draw a rigid boundary between the areas, rather, research can be found that includes concerns (and methodologies) from both.

The research reported in this thesis takes the latter approach. In accordance with Damásio's position, it acknowledges the role of emotions in decision-making, leaving out concerns about externalization of emotions. So, instead of modeling emotions as a module to be added to an already functioning cognitive architecture, this thesis addresses emotions at the core of the decision-making process. Nevertheless, the goal is not to replace existing A.I. methodologies, but rather to search for new models and new tools that can bring innovation to the field.

This thesis is about Artificial Intelligence, having in mind the applicability of the model in robots. Although the experiments presented here do not involve robots, several related efforts supported by the presented model employed robotic scenarios, both in simulation and in real robots (section 4.7.1 provides a review).

1.3 Objectives

This thesis addresses the formulation of an agent model biologically inspired by mechanisms involving emotions and decision-making.

At the core of the proposed model lies a double-representation paradigm, in which the agent represents stimuli internally under two different perspectives: one aiming at fast response, and thus it is a simple, low dimensional representation, and the other aiming at cognitive activities, such as recognition, reasoning, planning, and so on, being a complex, high dimensional one. The model addresses the consequences of associating instances of each one of these representation schemata. This association is inspired by Damásio's Somatic Markers, according to which high level cognitive representations are associated with low-level body states. Then, a set of mechanisms are hypothesized that deal with these representations: how these representations are created, how they are associated and stored in the agent's memory, and how they are utilized when the agent is faced with new stimuli.

Having formulated the model at a conceptual level, two distinct research paths are explored. The first one concerns the role of the model in the formulation of causal models about the environment, as well as the interaction of the agent with it (chapter 5). The drive for this research comes from Damásio's proposals regarding the role of emotions in the anticipation of future consequences. According to Damásio, this anticipation occurs at two levels: at the cognitive level, and at the body level.

The second research path picks up one of the hypothesized mechanisms of the agent model, and explores it from several points of view (chapter 6). Several goals are sought: (1) to provide a formal approach to this mechanism, (2) to devise algorithms to explore the benefits in terms of efficiency of the double-representation paradigm, and (3) to employ one of the representation schemata to identify relevance on the other, using two strategies: metric adaptation, and the creation of new features.

1.4 Contributions

The conceptual model was initially presented elsewhere [200, 196], and has evolved since in several directions (a review can be found at the end of chapter 4). This thesis presents and discusses the formulation of the core aspects of the model: the double-representation paradigm, together with the desirability vector, and the hypothesized associated mechanisms. Then, two implementations are presented, one dealing with anticipation, and the other exploring mechanisms to formulate causal models about the agent interaction with the environment. One of the proposed mechanisms (termed indexing) is then explored, resulting first in a probabilistic analysis of efficiency. Then, it is assumed that the internal representation schemata have metric structure. On the one hand, the efficiency of the indexing mechanism is theoretically analyzed, followed by experimental results corroborating the analysis. Also under the metric structure assumption, an algorithm is devised to extract relevance, and to contribute to the creation of new features, thus improving the representation schemata employed by the agent. This algorithm is based on Multidimensional Scaling methodologies.

1.5 Structure

First, some background covering the issue of emotions in several scientific domains is reviewed in chapter 2. This material has served as inspiration to most of research performed in the field of A.I. and emotions, which is reviewed in chapter 3. The conceptual model is presented in chapter 4, followed by chapter 5, in which some experimentation on anticipation and on the formulation of causal models is presented. The main contribution of this thesis can be found in chapter 6, where the indexing mechanism is explored from several different perspectives. This thesis closes with chapter 7, in which conclusions of the presented work as well as possible future research directions are discussed.

Chapter 2

Background

2.1 Introduction

The theme of emotions has attracted the mind of curious inquirers ever since the early times of civilization. This chapter reviews research on emotions in several scientific fields, since the early days of Greek Philosophy, up to recent neurophysiological research. This review begins with an account from philosophical thought, followed by a review of some relevant physiological models of emotions. More recently, mostly because of the advent of modern imaging techniques, neurophysiology has taken a prominent place, aiming to find the physiological correlates of mental activity, with emotions as an unavoidable subject. In this latter context, the research of António Damásio receives extensive treatment, since it constitutes a central inspiration for the conceptual model underlying this thesis.

2.2 Philosophy

The subject of emotions is clearly not a recent issue in philosophical thought. In fact, it has been embraced by philosophers from the dawn of philosophy. Early Greek philosophers, namely Plato and Aristotle, are referred by the literature as the first ones to discuss the issues of emotions in human life [119, 33]. The discipline of psychology is considered to have been born with these two Greek philosophers.

Plato's (427–347 B.C.) account of the universe divides it in two distinct realms: the realm of the ideas ("Forms"), immaterial and eternal entities, and the realm of earthly objects, material and imperfect reflections of the former entities. Human existence is composed, according to Plato, of a mind and a body, living respectively in the realms of the ideas and of the material objects. This view is usually referred to as a dualist one, in the sense of separating mind and body as two independent entities. Moreover, Plato saw human life as an ongoing struggle between reason and emotion, with each one reaching for dominance over the other. From this struggle resulted a view of emotions as something not easily controllable and potentially dangerous. Therefore Plato gave emotions a negative connotation, as something that stands in the way of the idealized world of pure rationality.

An opposing view was held by Aristotle (384–322 B.C.): human behavior resulted from the combination of a higher cognitive and a lower emotional life. Aristotle positioned emotions in the body proper, considering that our bodily responses resulted from the way humans view the world around them. Aristotle saw emotions as a not so negative aspect of life as Plato did. Depending on the point of view, Aristotle's account is seen as a dualist one [33], because he advocated a separation between mind and body as two separate entities, but also as non-dualist [119], since Aristotle put reason and affect at an equal footage.

During the Roman empire, the Stoics and the Epicureans were the two main philosophical currents whose ideas included issues regarding emotions. The Stoics (roughly 500–200 B.C.) saw emotions as self-centered phenomena, in the sense of being cognitively induced by one's beliefs. They treated the emotions in the context of ethics, where emotional behavior was considered morally subversive. In a similar way, the Epicureans (roughly 50 B.C.–A.D. 100) also held a suspicious view of emotions. The Epicureans stood by a materialistic view of the universe, as composed of nothing more than atoms. However, while the Stoics saw emotions as a lack of virtue, the Epicureans saw it as a lack of knowledge [119].

During the seventeenth century, two distinguished philosophers held opposite positions with respect to the issue of emotions: Descartes and Spinoza. René Descartes (1596–1650) centered his philosophy around the idea of an ultimate reliance on self-awareness of one's ideas, which is crystallized in the famous quote "*Cogito, ergo sum*." Descartes held a strong dualist view separating the physical world, governed by mechanical laws, from the spiritual world, where human reason lives. The connection between these two entities is fulfilled by an area of the body, somewhere close to the brainstem, which he called pineal gland. He placed the emotions in the body proper. The emotional experience — the feelings — occurred in the immaterial soul. According to Descartes, it is in the domain of the immaterial soul that everything really important takes place. This hard-edged separation between mind and body dominated the philosophical thought about emotions until present times, till the advent of behaviorism [119, 33].

Still during the seventeenth century, Baruch Spinoza (1632–1677) upheld a quite different characterization of emotions and mind phenomena. On the one hand, Spinoza considered mind and body as two aspects of the same substance — the individual's unity — corresponding to its internal and external manifestations, respectively. Mind is therefore an abstraction about a part of a *person.* And on the other, desire was seen as a drive for self-preservation, *i.e.,* a biological predisposition. It is interesting to note how this view predates Darwin's evolutionary theories. Only when one is capable of making reason prevail over emotion, which Spinoza saw as a struggle for survival, one truly reaches freedom to act. Otherwise, the biological predisposition for survival will drive ones actions. In sum, for Spinoza, emotions are something necessary, without which one's survival is at risk [33].

Out these two contrasting views, it was Descartes' that was to dominate philosophical thought for the centuries that followed. Several factors can be pointed out for this dominance. Firstly, Spinoza's writing style was less clear than Descartes', and secondly, Spinoza's denial of God went against the Christian religion, which dominated European culture at that time [119].

2.3 Psychology

The dawn of psychology can be traced back to the late 19^{th} century. Three movements contributed to the emergence of psychology as an independent discipline: physiology, which offered an understanding of the nervous system that did not exist before; psychophysics, which focused on the subjective experience of sensations, which was beyond physical considerations of biological sensors; and Darwin's theory of evolution by natural selection. These three trends shared an empirical approach to science, based on scientific experimentation, a legacy of the methods pioneered by Copernicus in the 16^{th} century [33].

One of the fathers of the field of psychology was William James (1842–1910), most notably with his book *The Principles of Psychology* [33] in 1890. James proposed a counter-intuitive model of emotions. According to James, an emotion follows the bodily changes provoked by a stimulus, rather than the bodily changes being a result of having an emotion [102]. He sustained that certain stimuli directly provoke bodily changes. The feeling of these changes is what he calls an emotion. For instance, one feels fear after detecting one's accelerated heart beat, increased respiration rate, sweaty hands, and so on. These later bodily changes are a direct consequence of being exposed to a stimulus. With these ideas, James put emotions at the heart of early days of psychology's enterprise, during its early days.

However, James' account of emotions relied mostly on introspection. In the late 19^{th} century physiology was making its first steps. But as scientific progress requires rigorous experimental methods, introspection had to be ruled out. It was following this line of reasoning that the so-called behaviorist movement emerged. J. B. Watson laid the foundations for a movement that dominated psychology for decades, with his paper "Psychology as a behaviorist views it" (1913). According to behaviorism, the only observable phenomena regarding one's behavior are limited to the stimuli one is exposed to, together with the actions one performs. Everything else is beyond the scope of scientific endeavor. In this context, emotions were described by Watson as nothing more than a pattern of physiological reactions to certain stimuli. This corresponds to reducing emotional phenomena to the bare minimum of what is externally observable [119].

Behaviorism came across several difficulties, namely, on the one hand, the inability to account for what Ryle called "Le Penseur," or in other words, an account for an internal point of view of experience. And on the other hand, the problem of diffuseness: the same situation and stimuli elicit different emotions in different persons, as well as the fact that different persons respond to the same emotion in distinct ways [119]. Trying to understand and to explain human behavior solely based on external measurable variables discards an essential component of it, namely the individual experience of situations from an internal point of view.

These limitations were some of the main factors that led to the reappearance of cognition in psychology. Cognitive approaches to psychology may be called *centralist* ones, in contrast with the behaviorist ones, which may be termed *peripheralist* [119]. However, it is important to note that cognitive psychology does not dismiss scientific methodology. Rather, it does not reject the description level mental construct, as behaviorism does. Cognitive Psychology took form during the 1950–1970 period. It was originally inspired by several sources: first, by research on human performance, mainly impelled by the Second World War efforts, where behaviorist accounts were insufficient to deal with issues like breakdown and attention; second, by information processing approaches, motivated by the emergence of computer science and artificial intelligence; and third, by linguistics, in particular the work of Noam Chomsky [1].

It was in the context of cognitive psychology that the first modern theories about emotions emerged. The following sections provide a brief overview of the most prominent theories, according to [50].

2.3.1 Appraisal theory

Today, a dominant approach to emotions in psychology is the family of appraisal theories. According to this approach, the elicitation of an emotion follows a process of subjective evaluation — appraisal — of a situation, object or event, with respect to a number of dimensions or criteria [165]. Appraisal theories cover a broad range of issues, such as the dimensions involved in the appraisal process, multiple appraisal stages, cultural differences, and pathologies, to name a few.

The appraisal theory emerged in a 1960 paper by Magda Arnold, describing the appraisal of an event with respect to three dimensions: beneficial vs. harmful, presence vs. absence of some object, and relative difficulty to approach or avoid that object. Since then, the appraisal theory has evolved significantly, due to contributions by Richard Lazarus, Nico Frijda, Keith Oatley, Philip Johnson-Laird, Klaus Scherer, Craig Smith, among many others.

According to Richard Lazarus, appraisal is not a one-step process, rather, it is composed of a primary appraisal, which evaluates an event in terms of positiveness or negativeness with respect to the person's well-being, and a secondary appraisal, which deals with the subject's ability to cope with the consequences of that event.

According to [165], four major trends can be identified in appraisal theory. One trend focuses on the issues of the *criteria* used by the appraisal process. This trend follows from the early works of Arnold and Lazarus. These criteria can be divided into these major classes: intrinsic properties of objects or events, e.g. novelty or agreeableness; significance of the event to the individual's needs or goals; the individual's ability to influence or cope with the consequences of events, and compatibility of events with social or personal standards, norms, or values. It is usually stated that the elicitation of a specific emotion follows from the profile of the appraisal outcome, according to a set of dimensions. Another trend deals with the issue of *attributing causes* to an emotion-antecedent appraisal. The *themes* trend studies the link between emotion elicitation and patterns of goal-relatedness of an event. And finally, another trend analyzes the *propositional* nature of semantic fields that underlie the use of specific emotion terms, with respect to their verbal definitions.

Although appraisal theories have been playing an important role in emotion research in psychology (and also in A.I., as discussed in the following chapter) they do suffer from several drawbacks. On the one hand, empirical evidence corroborating these theories comes mainly from verbal reports of the individuals experiencing the emotions. There is a vast literature that alerts to the pitfalls of introspection and verbal reporting. However, there is recent evidence from neuroimaging corroborating some aspects of appraisal theories (see [164] for a review). And on the other hand, appraisal theories cover solely the process mediating the perception of an event (either external, or internal, e.g., a thought) and the elicitation of an emotion. It is also important to stress at this point that appraisal, like all psychology theories, account only for humans beings. The issue of emotions in simpler biological beings is outside the scope of general psychology, and consequently of appraisal theories in particular.

Two appraisal theories have been particularly influential in the A.I. community, thus deserving a closer look here. Unlike many other psychology models, these ones present computational models, thus facilitating the process of bridging the gap between a psychology model and a machine implementation of that model. They are the OCC theory [142] and the Frijda's model [79], briefly reviewed in the following two subsections.

2.3.2 OCC theory

The name OCC theory stands for a theory of emotions proposed by Andrew Ortony, Gerald Clore, and Allan Collins, in their book published in 1988 [142]. One of its goals is computational tractability, thus making it a very appealing theory upon which A.I. models can be based.

The OCC theory proposes a taxonomy of emotions. Three classes of emotions are proposed, corresponding to reactions to *events* (e.g., pleased vs. displeased), to *agents* (e.g., approving vs. disapproving), and to *objects* (e.g., liking vs. disliking). Event-based emotions are further divided in events that involve others — including happy-for, resentment, gloating, and pity — and the ones that involve oneself — satisfaction, fears-confirmed, relief, disappointment. Reactions to agents are identified with an attribution group: pride, shame, admiration, and reproach. And reactions to objects are identified with a attraction group, which includes love and hate.

The theory is further developed on the aspects of the appraisal process details, as well as the variables that affect emotion intensities. It focus its attention on identifying the aspects in situations that imply the elicitation of specific emotions. For instance, literature often recurs to the description of situations that lead the reader to ascribe specific emotions to a fictitious personage. The name of the emotion does not have to be made explicit within the text. Another example one can consider is a sports match, where the outcome of the game may be appraised by each team's fans quite differently.

2.3.3 Frijda's model

The psychologist Nico Frijda proposed a model of emotions that also proved very influential to the A.I. community. One reason for this is the explicit addressing of emotions in robots in his writings.

Frijda takes a functional view of emotions, identifying the adaptive value as its main purpose [80]. The functions of emotions considered by Frijda are: signaling of relevance of events, detecting of difficulties in solving the problems posed by these events, providing goals for plans towards solving the detected difficulties, and to accomplish all of these in parallel. The theoretical notion of a concern, which is central in Frijda's theory, is used to model the relevance of events for a system at any given moment. Besides the affective component of emotional experience (e.g., pleasure or pain), Frijda also identifies an associated informational structure. For instance, issues like the specific way it is relevant (e.g., loss, threat, offense), the relevant properties when dealing with it, and so on. These cognitive aspects of emotions correspond to the *appraisal* concept from the appraisal theories. Frijda also identifies the awareness of one's behavioral impulses and one's bodily changes as further cognitive processes of the appraisal. The outcome of the appraisal processes are action readiness.

A computational model of Frijda's theory (ACRES: A Concern REalization model for emotionS) [79, 80] comprises several stages of stimulus processing, named core processes. Stimuli can originate from external events, as well as from internal ones, such as the case of a thought.

1. **Analyzer**: the stimulus is coded in terms of event types, as well as its implications in terms of causes and consequences;

2. **Comparator**: appraisal of the stimulus, with respect to the current system concerns. The output of this stage is an assessment of the stimulus relevance to the agent. In case of irrelevance, the process ends here. This stage can be identified with the primary appraisal concept referred in appraisal theories;

3. **Diagnoser**: a second appraisal of the stimulus, with respect to the ability to cope with it. This corresponds to the secondary appraisal in appraisal theories;

4. **Evaluator**: the urgency, difficulty, and seriousness of the event is computed at this stage. The outcome of this stage may lead to an action interruption;

5. **Action proposer**: generation of action readiness, leading to a possible action tendency, and/or change of the mode of activation;

6. **Psychological change generators**, in parallel with:

7. **Actor**: the action generator;

Stages 1–5 are subject to interaction with a set of regulatory processes. On the one hand, these processes influence the mode of processing of each one of those stages, and on the other, these processes are influenced by input stimuli, and by the core processes themselves.

2.3.4 Attributional theories

People tend to formulate explanations about events, especially negative events, in terms of what caused them. Attributional theories of emotion focus on the impact of these explanations on the likelihood, severity, and duration of emotional distress [91]. These theories are oriented towards clinical aspects of emotional disorders in patients. In this context, the hopelessness theory of depression, first proposed by Seligman in 1975, has gathered much attention of researchers in the clinical area. These theories have drawn little, if any, interest from A.I. research, possibly because attributional theories focus on pathological aspects of emotional disorders.

2.3.5 Network theories

Network theories propose that affect and cognition are intertwined by a network of associations [78]. It is well known that some affective states provoke physiological alterations in the body, such as accelerated heart beat, sweat, facial expressions, and so on. According to network theories, affective states can also have consequences upon cognitive activities, namely memory retrieval, selective attention, and association of concepts. For instance, depressed persons tend to recall negative events more easily than positive ones. Network theories were first proposed by Isen and her collaborators in 1978.

One elaboration of network theories is the Affect Infusion Model (AIM). According to this model, affective information selectively influences learning, memory, attention, and associative processes, and eventually "colors" the outcome of the person's deliberations in an affective-congruent direction. AIM proposes four information processing strategies the mind uses to accomplish a given cognitive task: *direct access strategy,* which uses crystallized, predetermined reactions and evaluations, normally used in familiar tasks;

motivated processing strategy, when there is a well-known objective or outcome, and there is little constructive processing; *heuristic processing strategy,* when there is no crystallized solution nor a defined objective, and therefore a heuristic strategy is employed, guided by a least-effort principle; and a *substantive processing strategy,* when recourse to pre-existing knowledge, as well as elaborate cognitive processes are needed to perform the task, being the most constructive processing strategy of them all. According to AIM, these four strategies are progressively more vulnerable to affect infusion. In particular, the direct access strategy is the least vulnerable, since it is based on pre-determined packages of responses, while the substantive processing one is the most vulnerable one. Many factors influence the choice of processing strategy, namely task familiarity and complexity, personal relevance and motivation, processing capacity, and mood.

2.3.6 Multi-level theories

The theories reported so far assume a single level of information processing. However, this assumption makes it hard to explain certain situations. For instance, the difference in emotional intensities while experiencing an emotion-eliciting event, and while remembering the emotion experienced in that same event. The multi-level theories propose that there are several levels of information processing in the brain [182], and thus the same topic may be represented in qualitatively different ways at different levels. Some of these levels are directly linked to emotion, while others are not.

To be more specific, four approaches are addressed. The first one is Leventhal's perceptual motor processing model, proposed in 1979. Three levels are proposed: at the lowest level, *sensori-motor* processing includes innate expressive-motor responses and feelings; then, the *schematic level* deals with memories of emotional experiences, including perceptual, motor, and affective information; and finally the *conceptual level,* consisting of propositionally organized memories, at the most abstract level. According to tis approach, emotional responses result from contributions from these three levels. However, there is a clear distinction between the schematic level, which include memories of emotional experiences capable of directly eliciting emotions, and the conceptual level, which includes memories about emotions, that only indirectly are capable of eliciting emotions (by posterior resort to the schematic level).

A second approach is the Interacting Cognitive Subsystems (ICS), first proposed by Barnard in 1985. Three simple ideas support this approach. The first one is that there are different kinds of representations (mental codes) covering distinct aspects of experience, divided in propositional and implicational (schematic) code patterns. The second idea is that there are specific processes that are able to transform one kind of codes into another kind. And the third idea is that different memory systems store different mental codes, separately. According to ICS, the elicitation of an emotion is a result of processing an appropriate pattern of implicational code. These implications codes constitute the "common currency" in which sensory contributions, on the one hand, and cognitive contributions, on the other, can be expressed, integrated, and capable of modulating the production of emotions.

A third approach, termed Multiple-Entry, Modular Memory System (MEM), advanced by Johnson in 1983, proposes that the memory is organized in distinct but interacting subsystems. It identifies two perceptual subsystems (P-1 and P-2), and two reflexive subsystems (R-1 and R-2). The basic aspects of perceptual processing are handled and stored by P-1, while the P-2 processes originate perceptual experiences of meaningful objects interacting in meaningful ways. The reflexive subsystems R-1 and R-2 are responsible for the executive and the strategic supervisory processes. According to MEM, any one of these four subsystems can contribute to the elicitation of emotions. Biologically primitive emotions arise from P-1 and P-2 processes, while the R-1 and R-2 ones are behind the elicitation of secondary (or derived) emotions[1].

Finally, a fourth approach, called SPAARS (Schematic, Propositional, Analogical and Associative Representation Systems) is the most recent one. The SPAARS approach, proposed by Power and Dagleish in 1997, aims at integrating previous contributions to multi-level theories, while rooted in a philosophical and physiological historical context. SPAARS distinguishes several representational systems: *analogical,* corresponding to sensory-perceptual levels; *propositional,* corresponding to a propositional-conceptual level; *schematic,* which is similar to the implicational schematic models of ICS; and *associative,* which provides associations among the previous three levels. The schematic level is the one primarily responsible for the elicitation of emotions. However, the associative level can also elicit emotions, without the recourse of the schematic one. Two routes for emotion elicitation can therefore be identified: one by the means of the schematic level, which comprises an appraisal process, and another by the hand of the associative level.

Multi-level theories bring forward the idea of distinct mechanisms working in parallel in the brain, with varied modes of representation. This is acknowledged by the modular nature of the above models. One interesting

[1]Secondary emotions are the ones derive from mental imagery (e.g., remembrance of a past emotional event). See section 2.4.2 below for further details.

aspect of these approaches is their ability to model conflicting issues, arising, for instance, when different levels provide contradictory assessments of a situation. Moreover, dealing with several levels also provides a means of bringing psychological models closer to the physiology of the brain. These latter models are unanimous in stating that information is processed in the brain at several levels, simultaneously, and with different modes of representation.

Joseph LeDoux [112] has performed extensive research on the physiological mechanisms of fear. The last section of [182] presents an interesting comparison of LeDoux's physiological model with ICS, providing a mapping of concepts between the two models. For instance, LeDoux's "high-road" (slow, elaborate) and "low-road" (quick, crude) levels can be mapped onto the different ways implication codes are derived. Such interdisciplinary studies are extremely interesting because they contribute for mutual corroboration of models, where each one arose from a different approach: one from psychological studies, and the other from neurophysiology.

2.3.7 Self-organization

The concept of self-organization in cognition-emotion interactions [116] is rooted in four core principles: *recursion,* in the sense of a process continually reworking and revising its own results; *emergence,* resulting from the interaction of parts within a complex system; *consolidation,* in the sense of a dynamical convergence of a complex system towards some stable state (an attractor); and different *time scales* involved in a complex system. According to the self-organization view of emotions, the appraisal process is not formed by a single step, but rather by a continuous interaction among cognitive appraisals of situations, as well as modulation of cognitive processes by emotions. Within this context, emotions guide and constrain cognitive processes.

2.3.8 Basic emotions

The goal of basic emotions research is to organize affective phenomena in a systematic way. The concept of basic emotions was introduced by Paul Ekman [71]. He identifies two meanings of the word "basic" here. One corresponds to the idea of a set of emotions that are distinguishable among them in one or more important ways. This idea contrasts with the views that all emotions are all essentially the same, only differing in degree in one or more dimensions. The other meaning of the word relates basic emotions with their adaptive value in dealing with fundamental life tasks [71]. Paul Ekman defends that there are nine characteristics that distinguish basic emotions from other affective phenomena [70]. These characteristics are: *distinctive universal signal,* such as facial expressions[2], *comparable expressions in other animals, emotion-specific physiology,* namely in terms of autonomic nervous system (ANS) patterns of activity, *universal antecedent events,* meaning common elements in situations that elicit specific emotions, *coherence in response systems* between autonomic (ANS) and expressive responses, *quick onset,* possibly arising before one being aware of them, *brief duration,* on the scale of seconds and minutes rather than hours or days, *unbidden occurrence,* since emotions do not happen by conscious choice (although people can put themselves in situations or thoughts that are likely to elicit a specific emotion), and *automatic appraisal.*

Research on basic emotions can provide useful knowledge for the computational modeling of human emotions. Some approaches to emotions in the field of Artificial Intelligence (A.I.) have used Paul Ekman's approach to designate a small set of essential emotions. However, one should be aware that these emotions refer to human ones, which may not be the most appropriate for machines, unless the plausible simulation of human emotions is intended.

2.3.9 On the cognition-emotion debate

In the 1980's a heated debate took place within the psychology community, involving the roles of cognition and emotion [111]. The debate started with Robert Zajonc's proposal that, first, emotions are fairly independent from cognition, and second, that emotions precede cognition. Such proposal received heated contesting from the cognitivist community, because they sustain (e.g., Richard Lazarus) that emotional judgments cannot take place apart from cognition and motivation, hence not independently from cognition. From a distance, this debate boils down to a matter of semantics: what each one means by the terms emotion and cognition. On the one side, one can take a broad view of the concept of emotions, together with a narrow view of cognition. In this sense, emotional judgments can be performed independently from cognition, leaving for the latter higher cognitive aspects, like reasoning, planning, and so on. However, taking the opposite approach, a cognitive nature in emotions is undeniable, since the way people respond emotionally to situations evolve with experience, through life. Psychology's understanding of cognition favours a broader view of cognition, embracing

[2]The study of facial expressions after the experience of an emotion constitutes a major source of empirical evidence behind Paul Ekman's theories of basic emotions. Ekman found out that at least some of the human facial expressions are biologically determined, rather than culturally.

all aspects dealing with knowledge (from the Latin word *cognitio*, meaning knowledge), while a neurophysiological background may hint in the other direction, since emotions and higher cognitive tasks are dealt with by different (but interacting) neural systems.

2.4 Neurophysiology

The study of the physiology of the nervous system goes back to Galen (200 B.C.), a Greek anatomist. His accounts included concepts such as nerves, transporting signals (psychic pneuma), and muscles performing movement [44]. The first scientific approaches to the localization of functional roles in the brain were proposed by Franz Josef Gall in the late eighteenth century [112]. These functions included sensing, feeling, speech, memory, and intelligence, to name a few. Unfortunately, Gall's ideas were followed in a distorted fashion, giving rise to phrenology: a non-scientific practice of ascribing mental faculties according to the bumps found on the head surface. Phrenology was later dismissed after severe criticism. However, the basic idea that specific functions can be identified with certain brain zones stands until today. Contemporary neurophysiology provides detailed maps of several cortical areas. To give a few examples, in very general terms, the frontal lobes (behind the forehead) are identified with higher reasoning; in the back of the brain lie the occipital lobes, which are identified with vision (optical nerves connect the retina in the eyes to this area); and temporal lobes (on the left and right sides of the brain) are identified with memory. Even though a quite detailed functional mapping of the brain has been achieved, it is still believed that very little is known about its functioning. Functional mapping can sometimes seem counter-intuitive. For instance, Hanna Damásio and her collaborators have identified different brain areas responsible for recognition of persons, animals, and tools [60]. Many other studies, proposing a myriad of correlates among neurophysiology and behavior (and mental phenomena), abound in the literature.

Going back in time, and focusing now on emotions, at the time William James proposed his theory of emotions [102], physiologists were mostly studying the motor and the sensory centers of the brain. Therefore, James' account was built upon a sensory-motor model of the brain: an emotion-eliciting stimulus is processed by the sensory cortex, triggering bodily changes which are enacted by the motor cortex; these changes are then percepted by the sensory cortex, resulting in the feeling of an emotion [112] (see section 2.3 above for further details).

Walter Cannon was a contemporary of William James. He became famous not only because he severely attacked James' theory of emotions, but also because of his own model of emotions. Together with Philip Bard, he formulated a model — known as the Cannon-Bard theory — based on the latter's experiments. They identified a particular brain zone responsible for relaying signals from the sensory cortex onto two other brain zones (see figure 2.1). This zone is called the *thalamus,* located in the inner core of the brain, close to where the brain connects to the spinal cord.

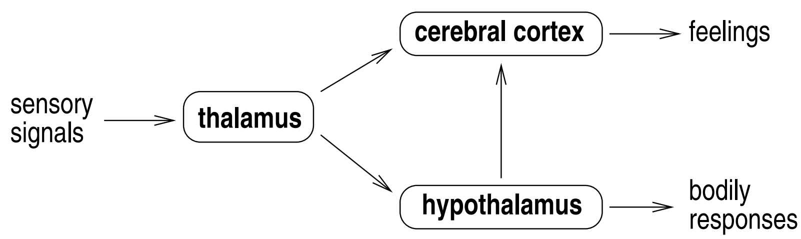


Figure 2.1: Cannon-Bard model of the brain emotional circuits (adapted from [112]).

Sensory signals are relayed to the cerebral cortex and to the *hypothalamus,* a structure near the thalamus. It is important to note that these signals are relayed *simultaneously* towards these two structures. Functionally speaking, the cerebral cortex is related with the recognition and understanding of stimuli, while the hypothalamus is related with the production bodily responses. Each one of these structures process in parallel sensory signals relayed by the hypothalamus. Moreover, the cerebral cortex also receives signals about bodily responses emitted by the hypothalamus. This simultaneous reception, after an emotion eliciting stimulus, constitute the *feeling of an emotion.*

This model was further developed and refined by James Papez (see figure 2.2). In 1937 he proposed one of the most influential models of the neurophysiology of emotions. In consonance with Cannon-Bard model, he proposed that the thalamus splits sensory signals in two streams: the stream of *thought,* and the stream of *feelings.* These streams follow distinct paths inside the brain. While the former activates the lateral areas of the neocortex, the latter activates the mammiliary bodies of the hypothalamus, which is responsible for bodily responses. An activation of this hypothalamus area results in the activation of the cingulate cortex, via the path of the anterior thalamic nucleus. The experience of feelings is attributed to the integration of the cingulate cortex activation, along with activation in the neocortex (stream of thought). Besides the anterior thalamic nucleus, which sends signals from the hypothalamus to the cingulate cortex, there is another structure — the hippocampus — which sends signals in the opposite direction.

According to Papez, this connection is responsible for controlling emotional responses, after one experiencing a feeling. On other words, it conveys the control the cortex exerts on the expression of the emotion.

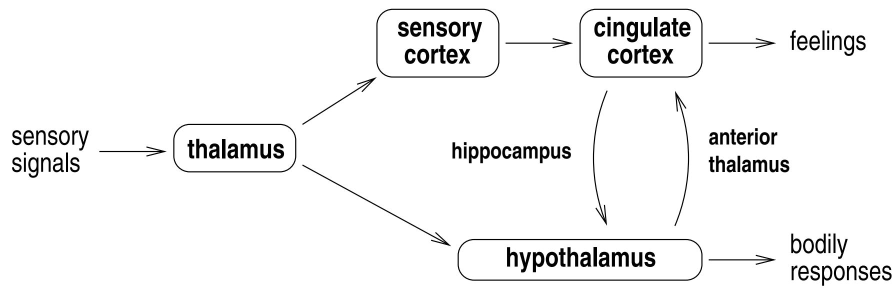


Figure 2.2: Papez model of the emotional circuits in the brain (adapted from [112]).

In 1952 Paul MacLean introduced the term "limbic system" to denote a reunion of several brain areas. He included the areas responsible for emotions identified by Papez, as well as the amygdala, the septum, and the pre-frontal cortex. According to MacLean, the limbic system contains the areas and the connections that mediate emotions in the brain. Later, in 1970 he introduced the idea of the triune brain. Comparing the brain's general structure across different species, MacLean proposed three major zones of the brain, corresponding to three grand evolutionary stages. The oldest one comprises the *reptilian brain,* which the lower vertebrates (e.g., reptiles, birds, amphibians, and fishes) share with all other species; then, there is the *paleomammalian brain,* that can be found in lower mammals, but is absent in reptiles; and finally the *neomammalian brain* can only be found in the highest primates, namely humans. Each one of these "brains" has its own properties regarding the nature of intelligence, memory, perception and motor functions, and so on. The underlying conjecture is that which are present in the brain of any given animal determine, to a large extent, the level of complexity of its behavior.

Subsequent research revealed flaws in the MacLean triune brain model, in the sense that those areas are not as clear-cut as his model suggests, both from a morphological and from a functional point of view. For instance, regions identifiable in function with the neocortex have been found in many primitive creatures. Moreover, some researchers, including Joseph LeDoux, state that there is no such thing as a limbic system [112].

2.4.1 Joseph LeDoux

Joseph LeDoux's contribution to the field of emotions has focused on the detailed study of the brain circuits of fear [113, 112]. In order to avoid the subjectivity LeDoux considers inherent to the study of emotions, He opted to choose a well defined emotion — fear — and to study its mechanisms exhaustively. Fear conditioning is a behavior paradigm introduced by Pavlov in 1927 [143]. The paradigm is based on two kinds of stimuli: an unconditioned stimulus (US) that prompts an innate response by the subject, and a conditioned stimulus (CS), which is meaningless to the subject prior to the experiment (or at least with respect to the US).

In Pavlov's well known salivating dog experiment, the dog is repeatedly exposed to a piece of meat along with the sound of a bell. The dog starts salivating after the sight of the meat. However, a bell is tolled every time the meat is shown. After several such pairings, the simple exposure to the sound of the bell is sufficient for the dog to immediately start salivating. In this example, the unconditioned stimulus (US) corresponds to the meat, which unconditionally makes the dog salivate. The bell, since it becomes associated with the dog salivation, is the conditioned stimulus (CS). This association is called *conditioned reflex.*

Research conducted by LeDoux and colleagues focused on tracing the activation patterns and the pathways in the brain involved in these conditioned reflex processes [113]. They found (see figure 2.3) that the sensory signals from the US and the CS meet at the amygdala, in the lateral nuclei (LA). In the case of an emotion eliciting stimulus, the LA activate another amygdala area — the central nuclei (CE) — which trigger behavior, autonomic, and endocrine responses. In the case of the US, as well as of the CS, there are pathways originating from the cortex and from the thalamus. Each one of these pathways has particular characteristics. The pathways from the cortex appear to be connected with stimulus processing of higher complexity. Moreover, the projections of these pathways towards the amygdala have different plasticity properties: the ones coming from the cortex learn more slowly than the ones originating in the thalamus. This indicates that fear conditioning occurs initially through the thalamus pathways. There is a third party found to be involved in fear conditioning: the basal (B) and accessory basal (AB) areas of the amygdala. These areas receive projections from the hippocampus, which is believed to be related to contextual information. Fear conditioning experiments with rats, for instance, have shown that contextual information intervenes in the process. Rats that return to a chamber where tone (CS) and shock (US) have been paired, exhibit fear responses even without any other stimulation.

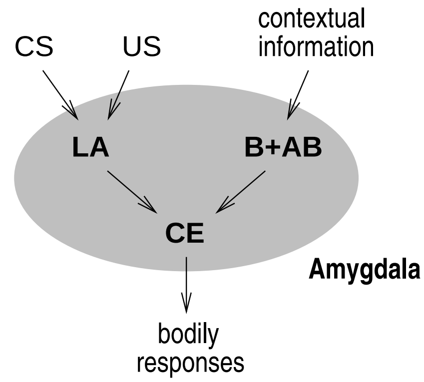


Figure 2.3: Schematic view of emotion circuits involved in fear conditioning, according to LeDoux [113]. Abbreviations: conditioned stimulus (CS), unconditioned stimulus (US), lateral nuclei (LA), central nuclei (CE), basal (B), and accessory basal (AB) areas.

Besides the pathways projecting to the amygdala, originating from several other brain areas, the amygdala projects signals to several brain zones. These projections are related with the influence that the emotional state (the activation patterns in the amygdala) exerts on other cognitive processes. LeDoux refers several examples: although the amygdala receives signals from the later stages of sensory processing from the cortical areas, it projects back to its earlier stages. This allows the amygdala to indirectly modulate the kinds of inputs it receives from those areas. Another example referred by LeDoux regards emotional memories: it is known that implicit emotional memories depend on the amygdala, while explicit ones depend on the medial temporal lobe[3]. However, the amygdala seems to modulate the storage of explicit memories, in a way that memories with emotional content tend to be longer lasting and more vivid than the non-emotional ones.

The working memory is a *locus* of integration of several brain mechanisms: sensory information, memory, emotions, and so on. There, information is contrasted and manipulated in several ways. LeDoux suggests that this integration gives rise to the conscious experience of feelings. Moreover, there are projections from the amygdala onto these areas, suggesting a modulating role in the functioning of the working memory.

[3]This dichotomy of memory, due to Daniel Schacter, is based on the conscious awareness of a memory: *implicit memory* (also known as *procedural*) depends on unconscious factors (e.g., skills, conditioned reflex), while the *explicit* one (also known as *declarative*) involves conscious awareness [112].

2.4.2 António Damásio

António Damásio has published several prominent books on the subject of emotions. Although his books have targeted primarily the general public audience, his ideas and theories have yielded a profound impact in many scientific disciplines [55, 56, 57], including fields not directly connected with neuroscience [175, 118, 90]. One of his main research areas involves the underpinning of the neurophysiological correlates of emotions in decision-making.

Early research by Damásio involved system level models of the brain [54, 53, 58]. The brain receives signals from various sensory modalities (visual, auditory, tactile, olfactory, and so on), and processes these signals through layers of neuronal structures. There are convergence zones in the brain that receive signals from these sensory processing regions. These convergence zones are amodal, in the sense that they receive signals simultaneously from various sensory modalities. The mechanism of integrating information from various sensor modalities into entities and events is termed the *binding problem*. According to Damásio, these convergence zones, not only perform this integration, but are also capable of retro-activating early sensor processing zones [54]. This means that, for instance, when a person recalls a situation, the brain activates several early sensory processing cortices, in a synchronous fashion. There are feed-back, as well as feed-forward projections among these early sensory cortices and the convergence zones. Thus, a recalled situation is represented in the brain, not only distributed over many diverse regions, but also in the low-level sensory cortices, from which those representations originated in the first place. For instance, the reader can perceive this by trying to answer this simple question: "how many windows does the house you live in have?" Can this question ba answered without actually reconstructing visually one's house inside the mind? Unless the unlikely case of the reader having this number memorized, one needs to actually reconstruct in her/his mind the spatial structure of the house, in order to count the windows one by one.

Somatic Marker Hypothesis

The somatic-marker hypothesis [59], which is central to Damásio's proposal, originated from the study of some particular projections among the frontal lobes and the central autonomic control structures. It is well known that the frontal lobes are related with high-level executive functions, such as reasoning, planning, decision-making, and so on. Damásio studied patients with damage in a particular zone of the frontal lobes — the ventromedial sector — which receives projections from all sensory modalities. Moreover, it was the only known source of projections, at the time this hypothesis was first proposed, from the frontal regions towards the central autonomic control structures [59]. These latter structures are responsible for many regulatory functions of the body. In particular, they are believed to control physiological body changes, namely the ones elicited by emotions. What is the role played by all these connections among the highest levels and the regulatory subsystems?

Damásio reports that patients with lesions in the ventromedial sectors of the frontal lobes exhibit a peculiar behavior. First, their intellectual capabilities were found to be intact, as was validated by various psychometric tests, ranging from standard IQ tests, to learning and memory tests (with the exception of social knowledge). Yet, these patients showed severe impairments in certain circumstances: they were unable to choose a course of action that would be clearly advantageous to them in the long-term; instead, they tend to plunger in endless debates over secondary matters, such as what to wear, where to shop, and so on; moreover, they also revealed impairment in social situations [59].

With the goal of providing an explanation the above phenomena, Damásio advanced the Somatic Marker Hypothesis (SMH) [59, 55]. In a nutshell, it proposes that decision-making in normal individuals is assisted by "the appearance of a somatic signal that marks the ultimate consequences of the response option with a negative or positive somatic state" (page 220, [59]). These somatic signals can be either conscious or covert, but they are physically measurable. One common such measure is the change in skin conductance. The measurement of these changes is called Skin Conductance Responses (SCR). Conscious effects of this somatic marking are, for instance, the "gut feeling" when a response option is considered, as well as attention focus on such responses. Covert effects include appetitive or aversive behaviors towards/away certain response options [59]. These mechanisms employ machinery supporting emotional processes.

Damásio divides emotions into two broad categories: *primary* and *secondary* emotions. The primary emotions are the ones corresponding to direct and immediate response to stimuli (e.g., following an unexpected loud sound), corresponding to the "low-road" in LeDoux's terminology. These responses operate before conscious awareness, and make use of the evolutionary older structures in the brain. The secondary emotions involve evolutionary newer structures (the LeDoux's "high-road"), and follow from thought processes, as in the case of the recollection of episodes eliciting associated emotional states. In this latter case, these emotions provoke actual measurable changes in the body state. The enactment of these body states depend on projections from the pre-frontal cortex, where the eliciting thoughts are formulated, towards the amygdala area, which promotes the body state changes. These projections originate at the above-mentioned ventromedial sector. The perception of these changes, together with the recalled images, is what Damásio calls the *feeling* of an emotion.

Two aspects are crucial to the SMH: one is that somatic signals are enacted *before* decisions are made, even before any cost/benefit analysis of the available options; and the other is that such enactment involves low-level, visceral, evolutionary old regions such as the autonomic control structures. Moreover, this enactment plays an influential role in the decision making process, with different degrees depending on the nature of the situation. Patients with lesions in the frontal lobes became unable to take certain common day decisions, while maintaining psychometric performance levels intact.

If this hypothesis is correct, thought can no longer be considered a disembodied process, separable from physiological mechanisms. This refutation of the Cartesian dualism motivated Damásio's book title "Descartes' error:" mind proper and body proper are two faces of the same coin, according to Damásio.

In sum, the high-level decision processes in the brain do not unfold isolated from the body, as in a metaphor of a computer inside a robot. Rather, decision processes consult the body, both by provoking body changes while analyzing certain options, as well as by being sensible to the resulting body changes. In result, certain options are rejected straight away, for being considered repulsive by the body, while others may receive salience, if found desirable by it.

This allows for a reduction in complexity of decision processes, as many available options can be put aside from the rational cost/benefit analysis that follows. The body can promptly reject certain options, either because of their undesirability, or their irrelevance. And because this pre-selection is performed outside of the scope of conscious thought processes, it is sometimes hard to explain to oneself in verbal terms. It is plausible to consider that human intuition falls into this category. Intuition is closely related to experience. It is not easily transmitted orally: one cannot simply teach intuition directly, at least without putting the students into situations where they can obtain that intuition themselves by experience.

Damásio also discusses an alternative mechanism to the body consulting process: the "as-if" loop. By using it, the decision making processes can skip the actual interaction with the body. The "as-if" loop mimics the body responses, without actually changing the body proper. However, note that what this mechanism does is just to emulate the body responses in those circumstances. And the way of doing this is to learn how the body actually responds to those situations. Therefore, from the conceptual point of view, it remains a consulting of the body response.

Testing the hypothesis

To test the plausibility of the SMH, Damásio and his colleagues conducted several experiments whose results are briefly reviewed below. One of them consisted of exposing patients with lesions in the ventromedial sector of the pre-frontal cortex, as well as normal controls, to three kinds of stimuli: (1) unconditioned stimuli, such as an unexpected loud hand-clap close to the subject's ears, (2) target pictures, such as ones of social disasters, mutilations, nudity, and (3) non-target pictures, showing bland scenery or abstract patterns. During the experiment, the skin conductance of the subjects was recorded (SCR).

All subjects behaved similarly on all tests: SCR was found in all subjects when exposed to unconditioned stimuli, while no significant SCR followed non-target pictures. However, patients with lesions did not generate SCR when target pictures were shown, unlike control subjects, where SCR was detected. According to the SMH, these lesions prevented signaling from the pre-frontal cortex to the autonomic control structures, thus preventing the reenactment of body states after the exposure to target pictures. These patients were therefore unable to express in the body the appropriate emotions following target pictures, thus not feeling in the way normal subjects felt. These patients reported later that they realized they were not feeling in they way they used to feel before the lesions [55].

A second experiment conducted by Damásio and his colleagues, designated the good-guy/bad-guy experiment [56, 187], can be described as follows. A patient with a specific kind of brain lesion preventing him from remembering certain forms of factual knowledge, namely faces, was subjected to the following situation: three persons, previously unknown to the patient, contacted him several times over a week, each one playing a pre-designated role. The first one (good-guy) did good things to the patient every time they met (e.g., smile, talk nicely), the second one was neutral, and the third one did unpleasant things to him (bad-guy). Every time the patient met one of these persons, he neither recognized them, nor recalled their previous encounters, because of the lesions to his brain. Pictures of these persons were then shown to the patient, and although he was unable to recollect any factual information about them, when asked who he thought was his friend, he tended to prefer the good-guy over the others about 80% of the times. The neutral-guy was choosen with a probability no greater than chance, while the bad-guy was almost never chosen. These intriguing results suggest a covert learning of affective valence independently from factual learning. This patient had no recollection of the encounters with those persons, still, he was able to perform affective evaluations solely based on their faces.

A third experiment was a card game designed to address the SMH [55, 23]. This game, posteriorly known as the Iowa Gambling Task (IGT), consists of four decks of cards, labeled A, B, C, and D. This is a one person game, where the player is given an initial loan of $2000. In each turn, the player is asked to choose one of the decks. A card is drawn from the chosen deck, and the player is informed of the amount of money either earned or lost, as a consequence of her/his choice. The game was designed in the following way: decks A and B usually yield earnings of $100, but occasionally there are cards that make the player lose as much as $1250, while decks C and D usually offer a more modest earning of $50 per card, and occasional losses are not higher than $100. The contents of each deck was formed in such a way that in the long run, decks A and B are disadvantageous, while the other two are advantageous to the player. The players are obviously ignorant of these facts before playing the game. Each game is made up of 100 turns. Two groups of subjects played this game: a normal control, and patients with lesions in the ventromedial sector of the frontal lobes. The results are remarkable: after sampling some cards from each deck, normal controls tended to prefer the advantageous decks C and D, while the frontal patients tended to prefer the other two (disadvantageous) decks. Frontal patients seemed to prefer the decks that gave larger immediate amounts, but most important of all, they seemed to be *insensitive* to the high risk involved with those decks (occasional high losses). This insensitivity to future consequences is referred to by Damásio as a “future myopia.” It was also found that normal patients showed SCR immediately before choosing a deck, while frontal patients did not show any significant SCR. However, both players showed SCR after knowing whether they had gain or lost money. The lack of SCR prior to deciding, in the frontal patients, is coherent with the SMH, in the sense that the absence of SCR prior to deciding prevents these patients from being sensitive to the future consequences, in this experiment, of their possible choices .

These results suggest that the pathways from the amygdala to the prefrontal cortex, signaling body state changes, have a determinant effect on the choice of the advantageous course of action. Pre-frontal subjects seem insensitive to the prospects of high losses, thus suggesting that the body plays a key role in representing those unpleasant prospects. The SCR prior the decision seems decisive to the appropriate decision of avoiding the disadvantageous decks.

The impairments manifested by these patients, in the first and third experiments, is not visible in well-defined tasks, since IQ tests show results within average. The kind of tasks revealing reduced competence concerns long-term decision-making. Damásio refers the example of the Phineas Gage case, a railroad construction worker who after severe lesions in the pre-frontal cortex, became unable to sustain his previously stable family life nor his professional career [55]. His life became a turmoil. Another example discussed by Damásio (referred by him as Elliot [55]) is of another patient sustaining similar lesions. Elliot exhibited a curious impairment to perform simple common day tasks, such as scheduling an appointment with a doctor, or simple organizational tasks in an office. In these cases, this patient seemed immersed in a sea of indecision, taking incredible (unreasonable?) amounts of time to make up his mind, pondering all the pros and cons of each option *ad nauseam.*

When the precise consequences of decisions are hard to predict in all of their details, and thus a capacity to ponder in general terms is at order, these patients seem unable to reach a decision. They seem unable to realize the relevant aspects of decisions when the consequences of the available options are not well-defined. The Somatic Marker Hypothesis advances an explanation to this impairment. When faced with a decision, the brain ponders possible courses of actions, in an means-ends analysis fashion. Their consequences of these are considered (and possibly anticipated), in turn. The first aspect is that the body responds to each of these possibilities. The SCR shown by normal patients reflects this consulting of the body. The second aspect of the proposal is that these changes in the body are signaled back to the brain. This signaling can have several effects in the decision-making process: some possibilities may be rejected without any further consideration, while other possibilities may receive further attention, held in working memory more vividly, and receive additional analysis. Although the influence of emotions is commonly been recognized in decision making, the novelty introduced by Damásio centers on two proposals: first, that emotions are essential for appropriate decision-making, and second, that they play a role even in those decisions not explicitly involving emotions.

Taking an extreme example, in the proof of a non-trivial mathematical theorem, even though the steps taken must obey strict mathematical rules, emotions may play a role in the choice of the followed path, the employed abstractions, etc. Illustration of this influence is visible when mathematicians talk about elegance of proofs and creativity. Henri Pointcaré once wrote that

> In fact, what is mathematical creation? It does not consist in making new combinations with mathematical entities already known. Anyone could do that, but the combinations so made would be infinite in number and most of them absolutely with-

> out interest. To create consists precisely in not making useless combinations and in making those which are useful and which are only a small minority. Invention is discernment, choice.
> [...]
> To invent, I have said, is to choose; but the word is perhaps not wholly exact. It makes one think of a purchaser before whom are displayed a large number of samples, and who examines them, one after the other, to make a choice. Here the samples would be so numerous that a whole lifetime would not suffice to examine them. This is not the actual state of things. The sterile combinations do not even present themselves to the mind of the inventor. Never in the field of his consciousness do combinations appear that are not really useful, except some that he rejects but which have to some extent the characteristics of useful combinations. All goes on as if the inventor were an examiner for the second degree who would only have to question the candidates who had passed a previous examination. (taken from page 188 of [55])

Automatic theorem provers exist nowadays, but they only seem capable to prove trivial theorems. These systems seem unable to scale with the domain complexity. A way to circumvent this limitation is to include additional techniques, such as including domain-knowledge, or well-crafted heuristics. But if it is the case, then one may question to what extent the actual intelligence resides with the machine and with the designer.

Questioning the SMH

The Iowa Gambling Task (IGT) is one of the pieces of supporting evidence for the SMH. However, the validity of the conclusions drawn by Damásio and colleagues from this experiment has been contested. For instance, Tiago Maia *et al.* dismiss not only any unconscious knowledge of the advantageous deck, but also the statement that patients with lesions under-perform because of the impairment of the somatic marker mechanism [124]. Maia reproduced the IGT under the same conditions (but not with patients with lesions), where the subjects were asked a more detailed questionnaire about their knowledge of the game. The results have shown that subjects have more conscious knowledge about the most advantageous strategy than suggested before, even prior of behaving according to that strategy. In a second line of reasoning, Maia cites experiments performed by Lesley Fellows and colleagues on the consequences of reward reversal [77]. Reward reversal corresponds to changing the learning conditions in a radical fashion mid-way an experiment, in order to assess whether subjects are capable of adjusting their behavior appropriately. For instance, switching the conditions on which positive and negative rewards are given. The results show that patients with the same kind of lesions as the ones in the original IGT demonstrate inability to adapt to the new conditions after reversal. Maia sustains that these results show that the reason behind the under-performance of patients with lesions is more related to their inability to adapt their strategy after getting a $-1250 card, than to any lack of a somatic marking mechanism [124].

Antoine Bechara *et al.* published a response [25] to Maia's publication, which was promptly replied by Maia [125] in the same journal issue. Bechara sustains that Maia's findings are not contradictory with the SMH. Conscious knowledge of the advantageous strategy is addressed by the hypothesis "by proposing that pure cognitive processes unassisted by emotional signals do not guarantee normal behavior in the face of adequate knowledge," while "cognitive processing assisted by emotion-related marker signals, conscious or not, contributes to the proper action being taken." [25]. Maia's response follows the idea that there are more plausible explanations than the SMH, not only to the IGT results, but also to other evidence [125] (not explicitly pointed out by Maia). For instance, behaving in a way incoherent to one's knowledge of the advantageous strategy can be explained as an exploratory behavior. The only issue both seem to agree on is the presence of many open questions on the subject.

The body

One central issue in António Damásio writings is the importance of the body, hence the suggestive title "Decartes' Error" [55] of his first book. René Descartes upheld the separation of mind and body as two distinct entities. Moreover, the body was considered pernicious to pure rationality. Damásio, on the contrary, reclaims the role of the body as a crucial one. First, mental imagery is not only interpreted at the cortical level, but also at the level of the body. In other words, the body responds to that imagery, independently and simultaneously to the cortex processing. And second, the higher levels of the brain hold, at the same time, their own representations at the cognitive level (factual information, memory, cost/benefit analysis of response options, and so on), as well as representations of the body changes incurred. Thus, these body changes provide a "coloring" of one's thoughts. Since each person's individual life experience is unique, Damásio sustains that subjectivity arises from the differentiated somatic marking of each individual. Exposed to the same situation, different persons have different subjective responses, since their bodies respond differently to it.

Chapter 3

Review of the state-of-the-art

3.1 Introduction

There is a large body of research on emotions in the field of Artificial Intelligence (A.I.). This research has evolved from the initial early implementations in the 1980's, to an expansion of the field beginning at about 1998. This chapter presents a review of the evolution of the field. Since several different approaches can be identified, a taxonomy of the field is also proposed here. The review of recent research is then presented under this taxonomy.

3.2 Artificial Intelligence

The field stems from the two concepts evidenced by the name Artificial Intelligence: the goal of creating *artificial* systems exhibiting functionalities classifiable as *intelligent*. While the latter refers to a natural phenomenon observable in biological beings (prominently in humans), the former refers to what Herbert Simon calls the Sciences of the Artificial [173]: the study of man-made artifacts, in contrast with the traditional sciences focused on natural phenomena.

A review of several approaches to the definition of Artificial Intelligence (A.I.) can be found in the Introduction section of Russell and Norvig's textbook "Artificial Intelligence: A Modern Approach" [161]. Broadly, the nature of these definitions differ along two dimensions, one concerns the perspective of thought processes — is A.I. attempting to mimic human thought? — and the other the perspective of behavior and acting — is A.I. attempting to mimic human behavior? Along the former dimension, A.I. can be understood as either the replication of the human cognitive mechanisms, or a rational mechanism undressed from biological plausibility concerns (e.g., a logic approach). Regarding behavior and acting, A.I. can be approached as either an attempt to replicate human behavior, or to behave in a rational way. These are, however, extreme cases. Introductory textbooks often address the issue invoking the "strong A.I." and "weak A.I." perspectives. The former corresponds to the ambition of human-like thought and behavior, by machines, while the latter focuses on whether machines act rationally [161]. While the "weak A.I." perspective embodies a pragmatical drive towards working systems, the "strong A.I." one represents the long-term goal of attaining truly intelligent systems, regardless of how nebulous this idea might be.

From a historical perspective, many scientific fields contributed to the emergence of A.I.: philosophy, mathematics, psychology, computer engineering, and linguistics. The first publication commonly accepted as being the precursor of A.I. was written by Warren McCulloch and Walter Pitts in 1943, proposing a model of artificial neurons. Each one of these neurons can assume one of two states ("on" and "off"), receiving signals from other neurons. They showed that this model could implement logical connectives, and even perform any computable function [161].

In the late 1940's Alan Turing wrote an influential paper (only published later [189]) claiming the possibility of machines exhibiting intelligent behavior, while presenting several arguments sustaining his claim.

The name of the field was however only coined in 1956 after a workshop in Dartmouth in the summer of 1956, organized by John McCarthy, Marvin Minsky, Claude Shannon, and Nathaniel Rochester. The list of participants of this workshop included names such as Allen Newell and Herbert Simon.

The field has evolved over several stages, from the early enthusiasm of naive systems and toy problems, to a maturity state where strong theoretical results can be found as well as real-world systems actively used in production by the industry. The A.I. now field counts a large community of researchers. It has remained a quite inter-disciplinary subject, with strong connections to the many fields of mathematics (logic, statistics), robotics, linguistics, neuroscience, systems and control theories, among others.

3.3 Early implementations

Human intelligence has naturally inspired most A.I. researchers. The early days of the field were particularly prolific in this respect. Since emotions play a major influence in human behavior, it is natural to assume that the design of intelligent systems should take those issues in account. Such a proposal was first brought forward by Herbert Simon in a paper from 1967 [172]. In this paper, Simon considers systems with multiple goals. Mechanisms for selecting which goal to seek at a given time are considered, such as a simple goal queuing strategy. However, when faced with real-time systems, where the survival of the system depends on its response time in certain situations, such mechanisms are inadequate. An *interrupt system* capable of interrupting current processing in order to attend to a real-time solicitation is considered by Simon as an emotional behavior mechanism. Such a mechanism requires a continuous monitoring system, running in parallel with the current goal-seeking mechanisms. When a particular situation is detected, either of internal or external origin, the current goal-seeking process is interrupted, and the solicitation is attended. This emotional mechanism opens up several possibilities for learning. Two possibilities discussed by Simon are: the capability of learning new associations among stimuli and interrupt mechanisms, as well as weakening such associations, and to acquire or modify packaged responses to these interrupts.

Although Simon does not expand on how to implement these ideas, it is interesting to note that the relevance of emotion phenomena in human behavior caught the attention of one of the founders during the early days of the field. The particular aspect of emotions that Simon found interesting was not the machine handling of specific human emotions, but rather the mechanism of interruption of cognitive processes.

With the provocative title "Why robots will have emotions" [176], Aaron Sloman and Monica Croucher sketched a complex (and sometimes confusing, in the attempt of broadly covering many aspects of human intelligence) architecture of the mind. In a similar fashion as Simon, emotions are taken to play the role of interrupting current processing in order to cope with the vicissitudes of a changing and partly unpredictable environment. Sloman and Croucher conjecture in their paper from 1981 that "interruptions, disturbances and departures from rationality which characterize emotions are a natural consequence of the sorts of mechanisms required by the constraints on the design of intelligent systems." These constraints are of many sorts (physical needs, mental needs, social ones, and so on). To name a few of them: non-static collection of motives, non-static environment, speed of computation, environment complexity, complex physical structure, etc.

Both Simon and Sloman approached emotions as mechanisms of a larger system. They claimed that a cognitive system complex enough to exhibit intelligent behavior at a level similar to the human one ought to incorporate emotion-like mechanisms. Note that the focus is not on specific human emotions, but rather on the mechanisms involved in emotion processing. Thus, this does not dismiss the possibility that one or more human emotions may uncover mechanisms generic enough to be useful outside of the biological context. To what extent human emotions are specific to biology and to specific environments remains an open question.

An alternative approach is possible, though: the design of a system endowed with representations and mechanisms closely based on human emotions. Early works on this line of research include the ones of Jaap Swagerman, based on the emotion theories of Nico Frijda, and a review made by Michael Dyer on previous computer models that exhibit comprehension and/or generation of emotional behavior.

Michael Dyer reviews in [69] three computer models — BORIS, OpEd, and DAYDREAMER — that he claims to exhibit comprehension and/or generation of emotional behavior. None of these implementations was designed to specifically address the issue of emotions. They all interact with the user via a natural language interface. The representations of emotions correspond to the ones captured by the corresponding words, which is the same to say that they aim at representing specific human emotions.

The BORIS implementation reads narrative texts and answers questions about probable emotional reactions of characters. The representations are based on affective units (Minsky's frames [133]). Affects are represented in terms of abstract goal situations, e.g. "unhappy" indicates a character has experienced a goal failure. Each affect unit includes also which character is feeling the emotion, a polarity (two values: negative or positive), and a person or object towards which the emotion could be directed (e.g., anger), among other slots. BORIS is able to infer emotional responses after a given situation, as well as to interpret a given emotional reaction in terms of its probable causes.

The OpEd implementation is an extension of BORIS, aiming at reading editorial texts and answering questions about them. The main innovation of OpEd is the introduction of beliefs. This way, OpEd is able to represent different points of view held by different characters. Beliefs are also represented as frames whose slots include the believer, the belief content, and pointers to beliefs that attack and that support it. Such an explicit representation of beliefs contributes significantly to the representational power of OpEd. For instance, it becomes able to understand emotional reactions resulting from a character's beliefs. An interaction example of OpEd follows (from page 333 of [69]):

> Q: Why have limitations disappointed Friedman?
>
> A: MILTON FRIEDMAN BELIEVES THAT VOLUNTARY EXPORT RESTRAINS ON AUTOMOBILES FROM JAPAN NEGOTIATED BY THE REGAN ADMINISTRATION WILL CAUSE LOSS OF JOBS IN THE U.S.

The goal of DAYDREAMER, unlike the above implementations, is to generate streams of thought (again, in natural language form). DAYDREAMER operates in two modes: an input mode, where plans are executed in the "real world," and a daydreaming mode, where a stream of thought is generated, based on the "real world" events. Several components constitute DAYDREAMER, to name a few: a personality module, comprising high-level goals (e.g., preservation of self-esteem, social status, etc.), a relaxed planning mechanism (more fanciful dreams result from relaxed constrains), and an emotional component that alters the process of daydreaming by feedback of dreamed emotions (e.g., imagining a goal failure makes DAYDREAMER "sad"). The daydreams are represented by conceptual structures, and are afterwards translated to English for output. An example daydream follows (from page 337 of [69]):

> I am disappointed that she didn't accept my offer... I imagine that she accepted my offer and we soon become a pair. I help her when she has to rehearse her lines... When she has to do a film in France, I drop my work and travel there with her... I begin to miss my work. I become unhappy and feel unfulfilled. She loses interest in me, because I have nothing to offer her. It's good I didn't get involved with her, because it would've led to disaster. I feel less disappointed that she didn't accept my offer.

All these three implementations are based on viewing emotions as patterns of beliefs, goals, and arousals. Michael Dyer provides some support for this view, citing Robert C. Solomon who argues that emotions are rational judgments. Solomon states that "raw" feelings are distinct from beliefs and goals, but require beliefs and goals in order to be identified as emotions. In urgent situations, for instance, emotions are still judgments, however they may seem irrational to us in the larger context.

Note that all of these implementations represent emotions at the verbal level. Moreover, they work with human emotions and require built-in representations of how they can be represented in terms of goals and beliefs. It is well-known that emotional phenomena often operate covertly, and therefore one's own interpretations may be misleading. There is evidence that the brain has a tendency to confabulate verbal explanations (see split-brain experiments [112, 87]). With this in mind, it may seem exaggerated to state that these implementations represent emotions *per se*. Rather, it seems more accurate to say that they work with representations of verbal interpretations of emotional phenomena. In other words, emotions are put in a rational and symbolic framework, within which inference mechanisms generate the programs' output.

The PhD dissertation of Jaap Swagerman describes ACRES, an implementation of Nico Frijda's architecture [81]. The fundamental issue is the realization of concerns. The term *concern* is understood in this context as "the system's disposition to evaluate events or internal conditions as desirable or as undesirable" [81]. These concerns are responsible for seeking the satisfaction of the system major goals. The system is based on Frijda's architecture proposed in his book [79].

The ACRES[1] program is written in Prolog, and interacts with a human operator via a text-based interface. It is endowed with the following set of concerns: continued operation ("avoiding being killed"), continuous operation (preservation of reasonable waiting times), reception of correct input, reception of interesting input, and continued unchanged operation (safety). This set of concerns is designed for the environmental niche the system resides in, *i.e.,* an operator-machine interaction. The program is composed of several task-components, which seek the realization of the above concerns. While the system seeks to realize all concerns simultaneously, some concerns take precedence over others. As input is received by ACRES, these task-components check for concern realization. The system learns from descriptions of emotions given by the operator, and stores knowledge about its own emotions, as well as other people's emotions.

To attain concern realization, ACRES has the ability to perform actions. These actions can either have external effects, or modify its knowledge base. It also disposes of pre-programmed action sequences, and is capable of planning sequences of actions. The detection of a concern relevance often leads to goals being set up. The choice of particular actions is based on the relevance evaluated for the present situation. ACRES builds memories of its own experience. Since these memories are formed by relevant aspects with respect to the system concerns, it is claimed that they constitute an emotional experience memory. Another aspect of ACRES is the use of its knowledge base to name emotions. The system has been tested with a set of 960 emotion profiles obtained from real life experiences of 32 subjects (30 emotions involved). ACRES was able to give correct first choices for 32% of the cases, while in 71% of the cases the correct one was among the top five choices [81].

The authors of ACRES acknowledge the limitation imposed by the system environment (text-based operator-machine interface) where, for instance, there is no concept of space or movement. This limitation prevents the supporting emotion theory to be fully tested. Even so, some concepts of the theory, such as action readiness, control precedence, hedonic signals, and concerns, could be implemented.

[1]ACRES: A Concern REalization model of emotionS.

3.4 Historical perspective

The evolution of emotions research within the domain of A.I., can be roughly divided into three major periods. First, the early days of the field, spanning from the 1960's until the end of the 1980's. Herbert Simon published his seminal paper about emotions and its relevance to the A.I. field in 1967 [172], and in 1981 Aaron Sloman (together with Monica Croucher) published "Why Robots Will Have Emotions." The 1980's saw the creation of the first computer programs dealing in an explicit manner with emotions, namely BORIS [69] in 1983, and ACRES [81] in 1987.

A second period can be roughly identified with the first half of the 1990's. This period is characterized by some dispersed research, during which the first applications of emotions were implemented. Two examples are the Oz project [21, 156, 155] in the field of believable agents, and MINDER1 [216], an implementation based on Sloman's architecture.

The year 1998 marked a new era for the field, with the emergence of several meetings dedicated to the topic, namely the workshop "Grounding Emotions in Adaptive Systems" in the Simulation of Adaptive Behavior conference (SAB98), and the AAAI Fall Symposium on "Emotional and Intelligent: The Tangled Knot of Cognition." Several others followed in the next few years. These gatherings aggregated hitherto dispersed research, allowing for the creation of a community of researchers working on this area.

However, it is hardly the case that the field is united among a common goal, nor a common theory, for several distinct research goals and interests can be identified. Moreover, as it became evident in earlier sections of this chapter, theories about emotions abound. These theories are not necessarily contradictory, but rather attempt to describe the emotional phenomena from different perspectives, and sometimes from different initial assumptions. They also vary in terms of level of description: from the physiological level, as in Joseph LeDoux's research, where fear has been singled out, to the higher psychological levels, as in Appraisal theories, taking a cognitive stance.

3.5 A taxonomy of the field

This section proposes a taxonomy of the field. At a first level, research is divided among two areas, where research is focused on *internal manifestations,* or on *external manifestations* of emotions. It is possible to develop research that intertwines these two concerns, but the large majority the work found in the literature chooses either one of them.

Even so, these are not self-contained separate areas. People working on believable agents, for instance, consider it important to base the internal architectures of their agents on sound emotion theories. Suspension of disbelief during an interaction with one of these agents may depend crucially on the consistency among events and displayed emotions. Similarly, validation of agent architectures based on emotions, for instance through social interaction with humans, may require sound display and recognition of emotions. So, it might be fruitful that people working in each one of these two major areas exchange ideas, models, and implementations.

At a second level, the proposed taxonomy further divides research in a set of subareas. At a first glance, one could attempt a division based on the nature of the emotion theories (e.g., appraisal theories, neurophysiological, and so on). However, such an approach easily results with very diverse research work ending up in the same subarea. For instance, appraisal theories cross-cut horizontally across projects with different goals. For this reason, an alternative approach was chosen. This approach is based on the research goal the authors propose themselves. With this idea in mind, the adopted taxonomy is the following:

- Internal manifestations of emotions
  - Architectures. Research in this area aims at a generic agent architecture where the internal mechanisms of emotions play a prominent role;
  - Robotics. The goal in this area is the construction of mobile robots whose behavior is determined by emotional components in the architectures;
  - Emotions Modeling. This area aims at the creation of models of mechanisms of emotions, not necessarily biologically-inspired;
  - Cognitive Modeling. The research here studies computational models of emotional mechanisms of the brain;
- External manifestations of emotions
  - Believable Agents. The goal in this area is to build interactive agents seeking suspension of disbelief with the user;
  - Affective Computing. Classical computing is based on interaction with the user on a rational basis, while affective computing focuses on affective interaction among users and computers. This includes two aspects: computers recognizing affective states of users, and computers expressing emotional states in a believable way;

The borders suggested by the above descriptions are not rigid. Hence, a few words are in order about what was kept in mind while reviewing the published work. Although the Robotics area can be seen as an application of the Architectures one, a given research work is here considered to belong to the former area whenever not only a physical robot is involved (includes simulation), with at least its kinematic aspects, but also when it is not evident how the proposed architecture can be deployed in different application contexts. The Emotion Modeling area is here distinguished from Cognitive Modeling one by the object being modeled: the latter aims at modeling cognitive mechanisms in humans, by the means of computational models, while the former one is here understood in the context of (abstract) artificial models of emotions. These models may or may not be biologically inspired. It may also become difficult to distinguish clearly between Believable Agents and Affective Computing: while the latter centers its attention on the emotional content of the interaction, the former uses emotions as a component, among others, contributing to the goal of believability.

3.5.1 Architectures

The goal here is to build an agent architecture where emotional mechanisms play an essential role. These architectures are usually generic enough so that they can be applied in a broad range of contexts.

One of the earliest architectures that accounts for emotional processes was proposed by Aaron Sloman [174, 175, 177]. It captures several architectural trends in the A.I. field. On the one hand, it is based on the classic "triple tower" architecture: perception, processing, and action. And on the other, there are three horizontal levels of processing: reactive, deliberative (planning, deciding, scheduling, etc.), and meta-management (reflective). By crossing the former three vertical "towers" with the latter horizontal layers, one reaches an architecture formed by the nine slots. This is generic enough to cover a broad range of approaches found in the agent architecture literature, but falls short on explaining how exactly all those processes interact in order to produce a coherent behavior. The emotional mechanisms are represented by an alarm system, functioning in parallel. The alarm system monitors activity throughout the architecture. The role of the alarm system is to detect certain salient events, corresponding to emotional states. These events can have origin on either the internal state of the agent, or on external events. The alarm system may respond to those events by provoking changes in the processes running in the system. For instance, processes may be interrupted, while others launched. This point of view is inspired by Herbert Simon's early ideas about emotions as interrupts, as reviewed in section 3.3

above.

During the early eighties, Masanao Toda developed a conceptual model of an artificial being which he called Emotional Fungus-eater [184]. The thought scenario is a team of such fungus-eaters mining for uranium on a distant planet. The idea consists of reducing mental phenomena to a set of *urges.* An urge functions as a motivational sub-routine. These urges range from emergency urges, such as "fear", whenever the fungus-eater encounters an object that may jeopardize its survival, and "startle," when an unexpected detection of a potentially dangerous object occurs, all the way to social urges. Examples of social urges are the "rescue" urge, when a mate fungus-eater needs help, and the "gratitude" urge, e.g. by the one being helped in case of having been rescued. More sophisticated urges include "rule observance" urges, for maintaining some kind of social structure, and the "confirmation" urge of social status, assuming a social hierarchy based on power demonstrations (e.g., the "demonstration" urge). Even though the names of these urges may suggest their identification with emotions, Toda states that there is no one-to-one correspondence among the fungus-eaters urges and human emotions, although some of their names were borrowed from them.

Michel Aubé worked on Toda's theoretical framework, and developed a conceptual architecture for an emotional agent [11, 12]. Aubé first divided Toda's urges in two distinct groups: needs, and emotions. Needs corresponds to basic urges, such as hunger, thirst, and fatigue, while emotions encompass social and interactive urges, such as anger and guilt. The former are satisfiable directly via first order resources, such as food and water, while the latter correspond to second order resources, which are only obtainable by commitments with other agents. Toda identifies the first order urges (needs) with autonomy. Effective management of these basic urges is essential for an agent to be autonomous. The second order urges (emotions) relate to social interaction. Restricting emotions to the social domain is an arguable claim, which Aubé discusses in his publications [11, 12]. Aubé's central idea is to view emotions as commitment operators: emotions form attachments among agents, and therefore an emotion implies a commitment among agents. These commitments constitute the abovementioned second order resources. The *raison d'être* for considering needs and emotions w.r.t. resources comes from viewing the management of needs as the management of basic resources for survival (e.g., food, water, protection), while the management of emotions as the management of commitments among agents. Aubé's architecture is based on two layers: an emotional layer, encompassing emotions and commitments management, and a needs layer, containing the management of first order resources and needs. According to Aubé's theory, the needs layer is capable of triggering emotions, while the emotions layer performs regulation of the needs.

Rolf Pfeifer proposed an A.I. model of emotions called FEELER [146, 147]. FEELER is based on a set of production rules, such as the following one (example from [146])

```
R1:    IF      current_state is negative for self
               and emotional_target is VARperson
               and locus_of_causality is VARperson
               and locus_of_control is VARperson
       THEN    ANGER at VARperson
```

The rules are based on an emotions taxonomy proposed by the psychologist Bernard Weiner in 1982. The current situation is represented on a pair of working memories, a cognitive memory (to store structures such as plans and goals), and a physiological memory (to represent physiological activity). In the simpler case, left side conditions refer to the current situation stored in the working memories, while the right side triggers specific emotions, possibly containing variables (e.g., `VARperson` in the above example) about the target of the emotion. This allows for feedback, in the sense that the elicitation of certain emotions can trigger the occurrence of other ones. The model also accounts for interrupts, for instance, in the case of sub-goal violation, which triggers the rule production system, and for dynamics of the emotions (e.g., for how long an emotion should last).

In his later research, Pfeifer departed from this methodology, motivated by the recognition that emotion should be an emergent phenomenon, rather than engineered into the agent as in FEELER [147]. Pfeifer then followed a "Nouvelle A.I." approach, pioneered by Rodney Brooks [34, 35]. Pfeifer's methodology followed a path of constructing robots exhibiting emergent behavior. According to Pfeifer, at a sufficient level of sophistication, emotions will become identifiable as an emergent phenomena, given that agents are provided with some basic built-in mechanism(s) allowing them to discern what is 'good' for them from what is 'bad'.

As in Pfeifer's FEELER, the Salt & Pepper architecture of Luís Botelho and Helder Coelho also uses a rule production system for emotion elicitation [29, 27, 28]. However, this architecture differs from FEELER in several aspects. The appraisal process is divided in two parallel processes, driven by two distinct set of feature extraction processes: a cognitive appraisal, dealing with explicit representations of cognitive structures such as plans, goals, and so on, and an affective appraisal, which is responsible for the generation of emotional signals and responses. The authors found inspiration in neurophysiology (e.g., research by Damásio [55] and LeDoux [112]) for this double representation scheme. The affective appraisal is implemented by a production system. The emotional responses generated by the firing rules affect the internal state and possibly are capable of triggering actions. Internal effects of emotions include changes in motivators (goals), and attention shifts, among others.

Alastair Burt's approach takes a three layer agent architecture (INTERRAP) as a stating point, consisting of a Behavior-based (lowest level), a Local Planning, and a Social Planning layers (highest level). Burt views emotion as a mechanism for internal resource management [37, 36]. The agent architecture modeling is based on logic. However, according to Burt, logic does not suffice for an agent to function, a motivational component is also required. This motivational component is responsible for tasks such as managing goal generation, goal weighting, and for the way the layers influence one another and lead to action generation. Agents with bounded resources require a scheme for managing these resources: the emotions.

Basing his framework on Minsky's Society of Mind [133], Juan Velásquez proposed the Cathexis architecture [193, 195, 194]. The two main components of Cathexis are the emotion generation and the behavior system. The former consists of a networked architecture of proto-specialists[2], one for each emotion family (e.g., fear, disgust). Each one has a scalar value representing its activation level. They receive inputs from emotion elicitors (external stimuli), as well as from other proto-specialists (with either excitatory or inhibitory gains). Each one of them has also a decay function that fades out the activation level, in the absence of input activity. A proto-specialist is further subject to a thresholding for saturation at a maximum level, as well as to a minimum above which the corresponding emotion is activated. The behavior system is based on a network of behaviors. These can be activated by releasers which, depending on the specific behavior, can be triggered by emotions or by external stimuli. Velásquez developed two implementations of this architecture: Simón and Yuppy. Simón [193] depicts a cartoon baby-face showing emotional facial expressions, and has a set of basic behaviors (e.g., sleep, eat, laugh, kiss). The stimuli are a set of graphical interface inputs, such as sliders and buttons. These inputs act not only as stimuli, but they also modify internal physiological variables of the agent. Yuppy is a physical robot [194] with various sensors, such as CCD cameras, microphones, and air pressure sensors. The above architecture was augmented for Yuppy with an emotional memory structure (comparable to Minsky's K-lines[3]). Little details are given on the mechanisms of this emotional memory. The pro-

[2]In Minsky terminology, proto-specialists are simple genetically encoded agents, specialized at innate basic needs, such as thirst, hunger, warmth, and so on [133].

[3]K-lines are memory structures that, during recall of an event, the mental state expe-

posed goal is to replicate Damásio's secondary emotions [55]: associating an emotional state with external stimuli.

Alexander Staller and Paolo Petta proposed an architecture (TABASCO) for situated agents based on appraisal theories of emotions [179]. TABASCO builds on a pair of three-level architectures for perception and for action. On the perception side, where the appraisal process resides, there are a sensory level (feature detection), a schematic level (schemata, e.g. associative network with spreading activation), and a conceptual level (reasoning based on knowledge and beliefs, e.g. BDI [154]). The three levels of the action side are based on Bonasso's 3T architecture [26]. The three layers of the 3T architecture are the deliberative one (planner), the sequencer one (based on reactive action packages), and the reactive one (sensori-motor processing). Mediating between the perception and the action sides there is an appraisal register [178], combining the appraisal outcomes from the perceptual layers, and influencing the three layers of the action side. Finally, there is an action monitoring mechanism sending results from the monitoring of the agent actions to the perceptual side, in order to be integrated with the appraisal process.

The TABASCO target implementations include a graphical/3-D setup, and a text-based environment. In the former case, the setup is based on a virtual mirror (camera and projector, similar to the ALIVE project [123]): a virtual character interacting with the user, with the character's emotions being expressed in both behavior patterns and texture changes [144]. This setup was built for a permanent exhibit at the Vienna Museum of Technology (Austria).

The text-based environment, a modified version of the architecture resulted from blending the agent architecture JAM [96] with components from TABASCO. An appraisal mechanism was added to JAM, receiving beliefs from JAM's world model, and generating (impulse) goals to JAM's intention structure, as well as appraisal outcomes to JAM's plan library [145]. The environment consists of a 2-D grid, where a society of agents is deployed, looking for scattered food. When an agent detects food being eaten by another agent, it can aggress it (the stronger one takes the food). Several strategies are possible, for instance: aggress always, aggress only if stronger, and social norms that assigns food to the ones that find it first.

---

rienced at that event is reconstructed [133].

3.5.2 Robotics

This section addresses agent models designed specifically for mobile robots, including both physical robots as well as simulation. The bottom line is the constraints imposed to the architecture by the kinematic structure of the robot and/or the spatial structure of the environment.

Sloman's architecture has been implemented, for instance, in Luc Beaudoin's NML1 [22] and Ian Wright's MINDER1 [216] agent implementations, in the context of their PhD theses (both under Sloman's supervision). The environment, common to both implementations, is a 2-D space, enclosed by walls, where the agent moves around and interacts with objects in it. There are static objects, such as fences and ditches, and simple agents that just wander around (minibots). The task is to "take care" of those minibots, such as trying to keep them out of the ditches, as well as charging their "batteries." The MINDER1 agent, which builds on NML1, implements aspects of all layers of Sloman's architecture (reactive, deliberative, and meta-management). The agent keeps a set of concerns (inspired by Frijda's theory [79]), that give rise to the generation of motives. These motives are filtered, according to a resource allocation scheme. A reactive planner is responsible for the agent's actions. The meta-management layer comprises the management of the motive filtering, and a detector of perturbances. Whenever the rate of motive rejection is above a certain threshold, a perturbance state is detected. Such states could be used to control management processes in other parts of the architecture, however, this mechanism was not implemented. It was claimed by the authors that these perturbances states can be identified with human emotional states, in the context of Sloman's framework.

The early work of Dolores Cañamero on emotions can be classified in the area of architectures [40], although her later work shifted towards the field of Affective Computing. In the former area, she developed an agent architecture for a robot living in a grid-world. This environment contains objects of several classes: inanimate blocks (obstacles), food and water resources, and living agents (Abbotts) implementing the proposed architecture. The architecture is rooted in the Society of Mind framework, proposed by Marvin Minsky [133]. Following this framework, an agent is formed by a society of simple specialized agents, where from the interactions among them, intelligent behavior ought to emerge. An Abbott agent is therefore constituted by a society of several agents of different classes. There are *sensor* agents, devoted not only for processing external stimuli, but also to monitoring the agent's body state (somatic sensors). The body state is composed of several physiological variables (e.g., adrenaline, blood pressure, energy, pain, etc.). There are also *recognizer* agents that process sensor data and perform object recognition, *direction-neme* agents that specialize on certain physical directions (e.g., top, top-left, left, etc.), *map* agents that construct a map of perceived objects found in the environment, *effector* agents that perform actions upon the environment, *behavior* and *manager* agents that are specialized on certain behaviors, *motivation*, and *emotion* agents. These two latter classes of agents are responsible for the emotional component of the Abbotts.

Cañamero distinguishes motivations from emotions in the following way: while the former aim at the homeostatic equilibrium of the body's physiological variables, the latter are able to amplify (or modify) the motivational state of the agent. Emotional states can be activated by either external events (e.g., achievement of a goal elicits happiness) or internal patterns of physiological variables (e.g., sustained high level of a variable provokes anger). The effects of an emotion are twofold: first, it can modify the intensity of current motivations, leading to modifications in the behaviors intensities, and second, it can modify readings of sensors that monitor the variables that emotion can affect.

Matthias Scheutz approached emotions from a bottom-up perspective. The goal is to develop simple affective states in a multi-agent setup [166]. He implemented an agent architecture living in a simulated 2-D environment. The agents are robots that move around, looking for food and water resources, and avoiding obstacles, as well as other robots. The agents are evolutive, in the sense that they live until they either collide or run out of energy. After some pre-determined time after inception, they give rise to an offspring of brand new agents/robots. Each robot has sonar sensors to detect other objects from a distance, smell sensors to detect water and food, and touch sensors to detect close proximity to other objects. When a touch sensor is activated, an alarm system triggers a reflex action to make the agent move away from the touched object. The motors are driven by a schema-based controller [4]. The sensors are able to identify and localize the four types of objects in the environment: food, water, obstacles and other agents. When detected, an object will produce a force vector, depending on the kind of object and on its relative direction with respect to the agent. Using a linear combination of these vectors, a direction of movement is computed, resulting in motor commands. The weights used in this linear combination are crucial. Depending on their sign, they can make the agent be attracted to or be repulsed from a specific kind of objects, with more or less intensity depending on the weight magnitude. These weights are the outputs of a neural network, which constitutes the affective system of the agent. The inputs of this neural network are: internal water and energy levels ("physiological" variables), and alarm signals triggered by close proximity of obstacles or other agents.

The motivations for these agents are to look for food and for water. The intensities of these motivations are determined by the weights that come out from the neural network. So, these values are said to represent the motivational state of the agent.

In summary, the internal "physiological" variables and the alarm signals triggered by external events are processed by the neural network, resulting in the weights that determine the behavior of the agent. Thus, the neural network parameters span a parameter space where a broad range of agent "personalities" can be found.

The parameters of the neural network are subject to evolutionary mutation. The number of agents in an offspring depend on the energy level of the progenitor at the time of the reproduction. Therefore, the surviving agents tend to be the ones with the neural network parameters most fit for survival on the given environment.

The experiments consisted of runs of a certain amount of simulation steps each. The network parameters of the survival agents were then analyzed. According to Scheutz, two classes of affective states can be identified [166]. One class corresponds to the signal path starting at the levels of the "physiologic" variables, which can be identified with the agent drives, e.g. a low level of water makes the agent being more attracted to water resources. A second class corresponds to the signal path originating from the alarm signals. In this latter case, according to Scheutz, the neural network seems to measure the frequency of encounters with certain kinds of objects, suggesting the emergence of emotional states. Scheutz identifies these states with primary emotions, such as fear and anger.

More recent research by Scheutz shows a shift of concern towards the analysis of agent architectures, in terms of methodologies to evaluate the utility of emotions [168, 167]. Scheutz's approach consists of a systematic analysis of a given agent architecture in terms of performance-cost tradeoffs, and the impact or having or not emotions. He introduces the idea of cost induced by the architecture [167]. He first maps agent architectures to a common abstract framework — APOC — that models an agent's internal structure in terms of various kinds of components and links [2]. The cost induced by an architecture results from the combination of structural costs, process costs, and action costs.

Building on the animat paradigm, Sandra Gadanho and John Hallam proposed an architecture targeting a physical robot (Khepera). Their architecture consists of an adaptive controller, based on the Q-learning paradigm [181], driven by an emotional system [85, 82]. The two main architecture components are the emotional system and the adaptive controller. The environment is a 2-D maze with energy sources (represented by light sources) scattered around. In order to obtain energy, the robot has to bump into those light sources. There are additional characteristics that complicate this task, thus requiring the robot to perform sequences of actions to avoid running out of energy over time.

The emotional system implements four valenced emotions[4] (happiness, sadness, fear, and anger), and seven feelings (hunger, pain, restlessness, temperature, eating, smell, warmth, and proximity), resulting in two numerical vectors of four and seven components. Although they correspond to the names of human emotions, the choice out of the full human spectrum was based on whether they made sense in the given environment. Moreover, there is a hormonal system consisting of a dynamical system with a state vector of seven components (corresponding to the seven feelings). This dynamical system implements attack and decay rates of each of the state vector components. The feelings vector is obtained by summing up the hormonal state vector (weighted) with the sensations vector. This sensations vector is obtained from the robot's perception (internal and external), e.g. the hunger component corresponds to the robot energy deficit, and the pain component is high whenever the robot bumps into something. The emotions vector is computed from the feelings vector by using a neural network with handcrafted weights, reflecting how specific feelings influence certain emotions. For instance, low energy, and not acquiring it, makes the robot sad, and even sadder if no light is sensed; bumping into obstacles provokes pain, inducing fearfulness, but less fear if the robot is hungry or restless. The emotions close the loop by activating the hormonal system (through another neural network).

The adaptive controller is responsible for selecting a behavior out of a set of three options: avoid obstacles, seek light, and follow walls. Three neural networks, one for each behavior, are used to estimate their Q-values. The final behavior is probabilistically chosen using a Boltzmann-Gibbs distribution over the Q-values. The reinforcement value is obtained from the dominant emotion, *i.e.,* the highest component of the emotions vector. Recall that the emotions are valenced: happy is positive, and the others are negative. This reinforcement value is used to train the neural networks using the Q-learning algorithm [181]. Thus, the neural network aims at estimating the Q-values, taking into account the reinforcement as well as the expected discounted cumulative reinforcement for the optimal policy.

For the system to function in a continuous time environment, there is the need of an event detector responsible for triggering the learning, as well as keeping each behavior for sufficient amount of time so that its effects can be

[4] The valence of an emotion amounts to its positiveness or negativeness value. For instance, happiness is said to have positive valence, while sadness a negative one.

felt. For this purpose, two event detector mechanisms were tested: event-triggered (significant changes in the emotional state), and interval-triggered (fixed interval).

The authors performed several experiments in the above-mentioned simulated environment, using a Khepera robot simulator. Several variations of the architecture were experimented with, where certain emotional components were replaced by non-emotional counterparts. The results of the comparison showed that the emotional version of the architecture performed better in terms of collisions and number of events. The authors attributed this result to the event-triggered mechanism, which depends on changes on the emotional state.

Sandra Gadanho continued working on this architecture during her post-doctoral research studies. Enhancements on the architecture include modifying the emotional system in such a way that the goals are made more explicit, by using a set of homeostatic variables [84]. Further research also includes augmenting the architecture with an adaptive rule-based cognitive system, with the goal of a cognitive level decision-making mechanism [83].

Piotr Gmytrasiewicz takes a decision theoretic approach to emotions [88, 68]. Decision theory states that the optimal course of action a^* from a set A of possible actions, given an utility function $U(s)$ mapping the set of possible world states S to real numbers, is given by the following equation equation[5]

$$a^* = \arg\max_{a_i \in A} \sum_{s \in S} p_i^j U(s^j) \tag{3.1}$$

where p_i^j is the probability of finding the world state s^j after performing action a_i. The agent's knowledge of the world state is thus assumed to be probabilistic. Given a current state probability distribution $P_c(s)$ over the state space S, together with an action a_i to be performed next, there is a function that maps this distribution onto a next state distribution $P_i(s)$, where consequently $p_i^j = P_i(s^j)$. This function is designated probabilistic temporal projection. The quadruple formed by this projection, together with the current state distribution $P_c(s)$, the action set A, and the utility function $U(s)$ is termed a decision-making situation.

Transitions between emotional states are represented by a finite-state machine (FSM) whose transitions depend on environmental input. The central idea of Gmytrasiewicz's approach is that each state can perform transformations on each one of the components that form the decision-making situation quadruple of the agent. Namely, transformations on the action space A, on the utility function $U(s)$, and on the state probabilities are discussed [88]. An

[5] Assuming a properly formed utility function $U(s)$.

example of a transformation of the action space is the narrowing of possible actions. This can be compared to Frijda's action tendencies [79], when there is a predisposition for a small subset of actions. Emotions and feelings are said to implement and modify the utility function, for instance. And finally, transformations to the probabilities of states, for instance, by simplifying them, can allow the agent a quicker decision-making process (e.g., taking into consideration the most probable next state only).

Experiments were performed in a Wumpus world scenario [161], which resulted in interesting preliminary results [68]. The Wumpus world consists of a 2-D grid world inhabited by a moving creature called Wumpus, which kills the agent whenever they collide. The goal of the agent is to grab a piece of gold located somewhere in the grid, while avoiding being killed by the Wumpus. The agent can move in all four directions, as well as collecting gold wherever it finds it.

The agent decision-making used in this case was based on a partially observable Markov decision process (POMDP) approach [41]. Computing the optimal policy for a POMDP is computationaly hard. Therefore, methods to trade-off computational complexity against solution quality are interesting from a pragmatical point of view. The proposal Gmytrasiewicz brings forward is that emotions can provide such a mechanism. The idea is to use emotions to simplify the POMDP model, hence facilitating the task of finding the optimal policy.

Each one of the emotional states considered in the FSM — contended, elation, fear, and panic — perform transformations on the POMDP model: a contended agent takes the maximum possible time in computing the optimal plan; fear makes the agent restrict the time horizon of the planning process; panic is like fear, but additionally reduces the action space to movement; and during elation the agent attempts to reach the goal regardless of any danger of colliding with the Wumpus, *i.e.,* the utility function is insensitive to the case of such a collision.

Along with the described agent, two more agents were tested in several randomly generated Wumpus worlds: a cut-down non-emotional version, and one always in the panic state. The performance was evaluated by measuring the average of the accumulated rewards (of the POMDP) over a number of runs, in a set of three kinds of worlds, differing in the speed of the Wumpus. The simulations took into account the time the agent takes to perform planning using the POMDP model, and thus the speed of movement of the Wumpus has an impact on the performance. The preliminary results show that, while the non-emotional agent performs marginally better with a slow Wumpus, the emotional one exceeds it as the Wumpus moves more quickly. Moreover, the agent that is always in a state of panic did not perform better in any scenario. These results suggest a clear benefit of using emotions in such a rational decision-theoretic agent model.

Luís Morgado and Graça Gaspar use a signal processing approach to emotions [134, 135]. The emotional disposition of an agent is modeled by a two dimensional dynamic vector $ED \equiv (\delta P, \delta C)$

$$\delta P = \frac{dP}{dt} \qquad \text{and} \qquad \delta C = \frac{dC}{dt} \tag{3.2}$$

where the components are the time derivatives of two quantities: the achievement potential P, representing the potential the agent is able to produce in the environment, and the achievement conductance C, representing how much the environment allows for that change to occur. The position of the ED vector in 2-D space defines an emotional quality tendency, namely quadrant I (both δP and δC positive) means Joy, quadrant II (δP positive and δC negative) means Anger, and quadrants III and IV mean Fear and Sadness. In other words, the rates of change of the achievement potential and of the conductance determine the agent emotional quality. On top of this structure, they define cognitive elements as vectors in a multi-dimensional space. These elements can play the roles of motivators, mediators (transform motivations to actions), achievers, and observations (related to perception). Experimental case studies include, for instance, a robot moving in a 2-D space, looking for food, and avoiding obstacles [134]. The results show that not only the robot is able to avoid colliding with an obstacle, but also that it exhibits adaptive capacity, showing improvement along successive runs in the way the obstacle is efficiently contoured.

3.5.3 Emotions Modeling

The approaches reviewed in this section address models of emotional mechanisms, including computational models, but also mathematical ones. This accounts for models that do not necessarily mimic the biological counterparts, but rather aim at incorporating a set of characteristics found in emotions.

Zippora Arzi-Gonczarowski employs mathematical Category Theory [110] as a tool for modeling affective phenomena, although the scope of her model is broader [9, 6, 7]. Her goal is to provide a rigorous mathematical formulation of the field. Category Theory, first introduced in 1945, is an abstract algebraic theory that unifies a broad range of mathematical constructs. Its principles are very simple: a category is composed by a set O of objects, a set A of arrows (also known as morphisms) associating ordered pairs of objects (notated as $f : a \to b$ for $a, b \in O$ and $f \in A$), and a set of properties these objects must conform to (such as associativity), giving it a minimalist mathematical structure. Many other constructs can be built on top of this minimalist definition (e.g., set theory, algebraic topology, logic). Arzi-Gonczarowski approach builds on a category of perceptions [10], where objects are tuples $\langle \mathcal{E}, \mathcal{I}, \rho \rangle$ (perceptions), and morphisms are transitions from one perception to another. World objects, exterior to an agent, are represented by a set $\mathcal{E}_0$, which is assumed to be common to all perceptions of the category ($\mathcal{E} = \mathcal{E}_0$). In a perception, the set $\mathcal{I}$ contains internal representations (connotations), and the mapping ρ associates the set of exterior objects $\mathcal{E}$ and the set of the agent connotations $\mathcal{I}$ by the means of a three-valued logic $\rho : \mathcal{E} \times \mathcal{I} \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$. For a world element w and a connotation α, a logical value of $\mathbf{t}$ means that w has connotation α, $\mathbf{f}$ means that it does not, and $\mathbf{u}$ means that it is unspecified (at least at the moment). Morphisms represent transformations from one perception to another. Several standard category constructs can be mapped to this category of perceptions [10].

This mathematical background has been used by Arzi-Gonczarowski to model several aspects of affective phenomena. Connotations can model agent reactions to certain stimuli. If α denotes a reaction, $\rho(w, \alpha) = \mathbf{t}$ means that the perception of w provokes a reaction α. Changes in perceptions are modeled by morphisms. An emotional response may include, for instance, changes in the way the agent reacts to perceptions (moods and attitudes). The theory was further extended, for instance, to explicitly include behaviors and action tendencies into the perception tuples. Other extensions include perceptions of perceptions, by allowing the world set $\mathcal{E}$ to be made of other perceptions. This construct was said to constitute self-reflection [8].

As explained in section 2.3.1, cognitive appraisal means the evaluation of an event, with respect to an agent's goals and expectations. The approach of Jonathan Gratch is based on the idea that classical planning algorithms provide a level of representation adequate for the development of cognitive appraisal models [92, 93]. Such a level of representation allows for what he calls a plan-based appraisal, that determines the significance of events in terms of the successful execution of a plan. For instance, an event may threaten a precondition of some step of a plan, and thus jeopardize the achievement of the goal. The appraisal process is based on construal frames (from Elliot's construal theory [74]). Events are matched against these construal frames, and emotions are elicited according to the OCC theory [142]. Emotions exert their influence on the agent's communications, planning process, and action selection. To illustrate these ideas, an implementation called Émile was devised. Émile agents dialogue with each other in natural language, each one with different personalities parameters, all trying to achieve their own goals (e.g., one wants to make money, while another wants to go surfing).

Together with Stacy Marsella, Gratch extended his research to include coping behavior [126, 94]. They devised a set of strategies the agent uses to cope in response to the appraisal of events. For instance, if a possible future event has desirable effects, a Planning strategy is used to assert the intention of that event to occur; or if an intended future state seems unachievable, an Acceptance strategy is called for to retract from that intention [126]. Their model was subject to an evaluation, where the results form their model and from a group of human subjects was compared [94]. The evaluation of the human subjects was performed by means of questionnaires about the subject's feelings, as well as the way they appraised the situation. Two scenarios were devised, one based on verbal accounts of a situation, and another where subjects watched a video clip of a conversation in a virtual reality setup. The authors claim that their model showed consistent results with the subjects answers.

Aiming at believable agents and interactive characters, Ian Wilson proposed the Artificial Emotion Engine to model affective behavior [211]. Affective behavior is represented at three levels of prominence: momentary emotions (high priority, short span of time), moods, and personality (low priority, longest span of time). The extensive research of the personality psychologist Hans Eysenck was used to define personality traits in a three dimensional space (Extroversion, Fear, and Aggression). The engine receives at its input punishment or reward signals for each one of the agent needs, as well as signals from sensors of its emotional reactions. At its outputs, the engine produces signals at different levels: body and facial expressions, the action plan to satisfy the agent needs, and the raw agent emotional state.

3.5.4 Cognitive Modeling

This section reviews computational models that explicitly aim at modeling emotional mechanisms from a cognitive point of view. Contrarily to the previous section, the following approaches aim at modeling affective phenomena as they occur on humans.

Taking a physiological perspective, Christian Balkenius proposed a computational model of emotional learning and processing [17, 16]. The approach aims at modeling several areas of the brain related with emotions, at a functional level, rather than at a neural level. The areas modeled are the amygdala and the orbitofrontal cortex, as well as the interactions among them, at a simplified level. The resulting system receives sensory input and a reinforcement signal (reward/punishment). The amygdala learns to respond to a stimulus with an emotional response of the same magnitude as the reinforcement signal. The orbitofrontal cortex compares the actual and the expected reinforcement signals, coming from the amygdala. If they do not match, learning is activated such that the orbitofrontal cortex is able to anticipate the emotional response of the amygdala. The orbitofrontal cortex uses this learning process to be able to inhibit emotional responses of the amygdala prior to any reinforcement signal. The amygdala and the orbitofrontal cortex receive sensory input with different granularities, where the orbitofrontal cortex employs finer input resolution. Thus, it is capable of a higher degree of discrimination. The simulations showed interesting results at several aspects of classical conditioning, such as habituation, acquisition, extinction, and blocking. Moreover, the model reproduced many effects caused by lesions in certain brain regions.

Dietrich Dörner and Ulrike Starker developed a model of human action regulation (Psi-model), integrating cognition, motivation, and emotion [67]. The key idea is to view emotions as a controlling system that takes into account the degree of uncertainty of the environment, as well as the organism's competence to tackle problems. Two examples: uncertainty leads to safeguarding behaviors, and low competence makes the organism avoid too difficult problems. The cognitive processes depend on an arousal parameter, related with the general preparedness for action, and a resolution level, which regulates how deep planning and perception processes should go. They devised an experiment using complex and dynamic maze-like environments, and subjected both humans and agents based on the Psi-model to it. Several performance metrics were used, such as the number of places visited, the number of breakdowns, the number of "nuggets" collected, among others. The results showed similar profiles across these metrics among human and machine subjects. They also tested the Psi-model with and without emotions. The emotional version showed better results in general, namely in terms of agents succeeding to preserve themselves (number of breakdowns), and of the number of collected "nuggets."

Drawing on neurophysiology, Jean-Marc Fellous proposes a view of emotions as dynamic patterns of neuro-modulation, rather than patterns of neural activity as it is traditionally viewed [75, 76]. Many neurological studies point towards an active involvement of many chemical substances (neuromodulators) in the dynamics of emotional states [114]. These chemical substances, elicited by emotional states, are able to modulate the functioning of the nervous systems in many ways. Emotions can thus be seen as a mechanism of conveying information across the brain. Although the nature of this information is poorer than in other cognitive capabilities, it has an high impact. Fellous proposed an organization of behavior on four levels: Reflexes (e.g., knee jerk), Drives (e.g., rage), Instincts (e.g., fear conditioning), and Cognitions (e.g., learning). These levels present an increasing potential of neuro-modulation susceptibility, from the reflex level, where this potential is small, up to the cognitive level, where the potential of neuro-modulation is the highest. Fellous defends that there is no emotion center in the brain, but rather that emotions involve the whole brain. Moreover, he denies any causal interdependence among cognition and emotion.

Together with Fellous, Michael Arbib put forward a proposal of emotions as a mechanism of arbitrating among several modes of behavior [3]. Four major modes of behavior (the four Fs) are widely acknowledged as universal: feeding, fighting, fleeing and reproduction. For a robot in an ecological niche of different nature, a different set of modes may be more appropriate than these four. Each mode encompasses a group of tasks. Arbib and Fellous view motivations as a bias to choose one mode over another, *a priori*, in a given situation. Moreover, emotions are considered an evaluative process of the consequences of that choice. Emotions are then capable of switching modes whenever appropriate.

Eva Hudlicka proposes a computational cognitive-affective architecture (MAMID) [97, 98], where the underlying idea is that affective states, together with personality traits (individual differences), manipulate a series of architectural parameters, such as the processing speed and capacity of a set of cognitive modules. For instance, anxiety provokes reduced attentional and working memory capacities. The MAMID architecture is based on six modules which process perception, in sequence, resulting in actions at the end. These modules are: attention, situation assessment, expectation generation, affect appraiser, goal manager, and action selection. The affect appraiser mechanism is based on multiple-levels and multiple-stages appraisal theories (see section 2.3.6). Stimuli are processed at two levels: a low-resolution level in terms of valence, and a high-resolution one in terms of four basic emotions (anxiety/fear, anger, sadness, and happiness). The appraisal process consists of three stages: an automatic appraisal centered on the properties of stimuli, an expanded appraisal which is centered around the agent's internal motivational context (including the agent's goals and expectations), and a current state modulator providing a smooth variation of emotional values along time (ramp-up and decay). The affective state has influence on various parameters of the cognitive architecture, namely on the goals and actions selection rules, on the speed and capacity of the modules, and on the ranking of the mental constructs.

The proposed model was evaluated in a peacekeeping simulation scenario, employing three types of agents (anxious, aggressive, and normal). The agents were exposed to a series of surprise situations (e.g., a destroyed bridge, an hostile crowd) capable of eliciting different responses depending on the agent traits. The results confirmed that appropriate behaviors were performed, according to the types of agents [98]. Although the agents performed according to design, Hudlicka aims at further validating the results by comparing them with empirical studies with human subjects.

An enhancement of the MAMID was proposed by Hudlicka to model meta-cognition within the framework of the architecture [99]. The idea is to model the Feeling of Confidence (FOC). A FOC attribute is added to each mental construct of the architecture. The architecture is augmented with a meta-cognitive layer, which is responsible for monitoring the cognitive processes, as well as performing control on the architecture at a meta level. The FOC attributes reflect the confidence in the corresponding mental constructs. Moreover, for each construct type, there is a threshold value: if the FOC is above that threshold, no further processing is done, since the level of FOC is considered adequate; otherwise, the meta-cognition layer is called in an attempt to either increase the FOC, or shift strategies entirely. Emotional states and traits have influence on the threshold levels used throughout the architecture.

Taking inspiration on the architecture of the human Autonomous Nervous System (ANS), Christine Lisetti proposed a neural network that captures emotion processing at the physiological level [117]. As with the ANS, the neural network divides into a sympathetic and a parasympathetic subsystem. The former is responsible (in humans) for emotional states such as anger and fear, while the latter has an antagonistic effect (e.g., calming down). The model targets several physiological organs such as the brain, throat, heart, stomach, and so on. Experimentation utilized neural networks, implementing a Boltzmann machine model, and using an Hebbian rule for weight updating.

Although still in its early stages, interesting research is being conducted by Licurgo de Almeida and colleagues, with the aim of developing a physiological model of the body [63]. This model is detailed down to the level of organs (e.g., heart, liver, reproductive organ, muscles) as well as the interactions among them (e.g., hormones, circulatory system). The goal is to construct a biologically plausible system of the body, on top of which emotional mechanisms can be implemented. This corresponds to the Damásio's concept of emotions grounded on body states [55].

3.5.5 Believable Agents

The following sections review research on the external manifestations of emotions branch of the proposed taxonomy. The goal of building believable agents is to construct agents that interact with the user in a way to attain suspension of disbelief. In other words, to give the impression that the user is interacting with another being, rather than a machine. This implies that, from an external point of view, the agent acts in a life-like fashion.

The Oz project at Carnegie Mellon University was created with the goal of building interactive believable agents [156, 21, 20, 155]. The architecture underlying the agents is called Tok. It is based on several modules. There is a sensory module which is responsible for processing sensing data from the environment, as well as maintaining an integrated world model. The remaining modules perform queries on this model to gather information about the world state, as perceived by the agent. The two main components of Tok are the Em module, which models the agent's emotions, and the Hap module, which manages the goals and outputs the agent actions. Based on sensed data and on the internal state, the Em module is able to generate emotions. Moreover, the Em module is also sensitive to social relationships with other agents, giving rise to a social aspect in the agent behavior. The Em modules closely follows the OCC [142] model, but several simplifications were introduced. The generated emotions are then communicated to the other modules. The agent's behavior is determined by the Hap module. The Hap module is goal-based, although it does not perform explicit planning: goals are divided into canned sequences of sub-goals and/or actions. Goal failures or successes are transmitted to the Em module, leading to the possible elicitation of emotions. The generated emotions influence the Hap module in various ways, thus modulating the agent's behavior. The interface with the world is symbolic. Natural language modules were developed to allow verbal communication. The Oz project led ultimately to the inception of the Zoesis spin-off company[6] which commercialized virtual characters.

Bill Tomlinson, together with Bruce Blumberg and others, developed the AlphaWolves project, consisting of a social environment of a pack of virtual gray wolves [185, 186]. External users can interact with the pack by uttering sounds to a microphone, while the wolves are displayed in a 3-D virtual reality installation. The wolves generate emotional memories by associating emotional states to stimuli sensed simultaneously, following Damásio's Somatic Marker Hypothesis [55]. The next time a wolf encounters the same stimuli, the memorized emotional state is reinstated. Since the system aimed at realism, it was evaluated in the following way: subjects watched a video from National Geographic explaining real gray wolves behavior, and then they interacted with the system. Afterwards, they were asked to fill a questionnaire [186].

Drawing on the OCC theory, Clark Elliott proposed a model of emotions, the Affective Reasoner (AR), with the goal of simulating several aspects of emotion processing in a multi-agent setup [74]. The model matches a given situation against the agent's concerns and personality. Using the

[6] URL: `http://www.zoesis.com/`

OCC theory, an emotion may be elicited, followed by an action generation. Each agent observes the actions performed by other agents, and then formulates structures representing the concerns of other agents, based on the observed actions. This system was evaluated in a multi-agent simulation, using about 40 agents [74]. Later work by Elliott covered emotion expression, using multi-modal output (speech based, morphed faces, and music). The system was given an emotion category and a text, while the other parameters (e.g., speech inflection, faces) were automatically selected. Human subjects evaluated, side by side, the system and an human actor. Interestingly, the results showed an overall better performance with identifying emotion scenarios generated by the system, than the ones performed by the human actor [72, 73].

Chisato Numaoka proposed a self-biased conditioning system, targeted for the design of a personal assistant in a virtual reality setup [140]. The scenario consists of assisting a user navigating in a 3-D virtual world. This world was constructed as an extension to the traditional text-based web pages. The goal of the personal assistant is to indicate potentially interesting navigation possibilities in the virtual world. The approach is based on associative neural networks that learn correlations among stimuli, motor responses, and user preferences.

Carlos Martinho, together with Ana Paiva and colleagues, developed a virtual reality installation for the Expo 98 World Fair, consisting of a pair of virtual dolphins interacting with the audience [127, 128]. Each dolphin has a specific personality. The emotion generation model is based on the OCC theory [142], as well as on the believable agent models from the Oz Project [155]. The main goal is to attain ethologically plausible behaviors. However, this goal sometimes contradicts the needs of the audience to have immediate feedback of the interaction. Hence, a balance among these two aspects is necessary.

The implemented dolphins are called Tristão and Isolda. The former has an introvert personality, thus avoiding contact with others, while the latter is more playful. Tristão has a "crush" on Isolda, so that it has to balance among its introversion tendency and the desire to interact with Isolda. Each character functions according to a four phase cycle. The first phase is designated the *perception phase*, where sensor data is filtered, followed by a world model update. Then, a *reaction phase* follows, where immediate emotional responses are appraised. Third, a *reasoning phase* performs a cognitive evaluation, taking into account the internal world model. Finally, an *action phase* generates the agent actions, based on the results from the previous phases.

Taking a behavioral approach, Ruth Aylett engaged in constructing virtual Teletubbies (based on the well-known homonymous TV series for children) [13]. Together with Carlos Delgado, they targeted collective behaviors of virtual sheep [14]. Simple (sensori-motor) behaviors are combined using a utility value assigned to each behavior (called behavior patterns). Then, these are activated by structures of pre-conditions (behavior packets). These structures can then be organized in sequences of behavior packets (behavior scripts). Internal motivations in the form of a set of drives push the agent's behavior. These drives control a set of priority queues that ultimately lead to a selection of a specific behavior script. Emotions are seen as internal behaviors, that are able to influence the agent's external behavior. There are three levels at which emotions can be modeled: as a behavior pattern, where motor outputs are other behavioral packets, as an internal sensor pre-condition, and as a modulation of behavior scripts.

3.5.6 Affective Computing

The term Affective Computing was coined by Rosalind Picard in 1994, and appeared in print for the first time in a Technical Report from 1995. She defines the concept as "computing that relates to, arises from, or deliberately influences emotions" [150]. This publication was followed by the formation of the Affective Computing Lab lead by her, and by an influential book [148] authored by her. The paradigm of affective computing is to shift the way humans interact with machines, from a traditional rational and deterministic basis, towards an interaction conveying affective content. This implies both the ability of detecting emotional states in the user, as well as conveying affectively loaded content to the user. Most of the initial research work at Picard's group addresses emotional state detection from physiological sensors [210, 152]. In one of the studies, physiological data[7] was systematically collected from an actress (stimulated with visual imagery), and many state-of-the-art pattern classification methods were tested, resulting in high recognition rates (about 80%) over a set of eight emotions. Other sensor modalities used for emotional state classification are, for instance, cameras for pupil tracking, and a sensor chair for posture recognition [104]. An application akin to Picard's group is the design of wearables — clothes with embedded active devices that interact with the wearer — such as an affective DJ that chooses music according to her/his affective state [95], and a glove that lights up a LED depending on the user's skin conductance [151]. Research efforts have also been directed towards what Picard, together with her colleagues at the MIT Media Lab, designate as Affective Learning [149]. The

[7] Four sensors measuring facial muscle tension, blood volume pressure, skin conductance, and respiration volume.

proposal of Affective Learning is to use the paradigm of Affective Computing to develop tools to aid in the learning process of persons. It is well known that affect plays a crucial role in learning, and thus a system sensitive to the user's affective state could promote a more prolific learning process. Moreover, the idea of a learning companion learning a certain subject together with the user, through interaction, is also part of the proposal.

Focusing on the social interaction between robots and humans, Cynthia Breazeal developed Kismet [31, 30, 32]: a robotic head sketching the salient features of a face, such as the lips, eyebrows, eyelids, and ears. Although the presence of these features gives it an anthropomorphic look, the face does not quite resemble a human face. The head is capable of expressing several kinds of emotions, by means of appropriate smooth movement of its parts in coordination. These movements can be represented as points in a three dimensional space. These dimensions represent high/low arousal, positive/negative valence, and open/closed stance. Ekman's six basic emotions [70] are mapped onto this space: joy, anger, disgust, fear, sorrow, and surprise. Kismet uses a microphone to recognize affective intent from the user speech. Depending on its nature and intensity, Kismet may express an emotional posture in accordance. This allows interaction with a user in the form of a dialog composed by user speech conveying emotional intent, answered back by emotion expression by the head.

Building on Ronald Arkin's AuRA robot architecture [5], Lilia Moshkina and Arkin developed a framework for affective robot behavior (TAME), with the goal of an increased ease and pleasantness of human-robot interaction [136, 137]. TAME models personality traits, attitudes, moods, and emotions. Each one of these aspects has different characteristics and impacts on the architecture, e.g. emotions and moods are dynamic, varying according to robot interaction, while the other model components are not. The lower levels of the architecture follow the schema-based paradigm [4], where the specific parameters can be altered by the modules implementing TAME. For instance, the object avoidance gain can be increased whenever the robot is in fear. To evaluate the impact of TAME on human-robot interaction, a study was conducted allowing human subjects to interact with a Sony AIBO (dog-like) robot. Two versions of the robot were tested, with and without emotional behavior. The subjects were asked to answer questionnaires after several sessions of interaction with the robots. Four hypotheses were tested, addressing the ease of use, the pleasantness of the interaction, the recognizability of emotion display, and the impact of the displayed emotion on the subject's own mood. Briefly, a statistically significant preference of the emotional version over the other was found in the ease of use hypothesis, while the other hypotheses showed no significant difference. However, a curious, statistically significant trend was found indicating that female subjects found the emotions expressed by the emotional robot more recognizable (when compared with the non-emotional version) than male subjects.

Cristina Conati developed a probabilistic model of a user while interacting with educational games [45, 46]. The goal is to develop a pedagogical agent that takes decisions about its course of action based on this model of the user. The model is based on Dynamic Decision Networks [161], which extend Bayesian networks with decision-theoretic behavior and with the modeling of changes over time. The user emotional state, as it changes through time, is estimated by looking at its possible causes (e.g., user traits, goals), as well as at its effects (bodily expressions, tracked by biometric sensors). Emotions are modeled according to the OCC theory [142], using a subset of six emotions (joy/distress, pride/shame, and admiration/reproach). The model was assessed in a pedagogical game based on factorization of numbers (Prime Climb). The model was used during a practicing phase, where the pedagogical agent served as an instructor helping the user attaining the game objective.

Chapter 4

Conceptual model

4.1 Introduction

This chapter presents the conceptual model that underlies the research presented on this thesis. The model was originally developed by a group of researchers (including the author) at the Institute for Systems and Robotics (IST) in Lisbon, Portugal. Several publications (reviewed in section 4.7, at the end of this chapter), MSc theses, as well as funded projects, reflect the development of the model along many directions over the subsequent years.

After presenting the scope of and motivation behind the model, it is presented from a conceptual standpoint, followed by a discussion of some hypothesized consequences. The above-mentioned research based on this model is then briefly reviewed, including several extensions and implementations.

4.2 Scope

As discussed in section 3.2, the name of the field Artificial Intelligence (A.I.) refers to both *artificial* systems [173], constructed by humans to achieve a set of prescribed goals, and about *intelligence* [138], which relates to biological entities, namely to human intelligence. Thus, thinking about A.I. appeals for these two perspectives. On the one hand, the need to satisfy a set of design goals, with the available methodologies and techniques, and on the other, the biological inspiration arising from human intelligence.

Taking into account the radical differences between biological (neurons, synapses, neurotransmitters, and a massive parallel architecture) and computational systems (gates, digital circuits, and serial computation), it is natural to assume that best methodologies to attain a given set of design goals with an artificial system ought to be radically different from the ones emerging from biological evolution. For instance, massively search algorithms have proved successful by beating the chess grandmaster Gary Kasparov in 1997 [39]. The machine (Deep Blue) is able to search over 100 million positions per second, while a human player supposedly examines far less positions. The comparison between the serial processing performed by computers, and the parallel nature of the brain [209] is often mentioned to justify such diverse approaches. However, the history of science has shown that, for centuries, biology has provided fertile ground of inspiration for the design of artificial systems. Consider for instance biologically inspired research for the design of marine and aerial vehicles, based on the biology of fishes and birds [188, 130]. The design of such machines differs in many aspects from their biological counterparts, although the study of the latter teaches invaluable lessons. The streamlined design of modern planes holds little resemblance with birds, however, bird-like wings are visible in Leonardo Da Vinci's drawings of flying machines.

Many A.I. researchers have shared similar concerns. On the one hand, a plethora of biologically inspired research can be found in the literature. This research attempts to reproduce in machines aspects only found in natural intelligence. And on the other, there is research oriented towards machines accomplishing well specified tasks, making use of methodologies where biological aspects are, at least directly, absent. However, it would be unwise to assume that these two approaches follow completely independent paths. Often, the initial steps of a new field resort to biologically inspired methodologies[1]. As that field matures, the focus often shifts towards methodologies orthogonal to the initial biological inspiration.

The approach followed in this research does not fall in either of the two above-mentioned approaches in *stricto senso.* Rather, it is based on a biologically plausible hypothesis, while aiming at a self-contained model of autonomous agents [214]. In other words, the proposed model does not attempt to closely mimic the biological counterpart, since the set goal is to come up with a self-contained model formulation.

The biological inspiration for the present research are the emotional mechanisms in the brain. The nature of the related background material reviewed in chapter 2 is largely descriptive. Its purpose is to provide descriptive models of how the mind works, with a particular focus on emotion mechanisms. Therefore, such models are empirical hypotheses supported by experimental data from human subjects, in order to draw conclusions about their validity. When the problem of designing intelligent machines is faced, a different approach is necessarily required, since the goal is not to model the human

[1]The RobotCub project [132] is a good example of what is said here.

mind, but rather to design systems that exhibit appropriate behavior (according to some performance metric) in the environments they are exposed to. Therefore, the goal here is to construct a *prescriptive* model, rather than a *descriptive* one. Choosing a prescriptive model, however, implies certain concessions to the biological exactness of the model, in order to make it as functionally self-contained and self-consistent as possible. Therefore, some aspects from biology were factored out, so that the model ends up being a partial view of biology. In sum, there is a trade-off between biological plausibility and structural consistency of the model.

4.3 Motivation

In the book "Descartes' Error" [55], António Damásio proposes that the mechanisms in the brain behind appropriate decision making crucially depend on emotional mechanisms associated with the body. Extensive clinical studies performed by his research group have corroborated his theory. This view challenged the traditional (Cartesian) view of a rational thought functioning independently of the realm of the body. According to Damásio, the body and the emotional mechanisms are a *sine qua non* condition for intelligence. In this thesis, the problem of appropriate decision-making is addressed, under the light of the biological inspiration provided by the role of the emotional mechanisms in the brain.

One of the first approaches to the study of these mechanisms was advanced by William James, with the counter-intuitive idea that the perception of an emotion followed the emotional response to a stimulus [102]. It was only after neurophysiological studies of the brain, performed by Cannon and Bard, that it was realized that sensory information is relayed by the thalamus, following two distinct and parallel branches in the brain [112]. LeDoux fleshed out the details of these two paths after his extensive study of the fear circuitry [113]. The idea of these two processing paths — termed by LeDoux as the high and the low roads — suggests that stimuli is represented by the brain in two different ways. On the one hand, a rough representation (allowing a quick response) following the low-road, and on the other, a elaborate and complex one (requiring a slower processing) following the high-road. Moreover, according to Damásio, these two levels are not independent, but rather the mental imagery holds a close relationship with the brain's representation of the body state, which is closely related with the low-road of LeDoux. The Damásio Somatic Marker Hypothesis (SMH) states that mental imagery are *marked* with representation of body states [55]. These associations are crucial for appropriate decision-making, according to Damásio.

The first decades of the A.I. field have been dominated by the modeling of high-level cognitive capabilities, such as reasoning, planning, problem solving, among others. In the human brain, this kind of tasks is performed by the cortex, the brain's higher and evolutionary newer levels. The shortcomings of this trend in A.I. surfaced when A.I. systems attempted to link with the real world via sensors and actuators, as in robots. It was realized that the problem of relating symbol systems with sensor data was harder than initially thought, for instance, by the physical symbol hypothesis [139]. This problem, often designated as symbol anchoring problem, has then captured the attention of several researchers [48].

As a response to these difficulties, Rodney Brooks advanced a radical new strategy [34, 35]. Brooks proposed to approach A.I. bottom-up, *i.e.,* starting with simple (reactive) behaviors, uphill towards complex (cognitive) ones. And the simplest ones are reactive, reflex-like, behaviors. This kind of behaviors is performed by the lower and evolutionary older levels of the brain.

In a sense, these two approaches in A.I. can be thought as approaches to intelligence from two opposing directions. These directions can be compared with the two levels of stimulus processing in the brain. In the A.I. community, these two approaches — the so called GOFAI[2], and the *Nouvelle* A.I. — raised much controversy in the field spawning a period of several years since the late 1980's. Between these two approaches lies a deep conceptual gap, concerning representational, methodological, algorithmic, and other issues. Many attempts have been made since then to bridge this gap, often in the form of layered agent architectures (see section 3.5.1).

4.4 Structure

This thesis addresses an agent model formulated taking into account the biological inspiration provided by the role of emotions in decision-making processes. The model departs from the traditional approaches in the sense that it focuses on the representational aspects of stimuli, rather than on the layering of levels. The approach taken here to introduce the model considers first a simple agent [205].

4.4.1 First level: Perceptual

The simplest agent architecture, which one can call intelligent (granting a minimalistic meaning of the word) consists of a sensorimotor map among

[2]Good Old Fashioned A.I.

sensors and actuators. This level is here designated *perceptual level,* since actuation is directly derived from perception, *i.e.,* there are no intermediate levels of representation other than the ones directly derived from perception. At this level, certain sensor reading configurations trigger certain actions to be performed, according to a built-in mapping. It can be implemented, for instance, using a lookup table such that for each possible sensor reading, a corresponding actuation command is performed. The sensorimotor map of this level may either elicit a certain behavior for some stimuli, or may not respond at all to some other stimuli. This level can be genetically evolved in order to ensure the survivability of the agent in a given environment niche. Some properties shown by an agent with such level implemented are: simplicity, fast response, and robustness. This level is simple and fast because it can be implemented with a simple sensorimotor map (e.g., lookup table, neural network, etc.). The agent design encodes how the agent shall react to certain situations. This hard-wired encoding includes both the detection of situations as well as the actions elicited. Robustness, here considered in the sense of coping with varying environmental conditions, without explicit world representation, is a consequence of the reactive nature of this level. Properties of this kind of systems has been thoroughly discussed in the literature about reactive agents, e.g., Brooks [34, 35] is accounted for being the precursor of the idea of reactive agents, dismissing any explicit representation and reasoning about the world, and Kaelbling [157] has further developed these concepts, including formal approaches. Many publications following the reactive agents paradigm (*Nouvelle* A.I.) can be found in the literature.

4.4.2 Second level: Cognitive

Increased environment complexity on one hand, and the need of better competence on the other, demand, at a first glance, increased sensor diversity and richness, as well as more sophisticated processing of data. Together, these ideas imply that in order for such an agent to scale to more complex environments, the sensors, as well as the sensorimotor maps, have to be rich enough to capture the world's complexity. All relevant aspects of the world must therefore be visible to the agent's perception. Moreover, the sensorimotor maps have to accompany this complexity increase, since they need to encode all motor responses to the high-dimensional configuration space of the world. However, when the above methodology is attempted, the dimensionality increment of sensor data makes the design of sensory maps an intractable task. In addition, such built-in maps may show difficulties in accommodating environment changes (adaptation). In order to preserve performance, the sensorimotor maps ought to incorporate all that environment variety.

Let us consider an alternative approach other than incorporating additional complexity at the perceptual level. This approach consists of adding a second level of processing, that has to process stimuli simultaneously, in parallel, with the perceptual level. The parallel nature of these levels is essential so that the level of competence accomplished by the perceptual level alone is not compromised. This second level, called *cognitive level,* provides a complex, but consequently slower, processing of stimuli. It aims at attaining a cognitive level of competence, e.g. recognition, reasoning, planning, and so on. Consequently, this involves a rich and complex representation of stimuli.

It must be readily acknowledged that the names "cognitive" and "perceptual" fail to capture the full nature of each level described. The model discussed here would remain the same if we replaced them by any other pair of names, e.g., "high" and "low" levels, "complex" and "simple" levels, "first" and "second" levels, etc.

The parallelism of these two levels is a key point of the model. Other possibilities, such as a serial arrangement of the levels, bring conceptual problems. Considering first a perceptual level followed by a cognitive one, the former may throw away information which could be crucial for recognition purposes, because of its low dimensionality. In the alternative serial arrangement, since the cognitive level is slow because of the complex processing implied, the perceptual would lose its key property: fast response to ensure survivability. These two serial topologies can be compared with two classical models from psychology: the William James model in the former case ("we are afraid because we run"), and the appraisal theories in the latter, where the emotional response follows an appraisal process on stimuli[3].

The parallel nature of the model raises questions of synchronization, since the perceptual level is able to provide a response to stimuli faster than the cognitive one. One way to circumvent this issue is to allow the perceptual response to wait for the cognitive outcome, unless the situation demands an urgent response (as assessed by the agent). In other words, the perceptual layer is responsible for taking steps, either by waiting, preparing, or responding immediately, based on its assessment of the urgency of the situation.

[3]Some proposals account for a multiple stage appraisal process, hence accommodating both fast/rough and slow/accurate assessments of situations [165].

4.4.3 Double-representation paradigm

The distinct nature of the processing performed by each level demands that stimuli are represented differently in each one of them. Thus, while the perceptual level manipulates simple and basic representations, the cognitive one uses complex representations of stimuli. The former is oriented to capturing the relevant aspects of the environment, aiming at a quick response to urgent situations, while the latter is directed towards high-level cognitive processing. This constitutes what is here designated as the *double-representation paradigm* of stimuli. The complex representation is called *cognitive image*, while the simple one is called *perceptual image.* The first level provides a fast path from sensors to actuators. We consider that this mapping is performed in two steps: a first step that extracts a representation of the stimulus of reduced dimensionality, *i.e.,* the *perceptual image,* and a second step which maps the resulting perceptual image to action space. Simultaneously, a cognitive image is extracted from the stimulus, and subject to cognitive processing. These two representations are then associated and stored in a memory. By associating these two images, the agent establishes a one-to-one link between a rich representation and a basic one. When shall the agent associate and store these pairs of images? It should depend on a relevance assessment made by the agent, e.g., stimuli that elicit a perceptual response (a threat?), novelty, and so on.

Also, the agent requires a mechanism to permit the assessment of the desirability of stimuli. This is accomplished by the introduction of a third representation schema termed *desirability vector* (DV). This vector characterizes stimuli according to a set of dimensions relevant to the agent, such as dangerous/safe, interesting, demanding urgent action, threatening, to name a few. Following Damásio's SMH inspiration, the associated cognitive and perceptual images are further marked by this representation. This way, the recollection of a cognitive and perceptual images pair is accompanied by the corresponding DV, thus providing the agent with an assessment of the recalled images. It is a multi-dimensional vector, hence extending the traditional scalar utility value (valence). Based on the DV values of the available options, the agent is then able to choose the most desirable course of action.

Deciding according to an assessment of desirability provided by the DV representation raises a question of how the agent bootstraps upon initialization. To do so, the agent needs a built-in mechanism to assess DV values of stimuli. This model hypothesizes that, in addition to the built-in sensorimotor map performed by the perceptual layer, the agent should also be embodied with a built-in map between the perceptual image and the DV.

Since emotions correspond to internal[4] responses to events, from an individual standpoint, one can identify the DV components with this concept of emotions.

Let us now examine closely the consequences of associating a complex (cognitive image) with a simple representation (perceptual image) of the same stimulus. The purpose of storing these pairs (associations) is twofold: on the one hand, the agent may use the cognitive image extracted from the stimulus to search the memory for a pair containing a similar cognitive image — we call this *matching* —, and on the other hand, the perceptual image extracted from the same stimulus may be used to guide the matching mechanism — we call this *indexing.* The perceptual image, in such an association, ascribes a sensorimotor-based representation to a complex representation. For instance, imagine that the agent associates the shape of some object with certain features that triggered a run-away behavior (built-in). The cognitive image containing a rich representation of this shape (e.g., a bitmap) was stored together with the perceptual image representing the threat level that elicited the run-away behavior. Whenever the agent encounters a stimulus with a similar cognitive image (*i.e.,* a similar shape), this cognitive image is matched against the memory, and the previous association is recalled. Depending for instance on the degree of similarity, the agent may exhibit the same run-away behavior, even when the perceptual image of the stimulus does not trigger it by itself. Moreover, when the perceptual image is obtained prior to the cognitive one, the former can be used to guide the search for matching cognitive images. This corresponds to the *indexing* mechanism, and it is essential in order to make the search for a cognitive match computationally feasible, since the number of stored associations can become very large. We argue that this mechanism allows the agent to ascribe relevance to the stored associations, in the sense of constraining the search for cognitive matches to a subset. This subset corresponds to the associations indexed by the perceptual image extracted from the stimulus.

The idea of marking the cognitive image with a perceptual one is based on the Somatic Marker Hypothesis (SMH) developed by António Damásio, reviewed in the section 2.4.2 of this thesis. According to this hypothesis, the brain is able to associate in memory cognitive mental imagery with representations at the level of the body, and later, to enact these body representations after the recollection of that mental imagery [55].

[4]Possibly externalized, if one thinks about the etymological origin of the word. However, emotions also encompass those responses that are not externally visible (but nevertheless measurable, e.g. SCR).

4.5 Mechanisms

The operationalization of the above concepts is performed by three mechanisms: (1) the *marking* mechanism, which establishes and stores in memory associations between the cognitive and the perceptual images, together with DV marks, (2) the *matching* mechanism, which searches the memory for a previously stored association, with the same (or similar) cognitive image, and (3) the *indexing* mechanism which leverages the efficiency of the matching mechanism, by exploiting the different complexity levels of the associated schemata. The marking mechanism creates new associations, while the matching and the indexing ones work together to retrieve matching images from memory.

The marking mechanism creates an association between a cognitive image, a perceptual image, and a DV mark. The goal is to establish an association between instances of these different schemata. The mechanism is triggered according to built-in criteria, for instance, when the agent faces a novel situation. Only associations arising from situations relevant to the agent — with impact on its life — should be stored. Otherwise, the memory would be flooded with repeated and/or irrelevant associations.

The use of these associations, stored in the agent's memory, is done by the matching mechanism. This mechanism is activated each time the agent is exposed to a new stimulus. The goal of this mechanism is to identify the cognitive images, in memory, that best match the one extracted from the stimulus. Together with these matching cognitive images, the agent also retrieves the DV markings, thus providing an assessment of the stored association in terms of its desirability to the agent.

Taking into account that the perceptual representation is simple and fast to compute, a third mechanism is considered — termed indexing — that is used to efficiently perform matching. It involves a two step algorithm: in the first step, a perceptual image is obtained from the stimulus and matched against the perceptual images in memory. For the ones yielding a closer match, the agent, in the second step, matches the cognitive image extracted from the stimulus with those indexed by the closest perceptual images. Considering that the cognitive matching mechanism is an operation more complex than the perceptual one, this mechanism allows for a narrowing of the candidate cognitive images, thus providing an efficient algorithm to find cognitive matches.

The level of competence described so far leaves open several degrees of freedom. They can be exploited to spawn a diverse agent design space, from which varied agent personalities can emerge. Recall that the goal is to implement an efficient system, in the sense of responding adequately and in time to the solicitations of the environment, notwithstanding the complexity of the environment (the problem solved by nature). As for the marking mechanism: when shall the agent create the images associations? Possible answer: depending on the stimulus desirability. How can the agent prevent from storing repeated associations? In other words, a criterion is required to discard redundant information. How shall the agent update previous associations when faced with contradictory assessments? (e.g., contradiction between a stimulus DV and the on associated with a matching cognitive image) As with the matching mechanism: what are the criteria for a cognitive match (e.g., exact match? A metric among images? Until what degree of dissimilarity do two cognitive images match?) Since the perceptual level is faster, when/how shall the agent wait for the (necessarily slower) cognitive matching mechanism? An urgency assessment, using the DV, can be used here. What is the impact of this wait on the agent's performance?

A stimulus is relevant if either the perceptual image elicits a non-neutral DV, or a match in memory is found which is marked with a non-neutral DV. The salience assessment of stimuli is then provided by the perceptual layer, namely the DV. Thus, the relevance of stimuli is established by the DV, in the sense of filtering out irrelevant stimuli from further consideration.

A cognitive match gives the agent a more sophisticated range of possibilities. These possibilities include mechanisms like reasoning, planning, and so on, that can exploit the richness of the cognitive representation. If, on the contrary, no cognitive match is found, the agent is constrained to the limitations of the perceptual representation. Simple behaviors can result from the sole utilization of the perceptual layer. However, these may allow for a basic set of behaviors that ensure the survivability of the agent in an unknown environment.

Considering the matching mechanism in further detail, the following combinations arise from the matching process [198], as illustrated in table 4.1. A strong perceptual impression is here understood as a non-neutral DV value, *i.e.,* eliciting one or more DV components. When a stimulus has both a good cognitive match and a strong perceptual impression, one can say the stimulus is known and relevant. In other words, there is a coherence between the cognitive aspects of the stimulus and its meaning in terms of desirability (assuming of course a consistent environment). If the perceptual impression is strong, but there is no cognitive match, then the stimulus is unknown but relevant. For instance, a threatening stimulus, with undefined contours, so that the agent is unable to recognize its origin, but its perceptual impression impels it to respond to it, e.g. a flight-or-fight behavior. But if, on the contrary, there is a cognitive match, but no perceptual impression, the stimulus is considered irrelevant. Ascribing relevance using perceptual

cognitive match	perceptual match *none*	*strong*
weak	Unknown Irrelevant	Unknown Relevant
perfect	Known Irrelevant	Known Relevant

Table 4.1: The four possibilities arising from the cognitive and perceptual matches.

impressions can contribute to reduce the complexity — both temporal and spatial — involved in cognitive matching processes, since irrelevant images may be discarded (or at least deserve less attention). However, in the absence of other relevant aspects, the agent can invest its resources at finding a cognitive match. Finally, a stimulus both without cognitive match and perceptual impression is found by the agent to be unknown and irrelevant. In a complex environment, most stimuli may fall into this latter category.

The presence of a second level on top the first one may give origin to conflicting situations. These situations occur whenever the outcome of the perceptual level (action or behavior) differs, more or less dramatically, from the perceptual impression retrieved from memory. Different approaches to resolve this conflict can be used, leading to different agent personalities. For instance, an agent taking excessively into account the outcome of the cognitive match (memory), may experience difficulty with discriminating fine distinctions in the environment: since it tends to consider the perceptual impression of the associations in memory; the perceptual impression extracted from the stimulus tends consequently to be dismissed. On the contrary, an agent taking the cognitive match too little into account, may not be exploiting appropriately the "lessons" of the past: the tendency to only take into account the perceptual impression from the stimulus prevents the agent from anticipating future consequences, that could be exploited/avoided by the use of the stored associations.

4.6 Discussion

4.6.1 Meaning

Associating cognitive and perceptual images, together with DV marks, can be understood as basic meaning formation. Such basic meaning is primarily provided by the desirability vector. Take an example of a threatening stimulus, that elicits a desirability vector reflecting fear. The agent can respond to such a stimulus with an behavior concomitant with the threat assessment. In this case, one can say that the stimulus has a meaning to the agent.

Such ascription of "meaning" cannot be done in a lighthearted manner, since this is a controversial topic. For instance, John McCarthy legitimates the ascription of mental states to machines as far as it "expresses the same information about the machine that it expresses about a person." Moreover, he asserts that "it is useful when the ascription helps us understand the structure of the machine, its past and future behavior, or how to repair or improve it" [129]. Taking a different standpoint, John Searle discards the possibility of any meaning being accessible to a machine at all, at least as far as syntactic systems are concerned. To do so, he uses the argument of the Chinese Room [169]: Searle collocates himself inside a room where all connections with the exterior are done by the means of incoming and outgoing strips of paper. These strips contain Chinese symbols. In the room there is a book of rules expressing how to write outgoing streams of symbols, given the input. With the premise of the possibility of building a rule set such that the outgoing stream corresponds to answers to questions, about a given story previously provided, also in Chinese, Searle claims that although it may seem he understands Chinese flawlessly (given the correctness of the answers), he in fact does not understand Chinese at all. One can even imagine the same scenario, but with the story and the questions written in plain English. Assuming that in both versions the answers are sufficiently correct, Searle concludes that, even though from an external point of view the answers are equally good in any language, he understands the English story, while he does not understand the Chinese one at all.

Following Searle's metaphor, one can now imagine a Chinese room scenario where, instead of black characters on a white paper strip, some of those characters are colored. Assuming that these colors have a built-in meaning to the agent, e.g., blue is desirable and red is undesirable, such colored symbols acquire a primary meaning to the agent. The coloring of symbols can be taken as metaphors of DV marks. The symbols marked this way acquire a meaning for the agent: it can respond to them with some action, it can associate a symbol's visual shape with a certain DV, and so on. Moreover, one can consider more sophisticated processes of meaning creation, on top of such primordial ones, thus forming networks of associations bearing powerful cognitive capabilities to the agent [201].

4.6.2 Relevance

The ability to appropriately determine the relevant aspects of a given situation is a *sine qua non* condition for an agent to cope with complex and dynamic environments. The proposed model addresses this issue in the following way. At a primordial stage, the perceptual layer alone filters out irrelevant stimuli, by being only sensitive to stimuli directly related with the basic aspects (e.g., survivability). By associating together cognitive and perceptual representations, the agent can propagate relevance assignments to cognitive representations. Thus, the cognitive layer may consult the perceptual one, for instance, to guide cognitive processes on focusing search processes. Other possibilities include the attention focus being driven by perceptual assessment.

The sub-space of stimuli relevant to the agent is therefore dynamic, starting with a minimum set, and growing to include previously irrelevant stimuli. However, mechanisms are also necessary to allow for assigning relevance by alternative ways. For instance, curiosity (e.g., exploratory behaviors) can permit the agent to consider irrelevant stimuli. By exploring the possibilities of these situations, the agent can discover new potentialities in the environment. When this happens, the agent should establish associations with DV values accordingly, so that the related stimuli become relevant.

4.6.3 Bootstrapping from built-in structures

When the agent is first exposed to an environment, an amount of built-in structures (e.g., maps, associations) is needed in order for the process of interacting and establishing new associations to begin. Otherwise, the agent would stay passive and insensitive to the world. If all goes well, the agent will thereafter keep on constructing new associations and on improving its knowledge of the environment. This mechanism of initializing from a minimal set of maps is here called *bootstrapping.*

Obtaining a successful bootstrapping is not a trivial task. There are two problems that have to be addressed. On the one hand, there is the question of what are the appropriate built-in structures. These should be as minimal as possible, so that the adaptation capability of the agent is not compromised. Too many assumptions about the environment tend to constrain its adaptability in the face of unpredictable and/or changing conditions. And on the other hand, they should be rich enough so that the agent is impelled to explore the environment. A too basic built-in base may constrain the developmental possibilities of the agent.

The engineering of a set of built-in structures may employ evolutionary techniques (e.g., genetic algorithms [89]). Agents with different built-in configurations could be tried out, and evaluated in terms of survivability, adaptability (by exposing them to different environments with the same built-in structures), competence level, and so on. The fittest agent, according to a given criterion based on these aspects, could then be selected for further use.

4.6.4 Efficiency

One of the goals of the presented agent model is to cope with complex environments efficiently. Efficiency is here understood as the quality of utilizing as few computational resources as possible to reach a certain competence level. At a first glance, complex environments demand complex agents. However, Brooks' robots [34] have shown that simple (reactive) mechanisms can lead to robust and appropriate, although simple, behaviors in the presence of complex environments.

The agent model here proposed addresses efficiency from two sides. First, from the double processing of stimuli, two strategies for coping with a situation result: a simple and basic level of competence provided by the perceptual level alone, and a higher competence level resulting from the cognitive layer. In unknown situations (*i.e.,* no cognitive matches), the perceptual level ensures a basic competence level, aiming for survival, as well as bootstrapping mechanisms. In known situations, the cognitive level has the potential to provide mechanisms to exploit the complexity of the world (e.g., planning, reasoning, etc.). And second, the indexing mechanism provides an efficient way to find cognitive matches, saving the agent exhaustive searches over all cognitive images in memory. This efficiency gain results from both the double-representation paradigm, and the association between cognitive and perceptual images.

4.6.5 Homeostasis

Living organisms employ homeostatic mechanisms for maintaining certain physiological variables within appropriate values. Examples of such variables include body temperature, nutrient concentration in the blood, and so on. Whenever one (or more) of these variables become unbalanced, the organism responds to the event, either internally (e.g., muscle movement to increase body temperature), or by taking a course of action (e.g., looking for food). Thus, homeostatic balance can be seen as a source of motivation. The organism's behavior is determined by the need of maintaining homeostatic balance.

This concept can been transported to the autonomous agents field as a mechanism for motivation. To do that, first a set of homeostatic variables have to be established, together with its dynamics. These variables may depend on internal (e.g., battery charge, motors temperature), or on external conditions (e.g., light, wind). Then, the action selection of the agent is guided by the need to keep these variables balanced. The agent may employ a diversified set of strategies to cope with homeostatic unbalance, depending for instance on the world context.

Although homeostasis is an apparently basic mechanism, one can extend it to attain sophisticated levels of competence. For instance, curiosity to explore a part of the environment can be impelled by a homeostatic variable representing the familiarity of the surroundings. Homeostasis can be seen as an unifying concept, similar to the concept of energy in physics or value in economics, capable of embracing a broad range of applications.

4.7 Related work based on this model

The first papers proposing the principles of the agent model here presented were published in 1998 [200, 201, 197, 198]. These papers address the conceptual model, along with a discussion of some of the above-mentioned issues. Some initial implementations were then put together, allowing for some experimentation with the architecture [203, 202, 196]. A brief account of these implementations is presented next.

In the first implementation (called `damasio`), cognitive images — points in $\mathbb{R}^2$ — were associated with perceptual ones (points in $\mathbb{R}^2$ as well). Since this implementation pre-dated the introduction of the desirability vector, the components of the perceptual image have here a direct connotation in terms of desirability: the first component holds a positive desirability, while the second a negative one. After the presentation of a stimulus, the agent searches the memory for the image pair which cognitive image is closest to the one from the stimulus. Distances are evaluated by an Euclidean distance in $\mathbb{R}^2$. The agent's somatic response (same structure and interpretation as the perceptual images) is obtained by a linear weighting (with a parameter λ) of the perceptual image extracted from the stimulus and the one associated with the cognitive match. The associations are established each time the agent receives a stimulus. This simple structure was able to illustrate the marking mechanism, as well as the ability to discriminate among similar stimuli, and to generalize for dissimilar ones (a consequence of the instance-based nature of the learning mechanism employed).

The second implementation (`faces`) comprises stimuli formed by colored bitmaps (a pixel can be either background, black, red or green). The name *faces* arises from the fact that the system was experimented with stylized figures of faces. The cognitive and perceptual images extracted from a stimulus correspond to the bitmap itself, and to the total count of pixels for each one of the non-background colors (thus forming a vector with three components). The colors green and red were given positive and negative connotations in terms of desirability. For instance, a face with red strokes elicits a negative desirability. Depending on the assessed desirability, the agent can either *accept*, *reject*, or be indifferent to the presented stimulus. Like the previous implementation, this one also demonstrates the marking mechanism. For instance, after being presented with a certain face with green pixels (perceptual assessment responds with an accept; then, both images become associated in memory), the agent hereafter responds with an accept after the presentation of a similar face, even in the absence of any green pixel. This implementation also uses an indexing mechanism, utilizing the perceptual representation: the associations in memory with similar pixel count vectors are searched first for a match.

Recalling Damásio's Iowa Gambling Task (IGT), the third implementation aims at similar results, using an implementation of the agent model. Each time the agent is asked to choose among one of the four decks, it assesses each one of them in terms of desirability. An important innovation over the previous two implementations is the adaptability of the perceptual layer. The cognitive layer learning is event based, *i.e.,* it stores episodes of deck choices along with the desirability of the corresponding outcomes, while the perceptual one learns a mapping from the decks to a vector of desirability. The results obtained were similar to the ones published by Bechara *et al* [23, 24]. The frontal patients condition was simulated by preventing the agent from utilizing associations between the cognitive and perceptual images.

These three implementations, along with the presentation of the model, constitute the main contribution of the author's MSc dissertation [196].

4.7.1 The DARE project

The goal of the DARE project[5] included extending the above-mentioned model, aiming at its application to mobile robots. Under this project, several directions of research work were developed.

Several extensions were introduced by Márcia Maçãs *et al.*, and tested in successively more complex environments [120, 122, 121]. These extensions included the introduction of a homeostatic vector (HV), inspired by the idea of a physiological body state (see section 4.6.5). The agent seeks to balance this state to a pre-defined nominal operating point. The desirability vector (DV) is here understood as an assessment of the stimulus *per se,* thus regardless of the agent internal state (namely the HV). From the DV, a homeostatic change vector (ΔHV) is determined, depending on the current homeostatic vector HV of the agent, as well as on the stimulus DV. The homeostatic change vector represents the change on the HV inflicted by the current stimulus. The agent's decision about the action to perform depends on the evaluation of an estimate of ΔHV_i for each action a_i that it can perform. Each of the cognitive and perceptual layers contribute to these estimates, designated the *expected body change* after an action. Márcia Maçãs used the Damásio's Iowa Gambling Task (IGT) as an initial testbed. She experimented with several variations of her architecture [120], achieving in some of them results similar to Damásio's.

Another testbed where she applied the architecture was a 2-D maze world, where the agent moves around in a grid-like fashion. This world contains light sources, good and rotten food. The agent seeks light sources, but has also to satisfy its needs in terms of feeding (energy source). In this implementation, Márcia Maçãs used internal stimuli, to implement the agent's motivation to satisfy its energy needs. The agent was capable of achieving the expected results: it was able to learn to distinguish rotten from good food, using the shape of visual stimuli (and using the cognitive image for discrimination); the agent uses marks on the floor, previously deployed by it, to find its way home, as well as to sites where it previously encountered good food [122]. To do so, a memory structure records sequences of moves, in such a way that the agent is able to walk its way back.

Another extension proposed by Márcia Maçãs *et al.* targeted augmenting the model with an extra layer: a symbolic layer [121]. This extension was implemented in a market environment, where products are exchanged for money among agents. The agents seek survival, as well as the maximization

[5]DARE stands for Emotion-based Robotic Agent Development (the acronym derives from the translation to Portuguese), a research project funded by the Portuguese Foundation for Science and Technology in 1999–2004 (project PRAXIS/P/EEI/12184/98).

of profit from selling goods. This was the first implementation where the model was tested in a multi-agent setup. There is explicit communication among agents, in which the symbolic layer plays a central role. Moreover, it is also responsible for creating sympathy bounds among agents. In this framework, the cognitive and the symbolic layers distinguish themselves in the fact that, while the former is focused on individual behavior, the latter accounts for social issues. Social interaction enables an agent not only to take into account its own experience, but also the experience of others. The agents were tested in the following fashion: first, with the perceptual layer only, then with the cognitive added, and finally with all three layers. The agents showed improved performance levels, as higher levels layers are added, thus showing increased levels of competence.

Taking Márcia Maçãs' early work [122] as a starting point, Rui Sadio *et al.* implemented a variation in a real robot [162, 163], a Scout[6] platform. The task consists of surviving in a world with good and rotten food, represented by colored signs. There is also a ball with which the robot can play with. An internal energy level decreases as the robot moves around. The internal energy level is represented by a homeostatic variable which the robot seeks to balance. The implemented behaviors include *Approach* objects, *Deviate* from obstacles, *Play* with the ball, *Rest.* During the conducted experiments, the robot has shown appropriate behaviors. For instance, it is able to learn to distinguish between good and rotten food signs, and to choose among seeking food or playing with the ball, depending on its energy level.

With the purpose of exploring the learning possibilities of the architecture, Pedro Vale *et al.* experimented with introducing a reinforcement learning module in the architecture [191, 192]. The Q-learning technique [181] was employed, with state identification being provided by the cognitive[7] match mechanism. The action selection is based not only on the Q-values, but also on the anticipated desirability vector (DV) for each of the available actions. The latter also depends on the current body state of the agent, using the homeostatic vector (HV) idea previously presented.

Bruno Damas *et al.* approached the model by using an associative memory to implement somatic marking [51]. He has researched several methods of

[6]The Scout robots were manufactured by the former Nomadic company. They are differential drive robots, including, in the base platform, an on-board PC motherboard, and a sonar ring for obstacle detection. The employed units were upgraded with a CCD camera, and with wireless communication, which was used solely for development and debugging purposes; all processing was performed on-board.

[7]In some of Vale's publications the layer names were renamed: from *cognitive* and *perceptual* layers to *slow* and *quick* levels, and *cognitive* and *perceptual* images to *detailed* and *characteristics* images.

memory management, in order to retain in memory the most relevant information. Damas' work differs in several respects from the DARE architecture. Instead of the double-representation paradigm, a perception $P(t)$ is obtained from the environment. From this perception, a connotation vector $C(t)$ is derived, taking also into account the agent's internal state. The agent seeks to lead the connotation vector to an equilibrium value, in the same spirit of homeostasis. However, the performance of such a greedy approach would degrade in complex and challenging environments, so that this tendency is weighted with an exploratory behavior. Exploration is based on information theoretic measures obtained from the matching mechanism. Some interesting results were obtained in the domain of the RoboCup simulator [106]. In his MSc thesis, Damas has explored several further domains, such as the Blackjack game, and the inverted pendulum problem [52].

Chapter 5

Causal models and anticipation

5.1 Principles

One of the strategies the mind uses to make decisions is by pondering the consequences of the available response options in a given situation. To do so, it anticipates the consequences of each option under consideration, and then decides by evaluating these resulting anticipations.

Transporting this idea to the domain of A.I., *anticipation* is here understood as the capability of an agent to represent, internally, the consequences of its actions upon the environment, when exposed to a given situation. However, to do so, the agent requires a mechanism to formulate the consequences of a given action. In other words, it needs a *causal model* of the environment to derive the effects of its actions. This chapter addresses these two issues — the formulation of causal model, and anticipation — in the context of the conceptual model presented in the previous chapter.

With the purpose of approaching this problem, a simple experiment was devised, consisting in applying the emotion-based architecture to a problem of control supervision (section 5.2). The well-known problem of balancing an inverted pendulum was used as a testbed [199]. The explicit formulation of causal models was considered next (section 5.3), using a stochastic discrete event system as environment. In the experiments, the agent supports two modes of functioning: it is allowed to interact with the environment, collecting relevant information (online mode), and from time to time, it formulates and re-organizes its causal models (offline mode).

According do Damásio, emotions play a crucial role in decision making situations, where a person is faced with various scenarios, and various possible courses of action. The consequences of the available possibilities are then pondered in a means-end analysis fashion. When considering one of possibilities, the emotional circuits of the brain are capable of responding to it. Such responses are usually in the form of an actual (measurable) physiological change[1]. The body changes are then signaled to the brain, leading to either a prompt dismissal of some possibilities, or attraction towards others. This effect is often covert and unconscious, but sometimes it reaches consciousness, and people are then aware of those feelings. The novel aspect pointed out by Damásio's research was that such phenomena happen more often than was commonly thought. Even in apparently rational, non-emotional decisions, such body changes have an effect in the decision-making process.

5.2 A supervision control problem

5.2.1 Motivation

One of the mechanisms discussed by Damásio is what he calls the movie-in-the-brain [56]. Far from the ancient idea of the homunculus inside the brain watching a "movie" of sensory input while deciding what the body should do, this movie-in-the-brain (MITB) conceptualizes a structure registering, over a period of time, the sequence of perceptions, actions, and bodily responses. In Damásio's terminology, these form images representing objects outside the body, as well as representations of the body state. Such a structure allows an individual access to her/his recent history, encompassing not only her/his interaction with the world, but also how her/his body responded to it. Damásio discusses the movie-in-the-brain in the context of the formation of consciousness [56]. He views consciousness as a two step process, where first, a core consciousness is constructed, based on the individual experience. The movie-in-the-brain holds this experience in its various dimensions (perception, body state, and so on), as in an orchestra score containing different staffs for different instruments. On top of core consciousness, a second level is proposed — the extended consciousness — pertaining an autobiographical view of the individual. Such an autobiography is a second order account, made out of fragments of the movie-in-the-brain construed at the core consciousness level.

Daniel Dennett dismisses the Cartesian Theater idea of a place in the brain where all perception is put together. Rather, he recalls that the nervous system is made of several parallel paths from perception to action. Perception is made of various processes distributed over the brain. Thus, per-

[1] To be precise, Damásio also refers to what he calls the "as-if loop", that short-circuits the body loop. In other words, the brain simulates the effects in the brain, thus skipping the actual body changes. However, it is as if the body was involved anyway.

ception is formed by a continuous revision of perceptual data, what Dennett calls the Multiple Drafts Model [65]. Dennett goes a bit further, proposing that his model is also applicable to all varieties of mental activity. The Multiple Drafts Model cannot however be considered to stand in contradiction with Damásio's movie-in-the-brain, since the latter can be seen as a unified conceptualization of the representations held widespread among the brain. Recent neurophysiological evidence seems to corroborate these two views, as in Baars' Global workspace theory [15]. Information is distributed among various specialized brain structures, however, according to Baars, only a part of it is dominant at any given moment.

The design of the supervisor agent was biologically inspired by Damásio's movie-in-the-brain [56]. This computational implementation of the movie-in-the-brain (MITB) is employed here to store the sequence of the agent's perceptions and actions, together with their desirabilities. The agent's decision making is based on the information held in the MITB, hence, on the experience of its interaction with the world. To do so, the agent has to manage two kinds of activities. On the one hand, it must explore the environment by trying certain actions for which it is ignorant about their consequences, and on the other, it matches its current percepts against the MITB, with the purpose of choosing an appropriate action, taking into account its consequences after matching the current percepts with the stored MITB.

5.2.2 Testbed

The testbed employed in this experiment consisted in a simulation of the classic inverted pendulum control problem. A linearized version of this problem can be easily solved by traditional methods[2] using a state-space approach (example 10-2 from [141]). An in depth discussion in depth about related issues, such as stability analysis, can be found in [170].

The system was modeled as a non-linear dynamic system. Taking the variables indicated in figure 5.1, the system was modeled in the following way. The dynamic expressions for the Cartesian coordinates of the pendulum $(x_p(t), y_p(t))$ can be written in the form (the time dependency will be hereafter dropped from the dynamical variables, for the sake of readability)

$$\begin{cases} m_p\ddot{x}_p = -f_p \cos y + K_p\dot{y}\sin y \\ m_p\ddot{y}_p = -f_p \sin y - m_p g + K_p\dot{y}\cos y \end{cases} \tag{5.1}$$

[2] The system is controllable in state space, in terms of both car and pendulum position. To do so, the controller requires full access to the dynamic state of the system. In this case, the car and pendulum positions and velocities.

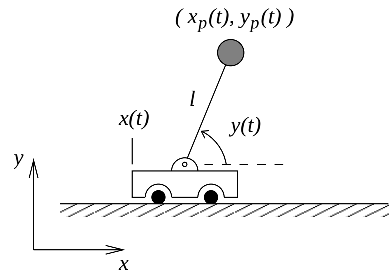


Figure 5.1: Illustration of the inverted pendulum system, with the kinematic state variables shown.

where m_p is the mass of the pendulum, $f_p(t)$ is the force with which the rigid bar (of length l) pulls the pendulum (balanced by the reaction force $f_c(t)$ exerted on the cart of mass m_c). There is a friction in the junction between the pendulum bar and the cart of magnitude $K_p\dot{y}$ onto the pendulum mass. The g is the constant acceleration due to gravity. Time derivatives are denoted with dots over the dynamical variables. Regarding the cart, which can only move along the x axis, one can write

$$m_c\ddot{x}_c = f + f_c \cos y - K_c\dot{x} \tag{5.2}$$

where $f(t)$ is the actuation force exerted on the cart (the system input), and K_c the friction coefficient of the cart wheels.

Two more equations are required to express the kinematic relationship between the cart position $x(t)$ and the pendulum coordinates:

$$x_p = x + l\cos y \tag{5.3}$$

$$y_p = l\sin y \tag{5.4}$$

which after double time derivation results in the expressions:

$$\ddot{x}_p = \ddot{x} - l\cos y\ (\dot{y})^2 - l\sin y\ \ddot{y} \tag{5.5}$$

$$\ddot{y}_p = -l\sin y\ (\dot{y})^2 + l\cos y\ \ddot{y} \tag{5.6}$$

Finally, the forces exerted on the pendulum bar have to balance each other:

$$f_c = f_p \tag{5.7}$$

Solving the above equations with respect to $\ddot{x}(t)$ and $\ddot{y}(t)$, it is possible to simulate the dynamics of the system. This boils down to a four dimensional state vector $\mathbf{x}(t) = (x(t), \dot{x}(t), y(t), \dot{y}(t))$. The state trajectories along time are described by an equation in the form $\dot{\mathbf{x}}(t) = \Phi\left(\mathbf{x}(t), f(t)\right)$.

A fourth order Runge-Kutta method [47] was used in the implementation to numerically solve the obtained non-linear equations. The parameters employed during simulations were the following: $m_p = 0.3$, $m_c = 0.2$, $l = 60$, and $K_c = K_p = 0.1$. The simulation step was set to 0.1.

The system is controlled by a simple proportional controller given by $f(t) = K\left[y_{ref} - y(t)\right]$, where y_{ref} is the desired angular position of the pendulum (e.g., vertical position) and K is the proportional gain of the controller. This is the same to say that the controller aims only at the vertical equilibrium of the pendulum, regardless of the car speed. This is a strong simplification, since when the pendulum is successfully balanced, the car usually keeps moving. Both balancing the pendulum and stopping the car would require a controller with feedback from the pendulum position time derivative and the car velocity (and position, if a certain final position was desired). In terms of the supervisor, the output would involve more than one proportional gain. However, it was decided to keep this example as simple as possible, and only the problem of balancing the pole, regardless of the car speed, was considered.

A control system supervisor (figure 5.2) is built by adding a module — the supervisor — which observes the state of the system, and tunes the controller parameter. In this case, the supervisor and the controller constitute the agent: stimuli are the system state and the actions are the new controller parameter (proportional gain).

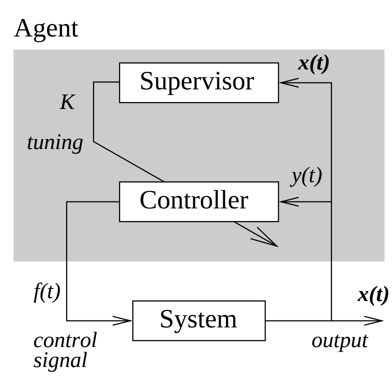


Figure 5.2: The complete system, including the controller and its supervisor, which embodies the emotion-based agent.

To get a realistic flavor of this setup, picture an agent watching the pendulum and measuring the objects' positions and velocities by means of its sensors, and trying to balance the pendulum by exerting a force on the car.

In this metaphor, the controller can be understood as a low-level reactive layer.

5.2.3 Supervisor design

Fundamentals

For each simulation step, the supervisor is stimulated with the state of the system, which corresponds to the vector $\mathbf{x}(t)$. Then, according to the architecture previously described, the cognitive and perceptual images are extracted. In this implementation, the cognitive image equals the state vector, $i_c(t) = \mathbf{x}(t)$ (f_c is in this case the identity), whereas the perceptual one has two components, $i_p(t) = f_p(\mathbf{x}(t)) = (i_p^1, i_p^2)$. These components are the deviation between the pendulum angular position and the vertical position (equilibrium) $i_p^1 = y - y_{ref}$, and the sum of the absolute speeds of the car and pendulum $i_p^2 = |\dot{x}| + |\dot{y}|$.

This choice of cognitive and perceptual images follows the observation that for a complex control problem (a plant, in control theory terminology), using the full state vector (considering it is directly observable or estimable in a reliable way) is intractable, either analytically or in a learning system (recall Richard Bellman's "curse of dimensionality"). Hence, the cognitive image reflects the full complexity of the system, while the perceptual one contains a reduced set of features.

The desirability vector (DV) components represent basic assessments of the desirability of stimuli. In this case, two components were considered: $v_d(t) = (v_d^{val}, v_d^{urg})$, denoting *valence* ($v_d^{val} \in [-1, 1]$, positive if $v_d^{val} > 0$, neutral if $v_d^{val} = 0$, and negative otherwise) and a degree of *urgency* ($v_d^{urg} \in [0, 1]$, 1 means maximum urgency).

The mappings f_d between the i_p (the perceptual image) and the v_d (the DV) are decomposed on two linear piece-wise functions as shown in figure 5.3. The valence is positive when the pendulum is near to equilibrium, and negative when far from it, thus depending on the first perceptual component $i_p^1 = y - y_{ref}$. The urgency increases, until reaching saturation, with the second perceptual component i_p^2, which heuristically assesses the general "speed" of movement (the sum of the absolute values of the first derivatives).

Movie-in-the-brain

For each time step, the agent stores a memory frame in the form

$$m_f(t) = \langle i_c(t), i_p(t), v_d(t), a(t) \rangle \tag{5.8}$$

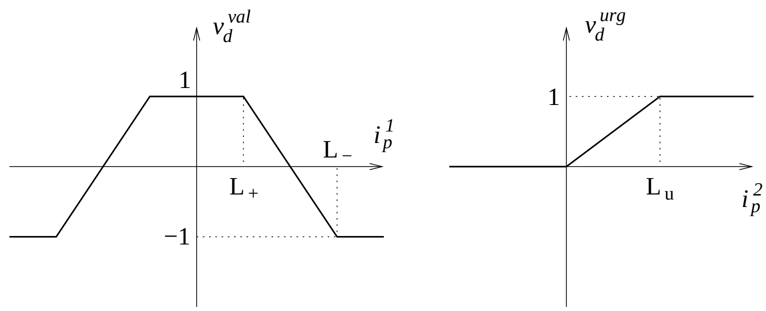


Figure 5.3: Profiles of the functions used to obtain the DV components ($L_+ = 0.1$, $L_- = 0.4$, and $L_u = 250$).

into the MITB. The MITB is a sequence of memory frames

$$\mathcal{M}(t) = [m_f(t_1), m_f(t_2), \ldots] \tag{5.9}$$

where $t_k < t$ for $k = 1, 2, \ldots$ represent the time instants at which the agent receive the stimuli, up to time t.

In the following, a generic time t is always assumed, and, therefore, the argument t will be dropped for clarity's sake.

Topographic map

Action selection is based on a mechanism called *topographic maps* (abbreviated to topmap), inspired by the homonym structure found in the brain: "neurons in the visual areas of the cortex [...] are arranged topographically, in the sense that adjacent neurons have adjacent receptive fields and collectively they constitute a map of the retina. Because neighboring processing units (cell bodies and dendrites) are concerned with similar representations, topographic mapping is an important means whereby the brain manages to save wire and also to share wire" (pages 31–32, [43]). However, the use of topmaps is here limited to the representation of actuation: "A similar hierarchy of multiple topographic maps is found [...] for muscle groups in the motor system" (page 33, [43]), among others. The biologically inspired idea of topographic maps has been used in neural networks [107], among other areas. However, in the context of this work, the idea of topographic maps was adopted in a different, simpler perspective.

The topmap is a function $\mathcal{T}(x) \in \mathbb{R}$ defined in a bounded interval $x \in [x_{\min}, x_{\max}]$. This function represents a map between a variable x and a real value. For instance, one of the topmaps utilized in the implementation maps values of the controller gain K to degrees of ignorance, so that the agent can identify which action it is most "ignorant" about, in terms of its effects.

Such a topmap function is obtained by combining a set of "building-block" functions $\psi(x)$ defined and parametrized as follows:

$$\psi(x; x_0, k, \tau) = k\, e^{-\tau \frac{|x-x_0|}{x_{\max}-x_{\min}}} \tag{5.10}$$

where x_0, k, and $\tau > 0$ are parameters. This function equals k for $x = x_0$ and decays exponentially with a decaying coefficient τ. The denominator $x_{\max} - x_{\min}$ makes the decay coefficient invariant of the variation range of x. The idea of the topmap is to find the argument x that maximizes this function, where ψ functions with positive k contribute as "attractors" and with negative k as "repulsors."

Given a set of contributions $\langle x_0^{(n)}, A^{(n)}, \tau^{(n)} \rangle$, with $n = 1, \ldots, N$, the resulting topmap is constructed in the following way. For each contribution (n), a function $\psi^{(n)}(x)$ is added (point-wise) to $\mathcal{T}^{(n-1)}(x)$.

$$\mathcal{T}^{(n)}(x) = \mathcal{T}^{(n-1)}(x) + \psi^{(n)}(x) \tag{5.11}$$

where the topmap is initialized with zero, *i.e.,* $\mathcal{T}^{(0)}(x) = 0$.

Each contribution is calculated from ψ by adjusting the amplitude parameter $A^{(n)}$ in such a way that $\mathcal{T}^{(n)}(x_0) = A^{(n)}$. This prevents that many ψ functions centered at the same point overload the topmap. Thus, given a contribution $\psi(x; x_0^{(n)}, A^{(n)}, \tau^{(n)})$, the topmap accumulates the following function:

$$\psi^{(n)}(x) = \psi(x; x_0^{(n)}, A^{(n)} - \mathcal{T}^{(n-1)}(x_0^{(n)}), \tau^{(n)}) \tag{5.12}$$

The parameters $x_0^{(n)}$, $A^{(n)}$, and $\tau^{(n)}$ parametrize each contribution individually. However, in this implementation, all $A^{(n)}$ and $\tau^{(n)}$ had the same value. Topmaps were implemented by quantizing the interval $[x_{\min}, x_{\max}]$ in equally spaced (small) discrete steps in x.

Decision-making

The agent's decisions are based on the recent history of its interaction with the environment, as recorded in the MITB structure. Briefly, it searches the MITB for situations similar to the present one. The consequences of the actions performed in those situations are then analyzed. The agent's decision is either to experiment a new action (not yet performed in similar situations), or to perform the action that in the past has led to the most desirable consequences.

A more detailed description of the decision making process follows. At each simulation step, the agent receives a stimulus, formed by the system state vector $\mathbf{x}(t)$. It then obtains the perceptual image i_p, followed by the desirability vector v_d, as defined in section 5.2.3. If the urgency desirability vector component is above a threshold, the perceptual layer takes precedence, and the perceptual action is immediately performed. The perceptual action consists in setting the controller gain to a fixed value if the valence DV component is negative, and zero otherwise.

Unless an urgent situation is encountered and addressed in such a way, the cognitive layer is called to decide whenever the DV values change beyond some other threshold. The DV change is calculated by measuring the Euclidean distance between two successive DV vectors. Additionally, to prevent too long periods without cognitive intervention, there is a maximum period over which a certain gain value is maintained. When this period expires, the cognitive layer is called to decide once again.

1. *Match the present cognitive image i_c against the MITB.* The metric used in this match is a normalized Euclidian distance. Given two vectors $\mathbf{u} = (u_1, ..., u_N)$ and $\mathbf{v} = (v_1, ..., v_N)$ the distance between them is given by

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left(\frac{v_k - u_k}{v_k^{\max} - v_k^{\min}} \right)^2} \tag{5.13}$$

   where $v_k^{\min}$ and $v_k^{\max}$ are the minimum and maximum values of the k-th component, among all cognitive images in the MITB.

   This matching process assigns to each memory frame m_f^M in the MITB, with cognitive image i_c^M, a degree of match given by (the inverse of) $d(i_c^M, i_c)$.

2. *Find local minima of the matching degrees.* The local minima are a subset of memory frames for which the matching degree is less than or equal to the ones of the predecessor and the successor (whenever more than one memory frame satisfy this condition, only the most recent in time is considered). These are the best matches, when looking locally at the MITB. Because of the smoothness of the state dynamics, this method yields a manageable amount of cognitive matches.

3. *Pick a sub-sequence after each cognitive match.* The sub-sequence of memory frames following a cognitive match represents the immediate future after the agent was previously faced with a similar stimulus. This sub-sequence is also dependent on the actions the agent took during that corresponding period of time. This sequence of associations

between cognitive and perceptual images, DVs and actions are the basis for the agent decision making process. There is a fixed parameter that limits the size of each sub-sequence.

4. *Evaluate each sub-sequence.* For the first frame of each sub-sequence, the DV and action are extracted (v_d^1 and a^1), then the first next frame for which the DV changes (in vector distance, with respect to v_d^1) more than a threshold is searched for (v_d^2 and a^2), *i.e.*, $\|v_d^2 - v_d^1\|$ greater than a threshold. If no such change is found, this sub-sequence is ignored. The amount of change is obtained from a weighted sum of the DV components difference, thus forming an evaluation of DV change

$$e_{ch} = \sum_k w_k \left[(v_d^2)_k - (v_d^1)_k \right] \tag{5.14}$$

In this implementation, the weights w_k determine to what extent the agent takes into account the valence or the urgency components of the DV.

5. *Construct action topmaps with respect to "ignorance" and to "evaluation."* Two topmaps are constructed: one called *ignorance*, $\mathcal{T}^{ign}(x)$, representing the degree of ignorance about the effects of a certain action $x = a(t)$, and another called *evaluation*, $\mathcal{T}^{eval}(x)$, representing whether the agent considers the effects of the action $x = a(t)$ desirable or not (positive values mean "desirable," while negative ones mean "undesirable").

   The "ignorance" topmap is obtained by combining a ψ function for the actions of each cognitive match (and local minimum). Each contribution ψ has a negative amplitude, so that each tried action functions as a repulsor. The "evaluation" topmap is obtained in a similar fashion, except that the amplitude of the ψ function depends on the evaluation (5.14).

6. *Action selection.* At this stage, the agent has to decide whether to maximize ignorance (exploration of the environment) or evaluation (exploitation). First, the ignorance topmap is maximized:

$$i_{\max} = \max_{x \in [x_{\min}, x_{\max}]} \mathcal{T}^{ign}(x) \tag{5.15}$$

   Then, if $i_{\max}$ is greater than a pre-specified threshold T_I, the agent chooses to try a different action (exploration). Otherwise, the agent chooses the action which maximizes the evaluation topmap:

$$a_c = \underset{x \in [x_{\min}, x_{\max}]}{\arg\max} \ \mathcal{T}(x) \tag{5.16}$$

where $\mathcal{T}$ is $\mathcal{T}^{ign}$ if $i_{\max} > T_I$, or $\mathcal{T}^{eval}$ otherwise. In this case, if there is at least one contribution with positive amplitude, the agent will choose the one with the greatest amplitude. If all contributions have negative amplitude, a new untried action will be chosen (the farthest apart from the tried ones, taking the absolute difference as distance function).

As explained previously, the agent action a is one of a_p (perceptual) or a_c (cognitive), depending on the urgency of response. In this implementation, a threshold T_U is used, such that if the DV component $v_d^{urg} > T_U$, the perceptual action is chosen ($a = a_p$), otherwise, the cognitive one is chosen ($a = a_c$). The function f_{ap} obtains the perceptual action a_p: it sets the gain to 200 if $v_d^{val} < 0$, and to 0 otherwise. In other words, this function simply turns off the controller gain when the valence component of the DV is not negative, and uses a large value otherwise. This results in a bang-bang kind of control, which purposefully promotes instability[3]. The challenge for the cognitive layer is to come up with gain values that balance the pendulum, thus yielding a better level of competence than the perceptual layer alone.

5.2.4 Experimental results

The inverted pendulum system described above is unstable in open loop. If we take a linearized, frictionless, model of the pendulum, it can be shown that it is unstable for any gain value K. However, in the utilized non-linear model, the friction coefficient contributes to stabilize the system, thus simplifying the control problem significantly. To illustrate this, figure 5.4 shows Monte-Carlo simulations of the system for three different friction coefficients (K_p and K_c). The Monte-Carlo method took the initial condition (initial angular deviation of the pendulum from the vertical position) and the proportional gain K as random variables.

The perceptual layer, *per se,* is unable to hold the pendulum straight, because it uses a too high gain value (200) when the urgency DV component is high enough, and zero otherwise. Thus, it performs a bang-bang kind of control. The goal of these experiments is to assess whether the agent is capable of experimenting with several gain values, and choosing the one yielding good results. Figure 5.5 shows two simulations, one with the perceptual layer

[3] The value of 200 was chosen for this purpose.

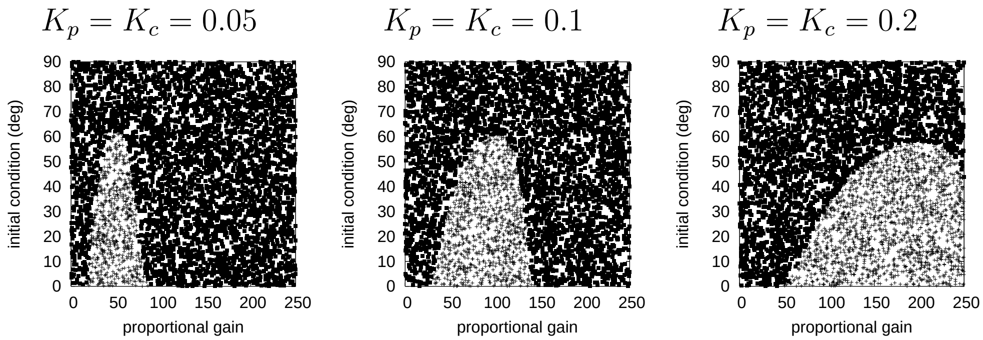


Figure 5.4: Impact of the friction coefficients on the stability of the system, using a proportional controller with fixed gain. The three plots depict Monte-Carlo simulations, for different friction coefficients $K_c = K_p$, with the initial condition and the proportional gain as random parameters. Plus signs ($+$) denote stable outcomes, while the small squares ($\blacksquare$) denote outcomes where the pendulum falls below the horizontal position. Each run took 200 simulation time units.

alone, and the other comprising the full architecture. In this example, the full architecture was able of stabilize the pendulum.

The following experiments exposed four versions of the agent (supervisor) to the system (pendulum+controller), assessing whether it was able to not let the pendulum fall down, for a pre-defined amount of time (termed *lifetime* below). About 1000 runs where performed, for each version, with random initial pendulum deviations within the interval $[0, 20°]$. The versions are described below:

1. *perceptual* — an agent formed by the perceptual layer alone, thus resulting in a bang-bang kind of control;

2. *random K* — a complete agent, where the cognitive layer always decides for a random value of gain K;

3. *full* — the complete agent, but initializing the MITB at the beginning of each run;

4. *full with MITB* — The complete agent, with a pre-loaded MITB previously captured. The pre-loaded MITB was obtained by running the complete agent for 100 times, accumulating the MITB each time. Each one of the 1000 runs began with the same pre-loaded MITB.

The results obtained are summarized in table 5.1. First, as expected, using the perceptual layer alone, the agent lets the pendulum fall at almost

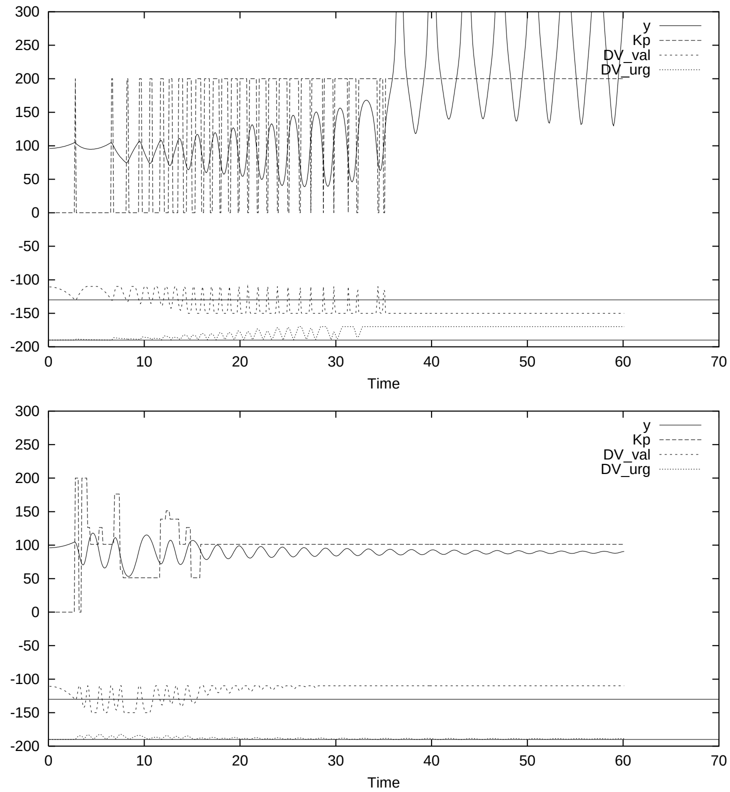


Figure 5.5: Two runs for 6^o initial deviation. The upper plot shows the performance of the perceptual layer alone, while the lower one shows the full architecture. The "y" labeled trace represents the angular position, in degrees (vertical position equals 90^o); the "Kp" represents the controller gain K_p, which is the action $a(t)$ performed by the supervisor; the values for each of the DV components v_d^{val} and v_d^{urg} are also shown as "DV_val" and "DV_urg", relative to two horizontal lines denoting the zero of each component.

every run. When longer lifetimes are used, the percentage raises to 100%, thus suggesting that the 99% value in the 60 time units simulations is only due to small lifetimes. With 60 time units of simulation, the full agent seems to outperform the random K version, although not by a large margin. But when longer lifetimes are tested, the performance of the former is much worse. Moreover, using a pre-loaded MITB seems to worsen results.

These results suggest that, although the agent seems to be able to experiment with several gain values, it has difficulties in utilizing the past experience to choose appropriate gain values. In the short-term it seems to perform better than random choices of the gain. But in the long-term, it performs as badly as the first version (perceptual layer alone).

version	lifetime	
	60	200
perceptual	99.0%	100.0%
random K	39.7%	54.4%
full	26.9%	84.0%
full with MITB	100.0%	99.8%

Table 5.1: Results obtained with the four versions of the agent (see text). The values express percentages of number of runs where the pendulum fell down below the horizontal position.

Running the experiments for a longer period of simulation time (2000 time units), no single run was able to stop the pendulum from falling. Results assessed in terms of time before the fall of the pendulum were collected in the form of the histograms shown in figure 5.6. These results were obtained for 1000 runs, except for the fourth one (full with MITB), which was limited to 100. The pre-loaded MITB employed in this latter version was obtained running 5 times for 200 time units (or less, if the pendulum falls before), accumulating the MITB over successive runs.

Statistics of the obtained lifetimes are shown in table 5.2. It is hard to draw clear conclusions from these results. The random K agent version has a very dispersed histogram, although there is a large number of runs with lifetimes smaller than about 10. The full architecture exhibits a less dispersive histogram, but in terms of the mean and the maximum values, it is worse than the former. Comparing the random K version with the full architecture with pre-loaded MITB, the results suggest an improvement (e.g., the mean is higher), but the histogram still shows a very erratic performance.

version	mean	stddev	min	max
perceptual	35.5	9.4	19.1	77.8
random K	261.2	286.1	2.8	1444.7
full	127.0	80.8	15.3	595.8
full with MITB	197.4	180.4	6.9	673.3

Table 5.2: Statistics of the attained lifetimes of the simulations.

The agent implementation used in the above simulations depends on many parameters. In the previous experiments these parameters were hand-tuned throughout the development of the implementation. Thus, there is a strong possibility that the agent's performance could be improved by a patient fine-tuning of these parameters. Taking a more systematic approach to improve agent performance, a Monte-Carlo simulation was performed, for random parameters values within a range of reasonable values. For each set of parameters values (instance), the performance of the system was measured by simulation.

There are six main parameters that were subject to this process. Two of them are introduced here, which in the past had shown no improvement empirically (and therefore not used in the experiments presented above): a low-pass filtering of the DV values, in the form

$$\tilde{v}_d^i[k] = \lambda_i \tilde{v}_d^i[k-1] + (1-\lambda_i) v_d^i[k], \quad i = 1, 2 \tag{5.17}$$

where $\tilde{v}_d^i[k]$ is the filtered version of the DV value $v_d^i[k]$, at simulation step k. The goal of this filtering is to smooth the DV variations over time. The two new parameters introduced here are λ_1 and λ_2. The remaining parameters are the threshold of DV change (`*dv-change-threshold*`), the maximum period for which an action is maintained before calling the cognitive layer (`*maintain-max-period*`), the threshold on the topmap maximum ignorance level above which the corresponding action is performed, instead of exploiting via an evaluation topmap maximization (`*max-ignorance*`), and the weight of the urgency DV component in the evaluation of the DV (`*t6c-dvw-urg*`).

The system was simulated for each Monte-Carlo instance to assess whether the pendulum fell for a pre-defined amount of simulation time (200 time units). Given the resulting cloud of instances, in parameter space, each one associated to a Boolean value (did the pendulum stay up during the simulation run?), the following procedure was applied: for each true valued point (*i.e.,* the pendulum did not fall), the number of true valued points within the N closest[4] ones was maximized. The point with the greatest number of true points within the closest N neighbors was then chosen. Two iterations of the Monte-Carlo process were performed. In the second iteration the intervals of parameter variation were narrowed, based on the maximization in the first iteration. The second iteration served for a final choice of parameters. The N value used was 100, although the solution was stable for values of N from about 50 up to about 300. Table 5.3 shows the best instance found by the Monte-Carlo process (rightmost column), as well as the intervals in the first and second Monte-Carlo runs.

Using the same methodology as in the above experiments, the histograms and the statistical results of the simulations with this set of parameters can be found in figure 5.7 and table 5.4. The perceptual and random K

[4] Euclidean distance in the space of parameters, after a normalization of zero mean and unit variance, for each one of them.

parameter	original	intervals		final
`*lambda-dv1*`	0.0	[0.0; 1.0]	[0.4; 0.9]	0.74600
`*lambda-dv2*`	0.0	[0.0; 1.0]	[0.0; 0.5]	0.14068
`*dv-change-threshold*`	1.0	[0.0; 2.0]	[0.0; 0.15]	0.037811
`*maintain-max-period*`	5.0	[0.0; 20.0]	[0.0; 15.0]	9.7428
`*max-ignorance*`	-0.2	[-1.0; 0.0]	[-0.5; 0.0]	-0.20046
`*t6c-dvw-urg*`	-0.5	[-2.0; 0.0]	[-1.4; -0.5]	-1.1801

Table 5.3: Parameters obtained using the Monte-Carlo method. Two iterations were performed, with the intervals shown in the two middle columns. The leftmost column shows the original hand-tuned parameter values, while the rightmost one shows the resulting parameter values (up to 5 significant digits).

agent versions were affected because of the DV low-pass filtering process, and in the latter case, also because the K change timing depends on the `*dv-change-threshold*` and `*maintain-max-period*` parameters. The full version has a statistically significant[5] higher mean lifetime value than the random K one, thus showing better performance. Moreover, the former mean lifetime is also higher than the one using hand-tuned parameters (also statistically significant[5]). Concerning the full version with pre-loaded MITB, although the mean lifetime is higher than for the plain full version, the histogram prevents us from drawing conclusions of statistical significance. Still, there is a clear improvement in two aspects: (1) much fewer runs with lifetimes smaller than 200 time units, and (2) an higher amount of runs with lifetimes greater than 800 time units.

version	mean	stddev	min	max
perceptual	20.9	1.8	18.0	30.9
random K	175.7	193.9	10.1	1032.5
full	264.4	108.8	31.8	679.7
full with MITB	368.8	242.9	150.8	1246.1

Table 5.4: Statistics of the lifetimes of the system in the simulations employing the parametrization of the agent from the Monte-Carlo method.

[5] Statistically significance asserted with 99% of confidence level, assuming normal populations, and using the approximation (for large populations) that $\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_x^2}{n}+\frac{S_y^2}{m}}}$ has a unit normal distribution (page 225, [158]).

5.2.5 Lessons learned

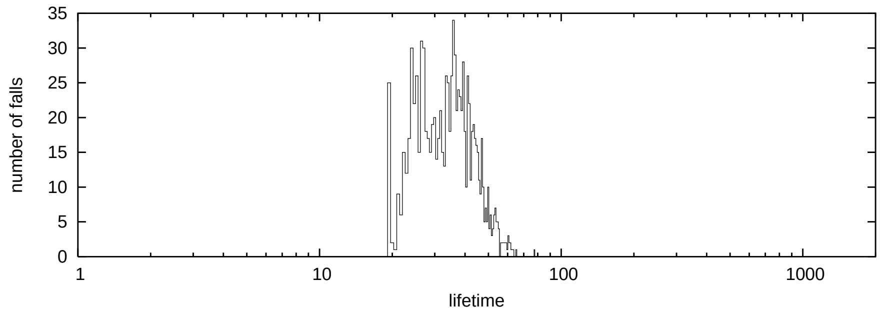
In a paper presenting the research presented in this section [199] a comparison with other models, namely experimental results obtained using a reinforcement learning approach [19], can be found. In the context of this thesis, however, the role of the supervisor experiment is not to propose an alternative approach to the control problem. Rather, its purpose is to identify the issues raised in the design of an agent, in a situation demanding interacting with an unknown world, learning through the interaction with it, and being able to act based on the anticipation of future consequences.

Several factors contribute to make it very hard to analyze the reasons behind the poor results.

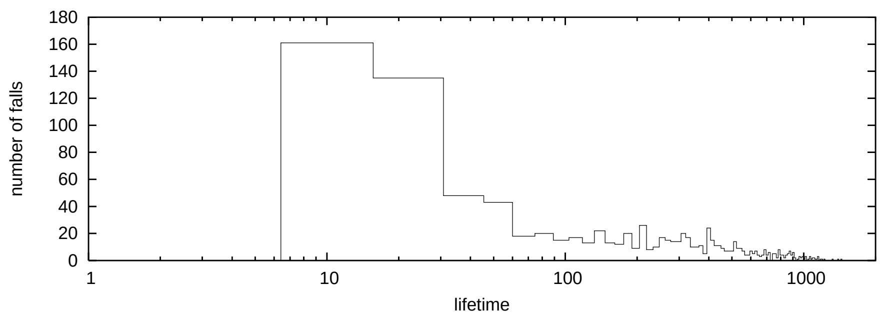
The first problems that arise when applying the agent model to a continuous time control system are the problem of identifying salient events, and the problem of action persistence. Even though the simulation provides a natural discretization, it is too fine grained and therefore not appropriate for decision making. On the one hand, the system state changes slightly for successive time steps. And on the other, the contribution of the actions in the dynamical system during a single time step is negligible, when compared to the net effect of actions performed in the past. Therefore, an action has to persist along a reasonable time interval, so that the consequences can be attributed to that action. In the implementation, events where detected using *ad-hoc* schemes: the agent decided based on significant DV change, and memory matches in the MITB were detected based on local minima of matching degree.

The second problem is how to extract meaningful information from the MITB. Experiments showed that the accumulation of experiences in the MITB worsens dramatically the agent's performance.

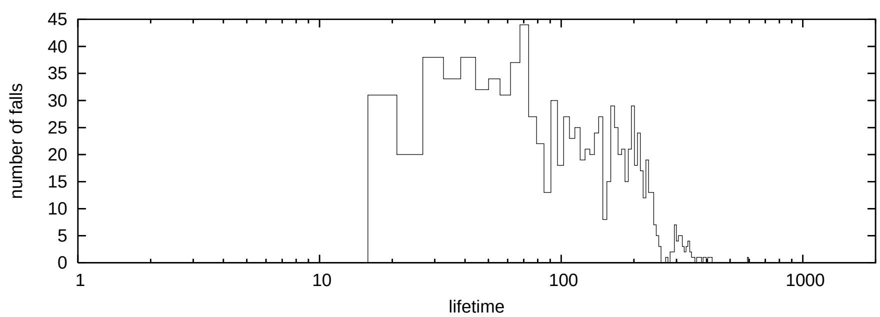
Taking into account that the goal of this research was to experiment with the proposed emotion-based agent model, this implementation has raised issues mostly out of the scope of the model, as the ones above-mentioned. Therefore, two research paths were investigated. One consisted in simplifying the testbed, so that conclusions could be more easily drawn. This corresponds to a back-to-basics approach, described in the following section. The second research path comprised the study of the proposed mechanisms. Chapter 6 describes the conducted research focusing on the indexing mechanism.



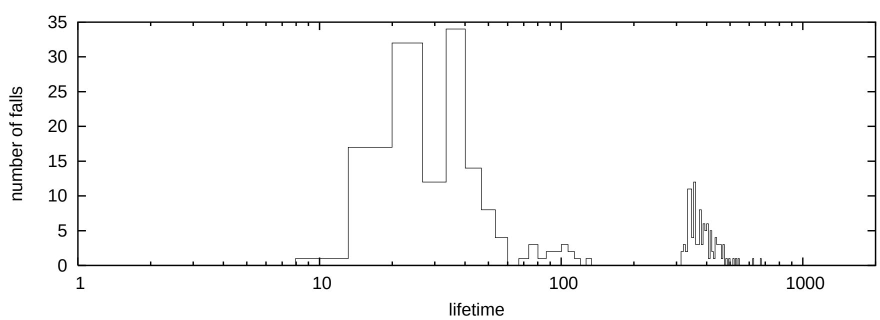
(perceptual)



(random K)

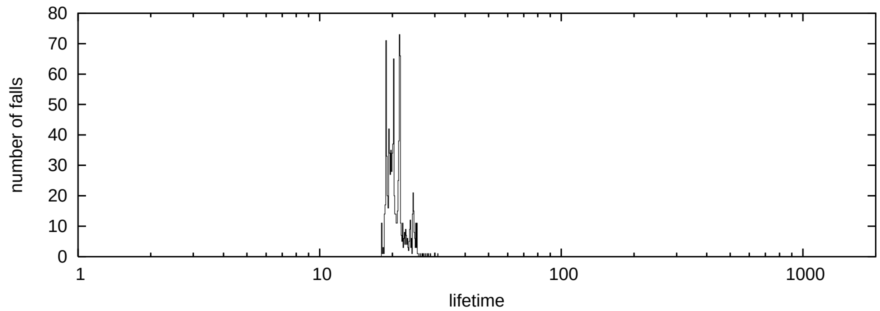


(full)

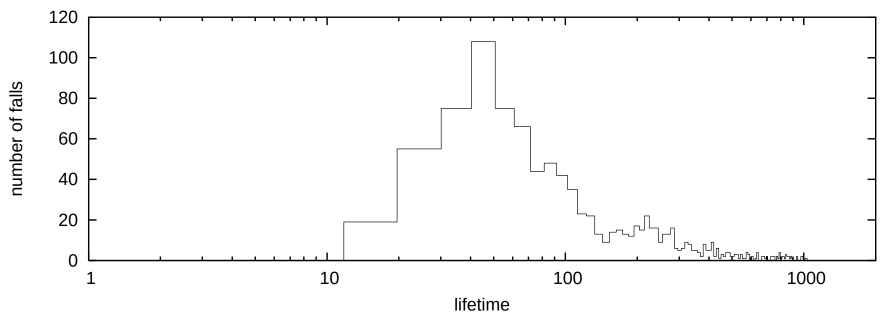


(full with MITB)

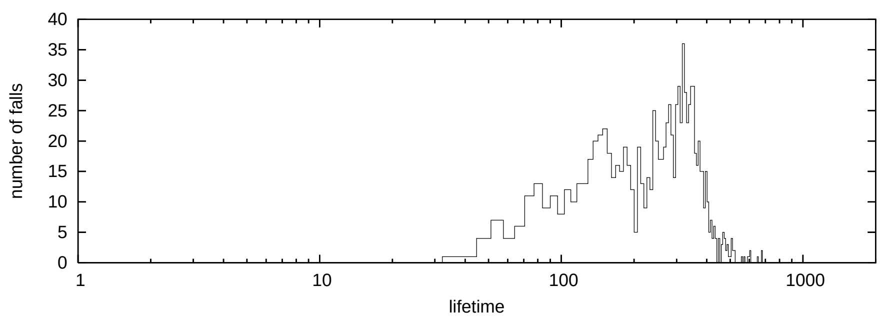
Figure 5.6: Histograms of the number of runs where the pendulum fell down with respect to the time units (log scale) elapsed before falling. The four experiments correspond to the agent versions described in the text.



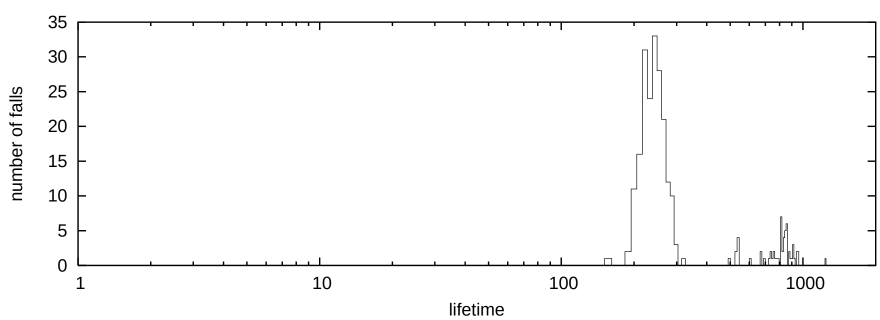
(perceptual)



(random K)



(full)



(full with MITB)

Figure 5.7: Histograms of the number of runs where the pendulum fell down with respect to the time units (log scale) elapsed before falling, employing the parametrization of the agent from the Monte-Carlo method.

5.3 Anticipations in a discrete event world

The continuous time system employed in the previous experiments proved too complicated for making progress. Therefore, a back-to-basics approach was taken here. The continuous time environment was replaced by a discrete one, where cause and effect could be made easier to identify. This allows to factor out the problem of identifying situations in a discrete way, given a continuous stream of stimuli.

A second simplification was to replace stimuli by discrete symbols. Some of these symbols have a built-in connotation (e.g., the symbol '`X`' is undesirable). The others are initially meaningless to the agent. The symbols are generated by a Partially Observable Markov Decision Process (POMDP) model. It is partially observable to prevent the agent to have full access to the environment state. In this way, synthetic environments exhibiting certain regularities can be constructed. These regularities are crafted in a way to provide enough information for the agent to formulate coherent causal hypotheses. For instance, it is possible to deterministically anticipate the undesirable '`X`' symbol, from the occurrence of certain previous symbols. In other words, relevant aspects of the world are made deterministic, while the rest remains random.

Moreover, the agent's actions ought to influence the world in a deterministic way. For instance, the '`AVOID`' action, which is employed by the agent once it anticipates the occurrence of an '`X`', always prevents its occurrence.

In summary, the world presents a sequence of stimuli to the agent, in the form of symbols, one at a time. Some of these symbols are undesirable for the agent. Although the sequences are stochastic, it is possible to model deterministically predictors of relevant stimuli, namely the '`X`' symbol. An `AVOID` action, performed in the step immediately preceding the '`X`', deterministically changes the course of events, *i.e.,* the next symbol is other than '`X`'. Finally, the world provides enough information to allow the construction of a causal model, allowing to anticipate undesirable outcomes.

In this context, what we are looking for is a mechanism capable of anticipating the effects of the agent's actions (and non-actions), given a situation percepted by the agent. Such anticipation can be made explicit, as in this research, or implicit, as in methods that learn based on propagating the utility of states. Examples of such implicit learning can be found in adaptive dynamic programming, reinforcement learning,, and so on. Instead, this research seeks to represent explicitly the relevant aspects of cause and effect, gathered through interaction with the environment.

With the goal of constructing and utilizing those causal models, two approaches were experimented with. The first one uses a decision tree as causal model. The agent interacts with the environment, and collects fragments (cases) of that interaction, using a certain policy. From time to time, it uses the set of collected cases to formulate, or re-formulate, the causal model (decision tree). In a second approach, the causal model is entirely supported by the double representation paradigm of the architecture.

5.3.1 First approach

Implementation

The agent begins its interaction cycle with a minimal *a priori* knowledge about the environment. In this implementation, the negative DV produced by the `X` symbol. Through interaction it formulates and puts a causal model into practice, which relates received stimuli and performed actions, with future consequences.

First of all, the agent has to be able to store, in memory, a sequence of the latest symbols to which the agent has been exposed so far. To do so, we use the same idea of a movie-in-the-brain (MITB), as in previous implementations. Next, the agent needs to collect and store cases, *i.e.,* subsequences of stimuli, which the agent finds relevant to formulate a causal model. Before being exposed for the first time to an `X` symbol, the agent never performs any action. When the first `X` symbol appears, the N previous stimuli (present in the MITB) are stored in a database of cases (N being a parameter: the size of the MITB).

The agent implements two distinct modes of operation: the *online mode,* where the agent interacts with the environment, collects cases when appropriate, and acts according to a previously formulated causal model (if any), and an *offline mode,* where all collected cases are analyzed, in order to formulate a new causal model, or to refine an existing one (if any).

We do not restrict the causal model to a particular technique. In this experiment we used a decision tree structure, using the C4.5 algorithm [153]. However we would like to stress that this implementation can use any other technique. In fact, it could be interesting to consider a portfolio of causal model mechanisms, which may be tested upon the environment, and chosen by the agent, depending on their performance. We believe that complex environments demand a broad variety of modeling techniques.

To build a causal model (as a decision tree) it does not suffice to take sub-sequences leading to negative DV stimuli. The algorithm also requires cases which do not end up in a negative DV. In the case of decision trees, the attribute space is partitioned according to the decision outcomes, requiring the training set to contain cases associated to all possible outcomes.

Therefore, the agent has to be also equipped with the capability of collecting counter-examples, namely those ending up in neutral stimuli. To do so, and avoiding the trivial solution of storing one case per stimulus, the following strategy was used: when a sub-sequence ending in a negative DV is stored in the database of cases, all symbols found in that sub-sequence are associated with that case. This way, symbols that occur before a negative DV stimulus become associated, each one, with one (or more) cases where they took part. When any one of those symbols is found, the stored case is recalled, compared with the past, and held for tracking: all differences the agent finds between the recalled case and the current one are registered. When the tracking of the recalled case ends, and if any differences were registered, a new case is added to the database of cases. The underlying idea is to associate the occurrence of a negative DV with the context provided by previous stimuli. By becoming associated with a negative DV case, all symbols participating in that context impel the agent to compare the present course of events with past cases, and to store any relevant disparities found, as new cases.

This database of cases is used next time an offline mode period occurs. When an offline mode occurs for the first time, a brand new decision tree is built. As mentioned, the decision tree is generated using the C4.5 algorithm [153]. The examples used by this algorithm consist of sub-sequences of stimuli. The attribute values are pairs in the form (n, v), where n is an integer representing the temporal position of either a stimulus $v = s$ or action $v = a$, in the sub-sequence. The outcomes are the DV values — negative or neutral — of the final stimulus in the sub-sequences. The final stimulus and action itself are not included in the attributes, since the decision tree is supposed to anticipate the DV *before* it happens. Finishing an offline mode period, the agent discards the database of cases. For the subsequent offline mode periods, an *ad-hoc* refinement algorithm was used, which will be explained below.

Once an initial causal model is formulated (a decision tree, in this implementation), the agent uses it to be able to anticipate negative DV situations. However, it may happen that the model fails to anticipate correctly a negative DV, or that it anticipates a negative DV that does not occur. In these cases, the model needs to be refined. To accomplish this, these cases where the decision tree fails are added to the database of cases, so that in the next offline period, the agent uses them to refine the causal model.

Refining a decision tree requires supplemental information, taken from the initial training set. For instance, statistical information extracted from the training set has been used to incrementally build a decision tree [190]. For the sake of simplicity, we use a simple *ad-hoc* scheme, that works as follows:

the agent adds to each leaf[6] the subset of cases associated with that outcome. The decision tree refinement is performed using these subsets. The algorithm consists of, for each example, starting at the root, and walking through the tree, until one of the following situations is encountered:

1. An attribute ramification does not account for the corresponding value in a given example: in this case, a new leaf is added at this ramification, associating the new attribute value with the example outcome;

2. An outcome leaf has a different outcome from the one of a given example: a new decision tree is generated, using the C4.5 algorithm, solely using the examples stored in that leaf, together with the given one.

This algorithm is not optimal; however, optimality does not concern us in this experiment. Other algorithms could replace this one. Or even a new decision tree could be generated, taking the union of all the examples stored in the tree, and the new examples, at the price of an additional computational cost. In sum, the idea of the offline mode is to implement the concept of knowledge re-structuring, whenever novel data is available to the agent.

The formulated causal model can be used to prevent exposure to negative DV stimuli. The agent is endowed with a built-in behavior consisting in performing a pre-defined action (symbol `AVOID`) once it anticipates a negative DV for the immediately following stimulus. If that action gives rise to a neutral DV, then this corresponds to an unexpected event (because a negative DV was predicted before). As mentioned before, this originates the storage of a new case, which will be used in the next offline period to refine the causal model. In the end, the causal model contains knowledge, not only about relevant stimuli which precedes a negative DV, but also about the actions capable of preventing negative DV stimuli. This allows the formulation of response options to certain situations. These options associate courses of action with future consequences (in terms of DV values), according to the causal model.

Figure 5.8 shows the architecture of the described agent. During the online mode, stimuli (1) are stored in the movie-in-the-brain (MITB). Under certain circumstances, sequences from the MITB are stored (2) and/or tracked (3). The tracking of cases is triggered by a match of a present situation with a previous case (4). Unexpected events during tracking are stored as new cases (5). During the offline period, a decision tree is constructed or refined (A). The decision tree is used (6) to anticipate (7) what can happen

[6] The ramifications of a decision tree correspond to possible attribute values, and the leafs correspond to possible outcomes.

next. This information is used to formulate courses of action (8), and to choose an action to perform (9).

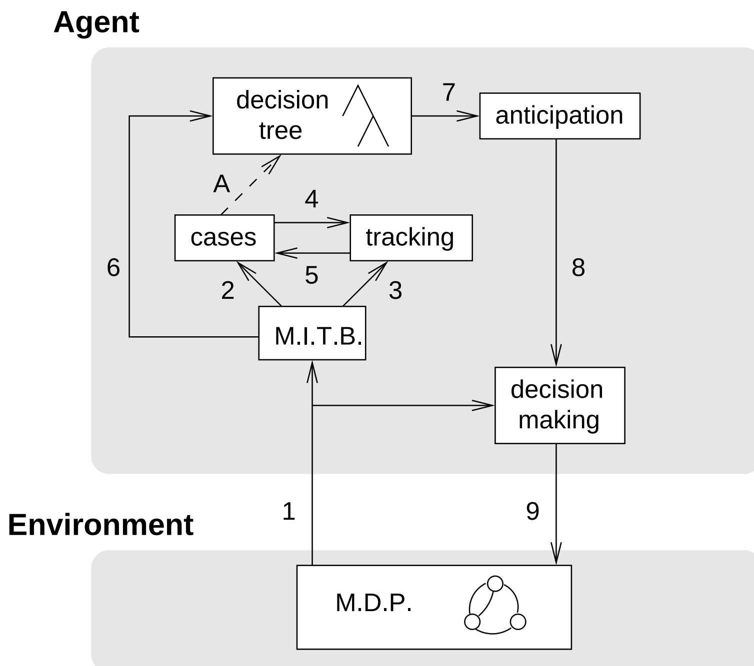


Figure 5.8: Agent architecture. The relationships among the several modules are: (A) decision tree generation or refinement; (1) stimuli the agent is exposed to; (2) storing a case, after an unexpected DV; (3) tracking the differences from a recalled case; (4) recalling a case; (5) storing a tracked case; (6) consulting the decision tree; (7) anticipating future consequences of actions; (8) using anticipations to choose a course of action; (9) action.

Experimental results

To test this simple agent, a synthetic environment was constructed, using a Partially Observable Markov Decision Process (POMDP) as a symbol generator. Each simulation run comprises five periods: three in online mode, interleaved by two offline periods between them. These three online periods have all the same number of stimuli (a parameter of the experiment). The POMDP state is not reset between the online periods, and no symbol is generated during the offline periods. The idea behind this scheme is to provide a first online period where the agent is able to collect cases, a second period to test the generated causal model, where the agent performs an (built-in) `AVOID` action whenever it anticipates a negative DV stimulus, and a third period where it decides according to the consequences collected during the second period. The difference between the second and the third online periods is that the causal model used by the former does not include the effects of performing the `AVOID` action (because none was performed during the first period), while the latter includes the refinements arising from the agent performing `AVOID` actions.

The POMDP used to obtain the results presented below can be found in figure 5.9, here designated *world A*. To correctly anticipate the `X` symbol, in this Markov chain, the agent just has to look for a `B` symbol, followed by any symbol (irrelevant), followed by a `D`, and followed by another irrelevant symbol. Whenever this happens, an `X` symbol follows immediately with probability equal to one, unless an `AVOID` action is performed. There is no other possible way of preceding an `X`. Note that the POMDP was crafted such that the `D` symbol can appear in either states (6) or (7), and that the symbol three time steps before an `X` can either be `A`, in state (2), or `C`, in state (3). An identical situation arises at states (9), (10), and (11). Since the causal model used by the agent does not account for uncertainty, the POMDP used here was crafted such that there exists a decision tree capable of correctly anticipating the `X` symbol (negative DV).

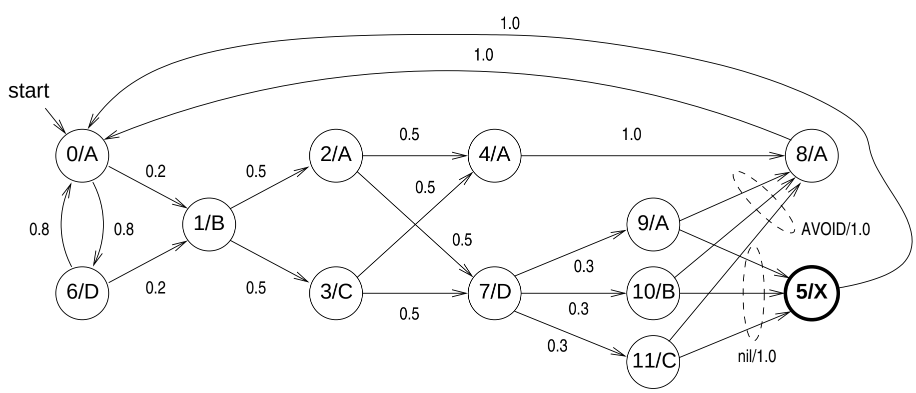
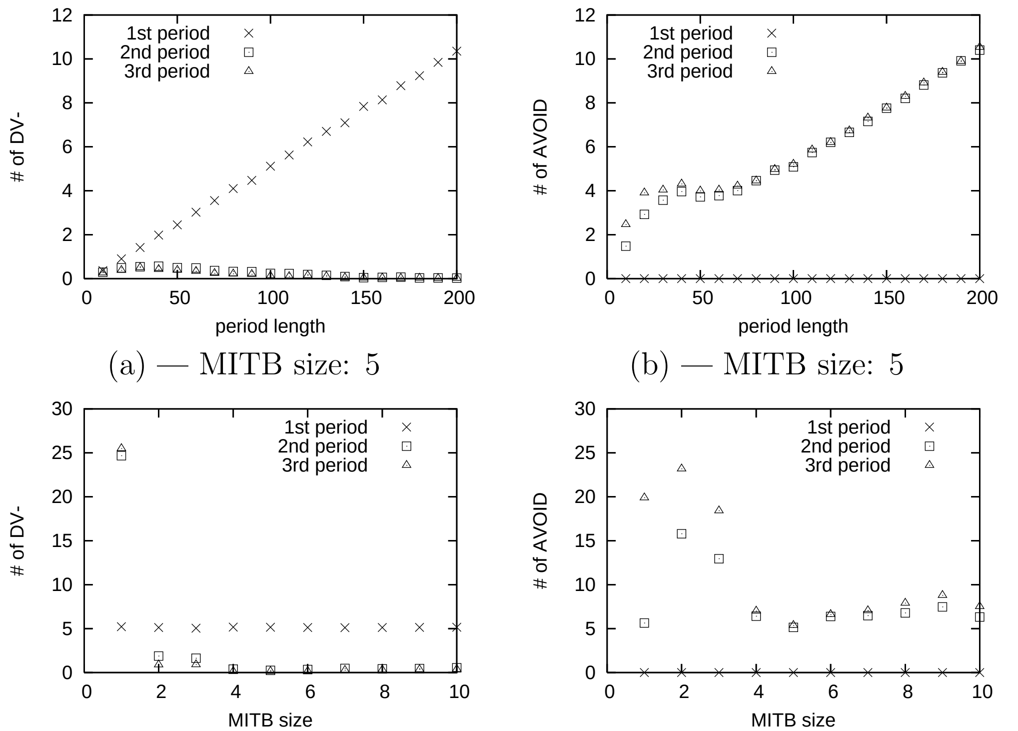


Figure 5.9: The Partially Observable Markov Decision Process (POMDP) used to generate the synthetic environment (world A) used in the experiments. The notation used is `state/symbol` inside the state circles, and either the transition probabilities of the corresponding arrows (when the action performed is irrelevant), or the corresponding action/probability pair (when the probabilities depend on the performed action). The initial state is (0), and the state that outputs an `X` is highlighted in bold. The transitions grouped with the dashed ellipses denote transitions sharing the same action/probability pair.

The criterion used to evaluate the agent performance is the number of negative DV symbols the agent was exposed to, during each experiment period. The results presented in figure 5.10 are plots of this number as a function of the number of stimuli that each online period takes, and of the size[7] of the MITB. In the topmost plots the MITB size was kept equal to 5, while in the ones at the bottom, the period length was set to 100. The results are presented as averages after running each experiment 1000 times.



(a) — MITB size: 5 (b) — MITB size: 5

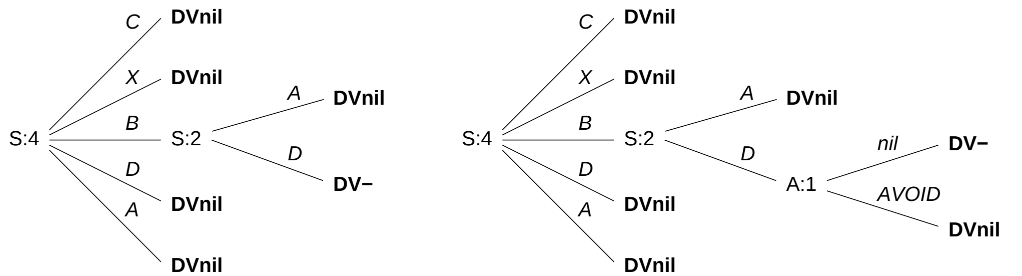
(c) — period length: 100 (d) — period length: 100

Figure 5.10: Sensitivity of the agent performance to the length of the online periods — plots (a) and (b) —, and to the size of the "movie-in-the-brain" — plots (c) and (d). Plots on the left show the agent performance in terms of number of negative DV stimuli perceived, while the ones on the right show the performance in terms of number of performed `AVOID` actions.

Performance is assessed in two ways: counting the number of negative DV stimuli, and the number of `AVOID` actions performed. In the ideal case, during the first period the agent receives a number of negative DV stimuli, while no `AVOID` action is performed. In the second period, no negative DV stimuli is perceived, while a number of `AVOID` actions are performed (approximately

[7]The number of stimuli taken into account to create a case.

equal to the number of DV stimuli in the first period, since the POMDP is the same). The results depicted in figure 5.10 show this trend. However, some deviations exist. Both in the cases of small period lengths and MITB sizes, performance degrades. This can be explained taking into account that, in the former case, there are too few cases to formulate a sufficiently accurate causal model. Concerning the latter case, the number of negative DV stimuli decreases down to around zero as soon as the MITB size is at least 4. This is a consequence of the way the POMDP was crafted: a four-step history memory is the minimum required to predict the occurrence of the `X` symbol. With a sizes of 2 or 3, the `D` symbol can help anticipating the `X`, but it can also occur during state (6). However, with a size of 4, the sub-sequence [`B`, (any symbol), `B`, (any symbol)] allows for a correct anticipation.

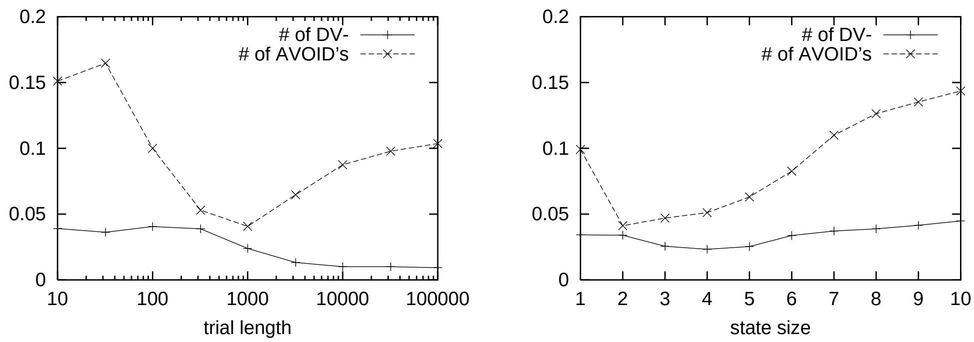


(a) (b)

Figure 5.11: Decision trees generated by the offline processing, after the first (a) and second (b) periods.

Example decision trees generated by the agent, after a single run, are shown in figure 5.11. The first tree depicts the generated tree after the first period. This tree is only capable of anticipating the negative DV stimulus, thus not accounting for any `AVOID` actions (because none was performed during the first period). During the second period, however, since the agent is now capable of anticipating the `X`, it attempts `AVOID` actions. Because these actions successfully change the course of events, deviations from the causal model are collected in the cases database. The offline period that follows refines the tree, adding the effect of the `AVOID` action, as the second tree in the same figure shows.

Figure 5.12 shows some results obtained from using a classic Q-learning algorithm [181] with the same POMDP used above as environment (world A). In order to do so, the Q-values were implemented by a table indexed by the string of the latest N symbols concatenated, where N is a parameter. Moreover, the experiments were conducted for two periods: during the first period (exploration), `AVOID` actions were randomly performed with probability of 1/2, and during the second one (exploitation), the action performed corresponded to the maximization of the Q-values. The reward is -1 for the `X` symbol, -0.1 whenever the agent performs an `AVOID` action[8], and zero otherwise. The plots in figure 5.12 show the sensitivity of the number of negative DV stimuli (same performance criterion as before) to the number of stimuli in each period, and to the N is a parameter mentioned above. The results are presented as averages after running each experiment 100 times.



(a) (b)

Figure 5.12: Sensitivity of the reinforcement learning agent performance to the trial length (a), and to the number of stimuli contained in the state representation (b). The agent performance is measured as the ratio between the amounts indicated and the total number of stimuli in each period.

On the one hand, it is interesting to note how performance degrades as the state dimension (*i.e.,* number of past stimuli included in the state representation) increases: an effect of the "curse of dimensionality." On the other hand, the length of the exploration period required for performance convergence is relatively high: about 3000. Note that one has to be cautious when directly comparing these results to the former ones, since the proposed architecture makes use of built-in knowledge about what to do whenever a negative DV stimulus is anticipated.

The environment used in the above experiments is very simple. So it was decided to construct another one, here designated as *world B*, with two innovations. First, the `X` symbol requires a two-step anticipation, under certain circumstances, to be successfully avoided with an `AVOID` action. In other words, an `AVOID` action performed one step before the anticipated occurrence of `X`, will not prevent it. Rather, only an `AVOID` performed two steps ahead is able to prevent the occurrence of `X`. The second innovation is that a symbol

[8]This penalization is needed to prevent the agent from performing `AVOID` actions all the time.

Y, having a positive DV connotation, was introduced. The goal here is for the agent to anticipate Y, and attempt an exploratory action to facilitate its occurrence. The devised environment allows for an APPROACH action to facilitate the occurrence of Y: the natural occurrence of Y has only 20% probability after state (0), while there is 100% probability of its occurrence after performing an APPROACH in that same state. The diagram of the POMDP implementing this environment is shown in figure 5.13 (world B).

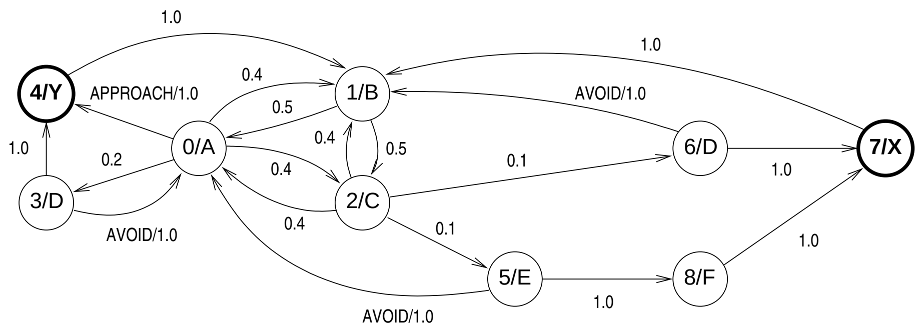
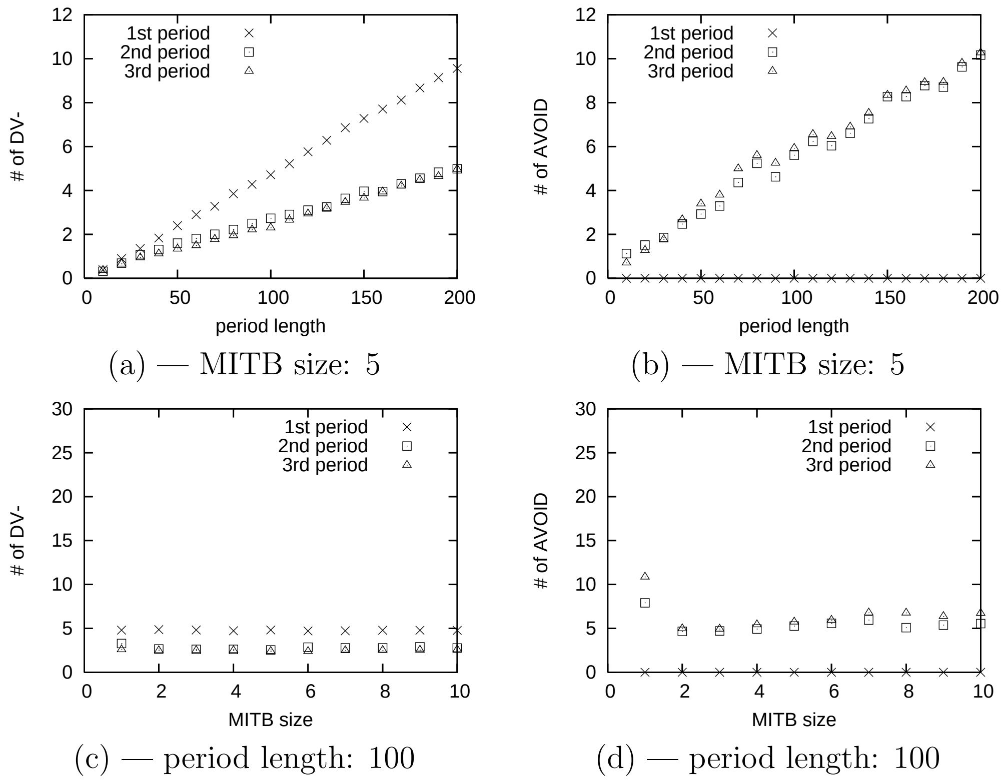


Figure 5.13: The POMDP of the world B environment. The notation used here is the same as in figure 5.9, with the addition that unlabeled arrows denote transitions with probability 1. The nil actions mean no action, as usual.

Using the same parameters as in the previous experimentation, the results show that the agent significantly underperforms. Figure 5.14 shows the results with respect to period length and MITB size.

The second and third periods show an improvement with respect to the first period, in terms of exposure to negative DV stimuli. However, the agent is not capable to anticipating all of its occurrences. The reason is simple: the agent is only capable of anticipating negative stimuli with a single step of antecedence. The POMDP branch leading to the X passing through the state (6) is successfully dealt with, while the one passing through state (8) is unavoidable with one-step anticipation. Since the probability of jumping onto any one of these states is the same (0.1), the number of negative stimuli in the second and thirds periods is roughly half the amount of the first period.

An example decision tree generated using the same environment is shown in figure 5.15. Note that this tree only allows avoiding the negative DV following the state (6). Since this tree uses the F symbol in state (8), it is only able to anticipate the negative DV that follows one step too late. A different strategy is thus required for this kind of anticipation.



(a) — MITB size: 5 (b) — MITB size: 5

(c) — period length: 100 (d) — period length: 100

Figure 5.14: Results of the experimentation using the world B POMDP. No results concerning positive DV are shown because the agent does nothing when it anticipates them.

Lessons learned

The role of the decision tree is to represent relevance. Its structure aims at representing the minimal knowledge necessary to successfully anticipate the DV of situations. The training set utilized to build the tree contains sequences of stimuli (and performed actions). These sequences include the perceived stimuli until (and including) one step prior to the relevant event. As a consequence, it is possible to deterministically anticipate negative DV situations, using the stimuli up to one step before it happening. In particular, consider a tree may containing an attribute value for a stimulus one step before the event of a negative DV, in one of its ramifications. If this is the case, negative DV anticipation through that ramification will require the perception of that stimulus, one step before the relevant event. Consequently, two-step anticipations through that ramification are not possible. This limitation results from the way the decision tree is constructed.

Two other difficulties arising from the usage of decision trees were also

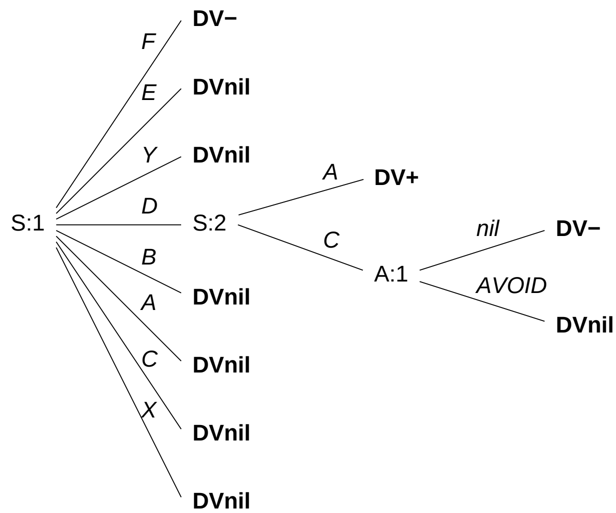


Figure 5.15: A decision tree generated by the offline processing, after the second period, using the world B POMDP.

found: first, a poor handling of uncertainty, namely if the occurrence of the `X` cannot be deterministically anticipated; and second, if for the same sequence there is more than one way of anticipating the `X`, the decision tree will account for only one of them[9].

In fact, it is possible that some, if not all, of these limitations could be circumvented by the use of sophisticated decision tree techniques. However, a different methodology was chosen. Instead of using a decision tree as representation model, the research was directed towards the investigation of techniques that used the double-representation paradigm. The results of this investigation are presented in the next section.

[9]In the POMDP of the world B one such case exists: once in the state (5), the symbol sequence that necessarily follows is [`E`, `F`, `X`]; since either `E` or `F` anticipate deterministically the `X`, a decision tree generated with examples from this POMDP will employ only one of these symbols.

5.3.2 Second approach

The goal of this approach is to overcome some of the limitations found in the previous one. The decision tree utilized proved hard to scale with environment complexity. The design choice consisted in replacing the decision tree by associations, as proposed in the double-representation paradigm of the model proposed in this thesis.

Implementation

The application of the double-representation paradigm to a problem requires the definition of the cognitive and perceptual schemata. The considered cognitive images represent sequences of stimuli (symbols), while the perceptual images represent sets of features extracted from the MITB. Each feature comprises a symbol and a temporal offset with respect to the end of the sequence (present time). For instance, a feature $\langle \texttt{C}, 2\rangle$ represents a `C` symbol percepted two time steps ago, with respect to the present time. Perceptual images can be more or less specific, depending on the number of features they contain. A perceptual match is only considered whenever all of its features match the MITB (conjunction).

A cognitive image represents a sequence of symbols. With respect to a reference time step, this sequence spans not only an amount of symbols preceding that step, but also an amount of symbols that follow it into the future. The symbols representing the past allow the agent to match this image with the agent's current situation, while the ones representing the future enable the agent to formulate predictions about the future occurrence of symbols. This latter feature can be used for planning purposes, although this possibility was not explored in the implementation. Once a cognitive image is stored, it is incomplete, since it cannot account for future stimuli. In the following time steps, that cognitive image is completed with the missing symbols, as the agent is sequentially exposed to new stimuli. The size of these sequences is fixed, and is specified by a fixed parameter.

The cognitive and perceptual images are associated in the following manner (see figure 5.16). Each perceptual image contains pointers to structures representing the consequences of performing certain actions (including no action). Each one of these structures is here termed *option,* since it represents an optional course of action available to the agent. Since the implementation uses only three actions (`AVOID`, `APPROACH`, and `REACT`), a perceptual image points to at most four options: one for each action plus one for no action (denoted `nil`).

Each option references a set of associated *outcome* structures, containing

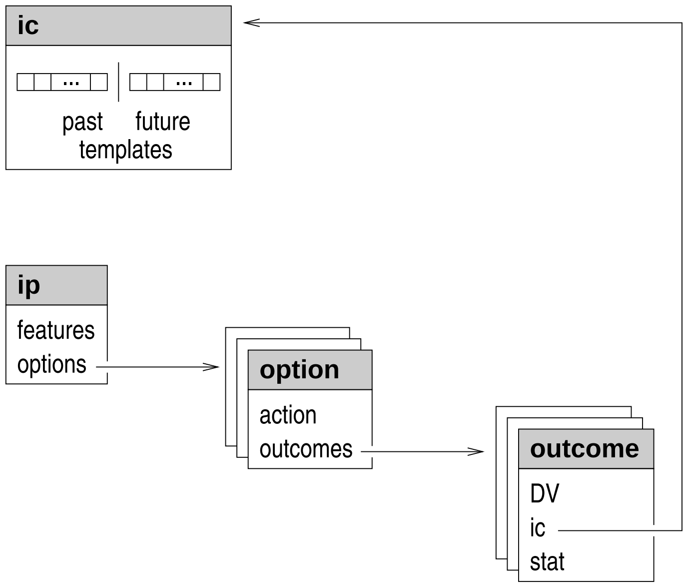


Figure 5.16: Schematic representation of the data structures in the implementation (ic: cognitive image, ip: perceptual image, stat: statistical counters; see text for the description of the remaining elements).

a specific desirability vector, and a cognitive image representing the past and future stimuli percepted in that situation. Each option points to at least one outcome. Various outcomes associated to the same option are possible, thus representing different possible outcomes of the same action in a given situation. For a given option, the corresponding outcomes are grouped by DV. For simplicity, the implementation employs only three discrete DV values — positive, negative, and neutral — and thus, for each option, up to three outcome structures can be found.

As the agent interacts with the environment, it constructs a web of these structures and associations. The perceptual images tend to over-generalize what features are relevant for a certain future consequence, while the cognitive ones tend to over-specialize. The degrees of over- and under-generalizations vary dynamically, as the agent compares its hypothesis with its world experience. Statistics are collected in several locations of this web. During development, the issues that required most effort were the management policies of the creation, refinement, and deletion of these images and their associations.

In an ideal case, a stable situation is reached as soon as perceptual images get as specific as needed, in the sense of providing correct anticipations. Then, the cognitive gets as general as possible, given the agent experience.

Taking the world A, for which the POMDP is depicted in figure 5.9, a perceptual image containing the feature set

$$\{\langle \mathtt{D}, 2\rangle, \langle \mathtt{B}, 4\rangle\} \tag{5.18}$$

has 100% success ratio in anticipating the negative DV originated by the symbol X. Cognitive images indexed by perceptual ones contain templates, for instance in this case a template

$$[\mathtt{B}, \mathtt{A}, \mathtt{D}, \mathtt{A}/\mathtt{AVOID}, \mathtt{A}] \tag{5.19}$$

can be indexed by the above perceptual image. This template may be generalized to

$$[\mathtt{B}, *, \mathtt{D}, */\mathtt{AVOID}, \mathtt{A}] \tag{5.20}$$

where $*$ denotes "any symbol," and a slash followed by a symbol denotes the action performed at that time step.

Taking now the world B (figure 5.13), a perceptual image with a single feature $\langle \mathtt{F}, 1\rangle$ anticipates the following X correctly, but an associated cognitive image with the template

$$[*, \mathtt{C}, \mathtt{E}, \mathtt{F}, \mathtt{X}/\mathtt{REACT}] \tag{5.21}$$

does not get any more generic than this, as long as no actions are performed prior to the REACT one.

In a given situation, the agent first selects the subset of perceptual images whose features match the agent's present stimulus together with its recent interaction history (as recorded in the MITB). For each of these matching perceptual images, all option and outcome structures are retrieved from memory. These structures can then be used to anticipate the DV of the next stimulus, depending on the agent choice of action. Moreover, each outcome structure indexes a cognitive image referencing a sequence template.

Prior to the agent's interaction, its memory is cleared. The only built-in association is between the X symbol and a negative DV, and between the Y and a positive DV. Until encountering the first stimulus eliciting a non-neutral DV (*i.e.,* one of these two symbols), the agent performs no action. Once a non-neutral DV is elicited, the agent forms an initial hypothesis of its cause. This involves the creation of a single cognitive image, containing the past sequence leading to the present stimulus. Moreover, in the next few steps, this image will record the sub-sequence that will follow. The number of steps considered, both before and after the present step, is a parameter of the agent (N_{mitb})[10]. Next, a set of perceptual images is created. For each feature found along a N_{mitb} number of preceding steps, a perceptual image is created. Recall that a feature is here considered to be a pair $\langle s, \delta \rangle$ where s is a symbol and δ an offset, as discussed above. Thus, a perceptual image (with the corresponding feature) is created for each symbol found in the window formed by the N_{mitb} previous steps. Moreover, each perceptual image points to a single option structure (with null action), which by itself points to a single outcome structure, containing the corresponding DV, as well as the cognitive image mentioned above. Note that the perceptual images do not share neither the option nor the outcome structures. This way, these structures register the agent's experience, separately, from each perceptual image point of view.

With exception of the trivial case of all hypotheses correctly anticipating the next DV, improvements of the memory structures are needed. This happens whenever an outcome structure (indexed by a perceptual image and option structure that match the situation) incorrectly anticipates the DV. These improvements boil down to a pair of processes. The first process takes care of updating the statistics of success of each outcome structure, while the second one refines perceptual images by instantiating new ones with more specific sets of features.

Statistics are collected in each outcome structure (slot `stat` in figure 5.16), counting the number n_{uses} of times the referrer perceptual image and option structure do match, as well as the number of times the DV does match (n_{succ}). The ratio of these two values expresses a success ratio of the anticipation proposed by that outcome structure. Moreover, the agent only considers an anticipation as reliable as long as n_{uses} is above a threshold (*i.e.,* maturity). It can be said that this constitutes a mechanism for handling the exploration vs. exploitation issue, since an outcome is only taken into account (for decision making purposes) after its anticipation hypothesis has been tested against experience at least a certain number of times. Note that more general perceptual images tend to be more used than the more specific ones. Consequently, the outcomes associated with them mature faster, thus providing the agent with a rough anticipative power, until the more specific images reach maturity.

The process of refining perceptual images functions by instantiating copies of perceptual images already present in the memory, and modifying them to make them more specific. This happens whenever the agent is unable to

[10]It was only for a matter of simplicity that the same parameter was used for the amount of symbols to retain before and after the present step. Nothing here constrains the equality of these two amounts.

correctly anticipate the current stimulus DV, *i.e.,* whenever there is a DV mismatch between an outcome structure and the current situation. Given a perceptual image i_p, pointing to an option a, referencing an outcome u, the agent compares the sequence template of the indexed cognitive image i_c against the MITB to derive discriminating features. The discriminating features are the ones that are present in one sequence and not in the other. This forks into two sets of features: S_1 with the features found in the MITB but not in i_c, and S_2 with the converse. For each one of the S_1 features, a new perceptual image is formed by adding that feature to i_p, and creating associated option and outcome structures reflecting the action performed and DV experienced. Then, *mutatis mutandis* for the S_2 feature set, where the option and outcomes structures reflect the situation represented by the cognitive image from the agent memory.

Take for instance the world A (figure 5.9), and a perceptual image with a single feature $\{\langle B, 4\rangle\}$, associated with the template

$$[\mathtt{B}, \mathtt{A}, \mathtt{D}, \mathtt{A}, \mathtt{X}] \tag{5.22}$$

Consider now that the agent's interaction results in the following sequence of stimuli:

$$\ldots, \mathtt{B}, \mathtt{A}, \mathtt{A}, \mathtt{A}, \mathtt{A} \tag{5.23}$$

When perceiving the last stimulus, the above perceptual image will match. Since the DV associated with the last stimulus is neutral, discriminations will be created using these two sequences. Apart from the last stimuli, which do not count for anticipation purposes, these sequences differ at the third step, and thus $S_1 = \{\langle \mathtt{A}, 2\rangle\}$ and $S_2 = \{\langle \mathtt{D}, 2\rangle\}$ are the discrimination sets, as defined above. Two new perceptual images (with associated structures) are created, one with the feature set $\{\langle B, 4\rangle, \langle \mathtt{A}, 2\rangle\}$ associated with a neutral DV, and the other with $\{\langle B, 4\rangle, \langle \mathtt{D}, 2\rangle\}$ associated with a negative DV.

As a result of these mechanisms, the number of perceptual images (and associated structures) grows with the agent interactions, until stability is reached. However, the agent only adds a new perceptual image to its memory unless there is already one with exactly the same set of features. Moreover, recall that the refinement process is only called if there are no perceptual images that correctly anticipate the DV.

The agent functioning works according to the following algorithm. Each time the agent is exposed to a stimulus, it first obtains a DV assessment of it. If the DV is negative, it performs a reactive response (a `REACT` action). Otherwise, it runs through two processes. The first one (called *anticipate*) performs a perceptual match, thus gathering all the perceptual images that match the current situation. Based on the estimated reliability of the anticipated outcomes, measured by the collected statistics, it decides on what action to perform. The second process (*update*) performs another perceptual match, but now checking whether the anticipations valid in the previous step, do or do not match the current situation. Depending on the result, more specific perceptual images may be created, as described above. These two processes are further detailed in the following paragraphs.

Given a situation, the decision-making is based on the construction of a set of anticipative scenarios, one for each possible action. Thus, for each action a, a set of $\langle i_p, u\rangle$ pairs is selected, among the perceptual images that match the current situation. For each set, a preference measure $Q(a)$ is calculated using the expression

$$Q(a) = \frac{\sum_k V(u_k) Rank(u_k)}{\sum_k Rank(u_k)} - C(a) \tag{5.24}$$

where $V(u)$ depends on the DV of the outcome u (plus one if positive, minus one if negative, and zero otherwise); $C(a)$ represents the cost of performing the action a (set to 0.2 for all actions other than `nil`), and $Rank(u)$ is a measure of the reliability of the outcome u. Since perceptual images of different degrees of specificity (*i.e.,* number of features) coexist in memory, for an outcome u, its $Rank(u)$ depends non-linearly on the statistics associated with the anticipated outcome. After experimentation, the following form of $Rank(u)$ showed good results:

$$Rank(u) = \begin{cases} 100 & \text{if } n_{\text{uses}} = n_{\text{succ}}, \\ 10R(u) & \text{if } 0.9 \leq R(u) < 1, \\ R(u) & \text{otherwise.} \end{cases} \tag{5.25}$$

where $R(u)$ stands for the ratio $n_{\text{uses}}/n_{\text{succ}}$. The above coefficients allow for an outcome with 100% success ratio to overcome all others, as well as giving preference to the ones with at least 90%.

When faced with a new and unknown environment, the agent needs an exploratory mechanism. This mechanism works as follows. From the above algorithm, the agent obtains the best value of Q by maximization with respect to the actions, thus obtaining the best Q value Q^* and the corresponding best action a^*:

$$Q^* = \max_a Q(a) \tag{5.26}$$

$$a^* = \arg\max_a Q(a) \tag{5.27}$$

If Q^* is below a (negative) threshold, the agent checks if it has performed the AVOID action before (*i.e.,* there is at least one anticipation for that action). If yes, the agent considers the current situation as too undesirable, thus changing the current DV value to a negative one (Note that this means overriding the DV assessed from the stimulus). If not, the agent "flips a coin" and performs the AVOID action with 1/2 of probability.

If Q^* is above a (positive) threshold, it checks whether an anticipation with APPROACH exists. If there is such anticipation, the agent changes the current DV to a positive one. Otherwise, the agent performs an APPROACH action.

Finally, if Q^* lays between the two thresholds, the agent performs the action a^*, which is the one that maximizes $Q(a)$.

This mechanism requires two commentaries. The first one is that it constitutes a mechanism of propagating DV values back in time. Without it, this agent would not be capable of anticipating with more than one step of antecedence. But once DV values can be propagated, the agent may end up propagating anticipations several steps back, and since the DV values are discrete, and there is no time discount mechanism (as in Q-learning, for instance), there is a plausible risk of propagating DV values indefinitely. Extensive experimentation was carried out to overcome this problem, adding continuous values of DV, but no satisfactory results were obtained.

The second commentary is that DV values can have origin in two sources, one is a direct assessment of the stimuli, and the other is internal. If the agent realizes that the best it can do, in a given situation, is too undesirable (or desirable above a threshold), it decides to declare that situation as having a negative DV by itself (or to be sufficiently interesting to be considered with a positive DV). In other words, this allows for the elicitation of DV values by internal causes.

In each time step, the update process runs through all matching perceptual images, descending through the outcomes associated with the performed action (option structure). First, the statistics of the outcomes are updated, according to the current DV value. Then, the agent evaluates whether more specific perceptual images have to be generated. If there are no matching perceptual images, a set of initial hypotheses is created. This means creating single feature perceptual images, one for each feature found in the last N_{mitb} MITB steps. When there is at least one reliable[11] outcome, the update process ends here. Otherwise, the agent generates a set of discriminations following the algorithm outlined above.

In summary, one can say that the *anticipate* process gathers perceptual

[11] An outcome is *reliable* iff it is mature and $R(u) \geq 0.9$.

matches aiming at anticipating the next step, while the *update* process looks one step back, comparing the anticipation with the current situation. This is why, for the same time instant, the perceptual match results in different sets of perceptual images for each one of these two processes.

Experimental results

The following experiments utilized the same POMDP worlds as in section 5.3.1 (worlds A and B, figures 5.9 and 5.13). Two batches of evaluation were performed: (1) the behavior of the agent along time, and (2) the final performance after a specified number of steps. In the first batch 1000 simulation trials were executed for each one of the worlds. The amounts of non-neutral stimuli (`X` and `Y`) and non-null actions were counted, in time intervals of 50 steps, and averaged over the 1000 trials. The results for the two worlds can be found in figures 5.17 and 5.18. Regarding the second batch of simulations, the agent was run for a pre-determined amount of steps, so that a steady state could be reached, followed by accounting the total number of events (non-neutral DV stimuli and actions) for another pre-determined amount of steps. As in the previous case, the results are presented as statistics over 1000 trials.

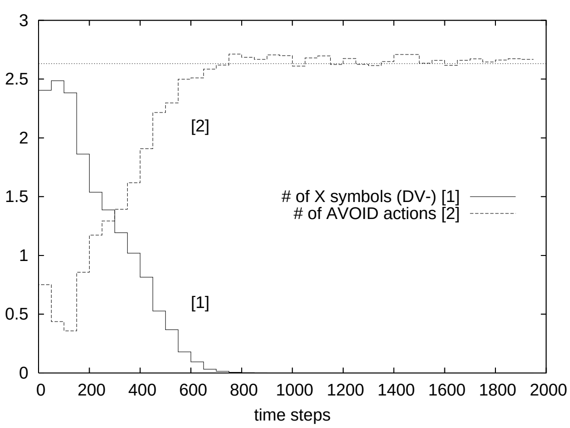


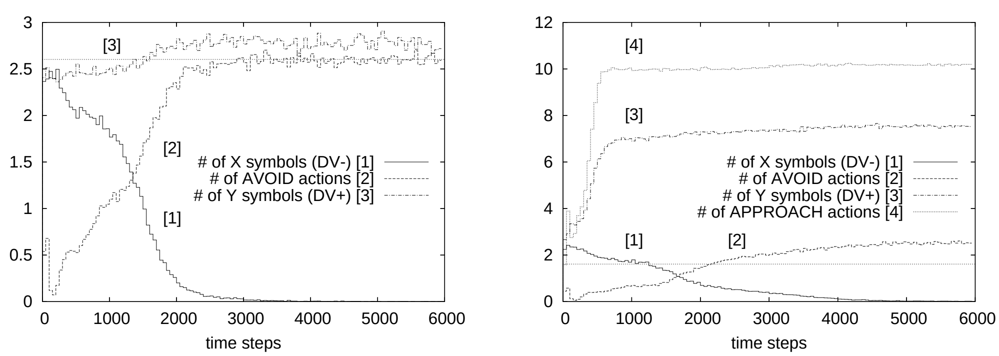
Figure 5.17: Performance of the agent when interacting with the world A (figure 5.9), measured in terms of number of `X` symbols perceived (negative DV) and number of `AVOID` actions performed. These quantities where accounted for successive 50 step periods. The theoretical minimum number of `AVOID` actions is indicated with a dotted line.

Analyzing the results for the world A (figure 5.9), after about 1000 steps, the occurrence of negative DV was eradicated, and the rate of `AVOID` actions approached the theoretical value quite closely. This theoretical value was

	mean	stddev	min	max
`X` stimuli	0.00	0.00	0	0
`AVOID` actions	53.23	17.82	36	590

Table 5.5: Statistics of 1000 trials, using the world A. The number of the above events were collected during 1000 steps, after running the agent 2000 steps. In this case, the theoretical expected number of `AVOID` actions is 52.63.

calculated from the stationary state probabilities of the Markov chain[12], obtained by performing the `AVOID` at states 9, 10, and 11 of the POMDP. This amounted for a probability of 0.05263. This probability times the number of steps in each period (50) gives the expected number of `AVOID` actions per period (2.631 actions). Table 5.5 shows the statistics after reaching a steady state: a first period of 2000 steps, followed by a 1000 step period in which the events were counted. The theoretical expected value of the number of `AVOID` for this second period is therefore 52.63.



(a) (b)

Figure 5.18: Performance of the agent when interacting with world B (figure 5.13). The performance is measured in the same way as in figure 5.17, taking also into account the number of `Y` stimuli and `APPROACH` actions. The theoretical minimum number of `AVOID` actions is indicated with a dotted line in both plots. These values differ because of the presence of `APPROACH` actions in the second plot. The (a) plot corresponds to the agent version without `APPROACH` actions, while the (b) one corresponds to the complete one.

[12]Given a Markov chain with a state transition probability matrix $P = [p_{ij}]$, where p_{ij} is the probability of going to state s_j given the current state s_i, the stationary state probability vector π is an eigenvector of the transpose of P, corresponding to the eigenvalue of 1. Each component π_i is the probability of being in state s_i, assuming stationarity. The actual probability values are obtained by normalizing the π vector such that $\sum_i \pi_i = 1$. For further details see the textbook [42].

	mean	stddev	min	max
`X` stimuli	0.026	0.3570	0	7
`AVOID` actions	52.35	6.622	29	77
`Y` stimuli	55.76	6.060	36	74

Table 5.6: Statistics of the results after exposing the degraded version of the agent to world B.

Figure 5.18 shows the results obtained from experimentation using world B (figure 5.13). Two versions of the agent were experimented with: a degraded version where the agent does not perform any `APPROACH` actions, and a complete one. The results of these two versions correspond to the (a) and (b) plots. In the former case, the agent is able to prevent the occurrence of negative stimuli after about 5000 time steps, with a rate of `AVOID` actions close to the theoretical expected value of 2.602 (probability of 0.05204 times 50 time steps per period). This value was calculated in the same fashion as before: `AVOID` actions were considered at states 5 and 6 of the POMDP. Note that in this world the agent needs to be able to anticipate with two steps of anticipation at state 5. Table 5.6 shows the statistics for this world using the degraded agent version. This data was collected in a similar fashion as before, except that the first period was raised to 6000 steps, since this world has shown a slower stabilization rate. A residual value of `X` stimuli can still be found. Note that the mean approaches the theoretical expected value of 52.04 (for a 1000 step period).

	mean	stddev	min	max
`X` stimuli	0.225	1.544	0	19
`AVOID` actions	52.73	59.00	11	447
`Y` stimuli	153.2	92.60	0	259
`APPROACH` actions	209.8	203.7	0	658

Table 5.7: Statistics for the results using the complete agent with the world B.

Taking the complete version of the agent into consideration, the results showed both the ability to eradicate negative DV stimuli, and the increase of positive DV stimuli, by means of `APPROACH` actions (plot (b) from figure 5.18). However, two problems are visible. First, the agent performs much more `APPROACH` actions than the ones required to promote positive DV situations. This is a result from the shortcomings of the way DV values are propagated back in time. And second, the rate of performed `AVOID` actions exceeds the theoretical ideal value. Some ideas to overcome these problems are discussed in the next section.

Table 5.7 shows the statistics collected using the complete version of the agent, with world B as above. Taking as ideal the behavior where the agent performs `AVOID` actions at states 5 and 6, and performs `APPROACH` actions at state 0, the probabilities of performing these two actions only on these conditions are 0.03226 for the `AVOID`, and 0.24194 for the `APPROACH` one. The theoretical expected value for the number of `AVOID` actions is therefore 1.613, for the 1000 steps for the second period. This value is clearly below the one observed in the simulations.

To gather an understanding of the agent memory structure, after interacting with the world, a sample run was performed using world A. At the end of 3000 steps, the memory contained 82 perceptual images. From these, only 17 were associated with at least one non-neutral DV outcome with 100% success ratio ($n_{\text{uses}} = n_{\text{succ}}$). The outcome with greater amount of matches (n_{uses}) is shown in figure 5.19. The number of matches of the no-action option is much smaller than the other ones, since as soon as the agent becomes aware of the benefits of performing the `AVOID` action, this outcome is no longer updated. Note how the cognitive images reflect correct generalizations, thus retaining the relevant knowledge about the situations they represent.

Lessons learned

The experiments described above have shown that this second approach was capable of anticipating with a two-step antecedence, thus overcoming a limitation of the first approach (section 5.3.1). However, this was accomplished at the cost of propagating DV values back in time. Since there is no time discount mechanism for the DV values, there is a real danger of the agent performing actions almost in any situation, because of non-discounted propagation of the DV associations. Moreover, the *update* process does not take into consideration whether the DV value it uses is a “real” one (originated by stimuli), or a DV assigned by the *anticipate* phase.

Several modifications to the presented implementation were experimented with, aimed at solving at the implementation level the above-mentioned limitation. However, we failed to achieve yield better results than the ones presented here. Research along this path was abandoned for several reasons. First, because the solutions being worked out differed in nature from the core issues of the agent model. And second, the increased complexity of the implementations made it very hard to draw plausible conclusions about conceptual aspects of the model.

$$
\left|\begin{array}{ll}
\multicolumn{2}{l}{\textbf{ip}} \\ \hline
\text{features}: & \{\langle \mathtt{D}, 2\rangle, \langle \mathtt{B}, 4\rangle\} \\
\text{options}: & \left\{\begin{array}{l}
\left|\begin{array}{ll}
\multicolumn{2}{l}{\textbf{option}} \\ \hline
\text{action}: & \textit{none} \\
\text{outcomes}: & \left\{\left|\begin{array}{l}
\textbf{outcome} \\ \hline
\text{DV}: \quad \textit{negative} \\
\text{ic} \longrightarrow [\mathtt{B}, *, \mathtt{D}, *, \mathtt{X}] \\
n_{\text{succ}} = 12 \\
n_{\text{uses}} = 12
\end{array}\right.\right\}
\end{array}\right. \\
\\
\left|\begin{array}{ll}
\multicolumn{2}{l}{\textbf{option}} \\ \hline
\text{action}: & \mathtt{AVOID} \\
\text{outcomes}: & \left\{\left|\begin{array}{l}
\textbf{outcome} \\ \hline
\text{DV}: \quad \textit{neutral} \\
\text{ic} \longrightarrow [\mathtt{B}, *, \mathtt{D}, */\mathtt{AVOID}, \mathtt{A}] \\
n_{\text{succ}} = 127 \\
n_{\text{uses}} = 127
\end{array}\right.\right\}
\end{array}\right.
\end{array}\right\}
\end{array}\right. \tag{5.28}
$$

Figure 5.19: Sample perceptual image and associated structures, obtained after a run of 3000 steps using the world A.

Related work

The idea of anticipatory systems is not new. Similar techniques, aiming for the explicit representation of anticipations can be found in the literature. Based on the work of the psychologist Joachim Hoffmann, Wolfgang Stolzmann introduced the Anticipatory Classifier Systems (ACS) in 1998 [180], where sensory input in the form of vectors of binary features is used to trigger production rules that anticipate future consequences, both in terms of sensory inputs and valence. These rules are learned as the agent interacts with the environment. ACS accounts for the generalization of rules, and also includes a chaining mechanism allowing for multi-step anticipations. However, unlike the model proposed here, Stotzmann's system uses an external reinforcement signal. Moreover, there are two separate learning mechanisms, one to reinforce the value of rules, and another one to learn anticipations. This model was further studied regarding non-Markovian environments by Mérivier *et al.* [131]. Butz added a genetic mechanism to improve generalization of rules [38]. Another example of research on anticipation is the one of Witkowski. He formulated a theory of anticipatory learning, based on animal learning studies [212].

Chapter 6

Indexing mechanisms

6.1 Motivation

The approach taken in the previous chapter was founded on the goal of experimenting with full systems, by closing the loop with the environment. This implies that all aspects regarding the functioning of the agent interacting with its world have to be tackled. Regardless of how many simplifications are made, in order to make the analysis of the implementation tractable, as well as permitting to draw conclusions from the ideas under scrutiny, issues like perception, memory management, decision-making, and so on, have to be implemented. Although this kind of approach provides fully working systems, the performance of which can be evaluated and compared, it poses many obstacles that deviate the attention from the core issue of the double-representation paradigm. In fact, during that research, much of the development effort was directed to problems unrelated with these core issues.

In contrast, this chapter presents research conducted under a different methodology: the focus is directed exclusively towards one of the mechanisms of the proposed model. In particular, the *indexing mechanism* described in section 4.5 was chosen as object of research. One of the hypothesized consequences of the association between cognitive and perceptual representations is the efficiency gain obtained by, first, a fast perceptual matching process, and second, the guided retrieval of the matching cognitive images from memory. The goal of the indexing mechanism is to obtain these cognitive matches in an efficient manner. The research presented in this chapter attempts to answer several questions: What is the magnitude of the efficiency gain? What is the price to pay for this efficiency gain, and how can it be mitigated? What is the adequate perceptual schema at the perceptual layer?

The research presented in this chapter comprises three different perspectives of the problem. The first one, presented in the following section, approaches it from a probabilistic standpoint, with the goal of expressing the efficiency benefits of the mechanism. Then, section 6.3 considers that the cognitive and perceptual images form two metric spaces. Given an assumption about the relationship between the metric spaces, in plausible agreement with the model properties presented in chapter 4, some theoretic conclusions are drawn, accompanied by illustrative experimental results. Finally, section 6.4 addresses the problem of constructing a perceptual representation (and metric) with the goal of improving the indexing efficiency. To do so, Multidimensional Scaling techniques were employed to devise an optimization algorithm.

It should be stressed that the research methodology taken here pursues, in addition, the goal of developing a formal approach to the presented conceptual model.

6.2 Probability analysis

6.2.1 Preliminaries

Let $\mathcal{S}$ denote the set of all possible stimuli the agent can be exposed to. According to the model, the agent represents stimuli using two different representation schemata. Let the sets $\mathcal{I}_c$ and $\mathcal{I}_p$ stand for the possible cognitive and perceptual images. The association among these images is represented by a set of pairs $\mathcal{M} \subset \mathcal{I}_c \times \mathcal{I}_p$. For simplicity, any given pair of images is either associated or not.

The matching processes between stimuli and images are here considered to yield Boolean results, *i.e.,* given a stimulus and an image, they either match or do not match. Two functions will be used to denote these processes:

$$m_c : \mathcal{S} \times \mathcal{I}_c \longrightarrow \{0, 1\} \tag{6.1}$$

$$m_p : \mathcal{S} \times \mathcal{I}_p \longrightarrow \{0, 1\} \tag{6.2}$$

For a given stimulus $s \in \mathcal{S}$ and a cognitive image $i_c \in \mathcal{I}_c$, $m_c(s, i_c)$ equals 1 if there is a match between the given stimulus and image, and 0 otherwise (and *mutatis mutandis* for the perceptual case).

Consider that the agent is exposed to a random source of stimuli. Therefore, for each possible stimulus $s \in \mathcal{S}$, there is an associated probability $P\{S = s\}$ of the agent being exposed to it, where S stands for the stimulus random variable. For clarity, this function will be written $p(s)$. The above matching functions (6.1) and (6.2) induce two random events, representing the match of a specific image to a stimulus, whose probability functions can be expressed as

$$P\{m_c(i_c)\} = \sum_{s \in \mathcal{S}} p(s)\, m_c(s, i_c) \tag{6.3}$$

$$P\{m_p(i_p)\} = \sum_{s \in \mathcal{S}} p(s)\, m_p(s, i_p) \tag{6.4}$$

since the agent is exposed to only one stimulus at a time.

For each image pair in memory, it is possible to write the joint probability distribution for both of the associated images matching a stimulus

$$\begin{aligned} P\{m_c(i_c), m_p(i_p)\} &= P\{m_c(i_c) \mid m_p(i_p)\}\ P\{m_p(i_p)\} \\ &= \rho\left(\langle i_c, i_p \rangle\right)\ P\{m_p(i_p)\} \end{aligned} \tag{6.5}$$

The term $\rho\left(\langle i_c, i_p \rangle\right)$ above is defined by the conditional probability

$$\rho\left(\langle i_c, i_p \rangle\right) \equiv P\{m_c(i_c) \mid m_p(i_p)\} \tag{6.6}$$

This conditional probability can be interpreted as a measure of the efficiency of a given association $\langle i_c, i_p \rangle$ in memory, since it represents the probability of the associated cognitive image matching the stimulus, given that the perceptual one does match. This quantity can be estimated during the agent's interaction with the environment, in a similar fashion as was performed in the implementation described in section 5.3.2.

It is assumed here that the agent memory consists of a set of associations $\mathcal{M} \subset \mathcal{I}_c \times \mathcal{I}_p$. The sets of all cognitive and perceptual images in memory are thus the projections of this set:

$$\mathcal{M}_c = \{i_c \in \mathcal{I}_c \ : \ \langle i_c, i_p \rangle \in \mathcal{M}\} \tag{6.7}$$

$$\mathcal{M}_p = \{i_p \in \mathcal{I}_p \ : \ \langle i_c, i_p \rangle \in \mathcal{M}\} \tag{6.8}$$

6.2.2 Indexing mechanism

The goal of the indexing mechanism is to identify a subset of cognitive images candidate to the cognitive matching process. Assuming that m_p is computationally much cheaper than m_c, such an indexing mechanism can provide an efficient way of finding cognitive images that match the stimulus, while reducing the number of computations of m_c to a number desirably much smaller than the total number of cognitive images in memory (*i.e.,* the cardinality of $\mathcal{M}_c$).

Noting that the number of cognitive images in memory that match a given stimulus s is $\sum_{i_c \in \mathcal{M}_c} m_c(s, i_c)$, the expected number of cognitive and perceptual matches can be written as

$$\begin{aligned} E_{CM} &= \sum_{s \in \mathcal{S}} p(s) \sum_{i_c \in \mathcal{M}_c} m_c(s, i_c) \\ &= \sum_{i_c \in \mathcal{M}_c} \sum_{s \in \mathcal{S}} p(s)\, m_c(s, i_c) \\ &= \sum_{i_c \in \mathcal{M}_c} P\{m_c(i_c)\} \end{aligned} \tag{6.9}$$

and *mutatis mutandis* for the perceptual case

$$\begin{aligned} E_{PM} &= \sum_{s \in \mathcal{S}} p(s) \sum_{i_p \in \mathcal{M}_p} m_p(s, i_p) \\ &= \sum_{i_p \in \mathcal{M}_p} \sum_{s \in \mathcal{S}} p(s)\, m_p(s, i_p) \\ &= \sum_{i_p \in \mathcal{M}_p} P\{m_p(i_p)\} \end{aligned} \tag{6.10}$$

Given a stimulus s, the first step carried out by the indexing mechanism is to perform a perceptual matching, thus obtaining a set $\mathcal{A}_p(s)$ of perceptual images (a subset of $\mathcal{M}_p$) which match the stimulus. This set is defined by

$$\mathcal{A}_p(s) = \{i_p \in \mathcal{M}_p \;:\; m_p(s, i_p) = 1\} \tag{6.11}$$

Then, the cognitive images from $\mathcal{M}_c$ having at least one association with the ones from $\mathcal{A}_p(s)$ are retrieved from memory. This set (a subset of $\mathcal{M}_c$) is here designated the *active set*, and is denoted $\mathcal{A}_c(s)$

$$\mathcal{A}_c(s) = \left\{i_c \in \mathcal{M}_c \;:\; \exists_{\langle i_c, i_p \rangle \in \mathcal{M}}\, i_p \in \mathcal{A}_p(s)\right\} \tag{6.12}$$

This set corresponds to the cognitive images *indexed* by the perceptual images matching the stimulus s. Next, the agent computes the cognitive match function m_c for each one of these images, resulting on a set of cognitive matches, denoted $\mathcal{R}_{IM}(s)$

$$\mathcal{R}_{IM}(s) = \{i_c \in \mathcal{A}_c(s) \;:\; m_c(s, i_c) = 1\} \tag{6.13}$$

The expected number of indexed cognitive images can thus be written as

$$E_{IM} = \sum_{s \in \mathcal{S}} p(s)\, \|\mathcal{R}_{IM}(s)\| \tag{6.14}$$

It is possible to find lower and upper bounds for E_{IM}, taking into account that a given i_c can be indexed by one or more perceptual images, *i.e.,* the cardinality of the $\mathcal{R}_{IM}(s)$ set can be bounded by

$$\sum_{\langle i_c, i_p\rangle \in \mathcal{M}} \frac{m_c(s,i_c)\, m_p(s,i_p)}{\|\mathcal{B}(i_c)\|} \leq \|\mathcal{R}_{IM}(s)\| \leq \sum_{\langle i_c, i_p\rangle \in \mathcal{M}} m_c(s,i_c)\, m_p(s,i_p) \quad (6.15)$$

where $\mathcal{B}(i_c)$ is the set of all perceptual images associated with i_c, here designated the *base set* of i_c

$$\mathcal{B}(i_c) = \{i_p \in \mathcal{M}_p \;:\; \langle i_c, i_p\rangle \in \mathcal{M}\} \quad (6.16)$$

The upper bound in (6.15) amounts for the total number of pairs that are simultaneously perceptual and cognitive matches, but since a single i_c can be shared by several of these pairs, this number overestimates $\|\mathcal{R}_{IM}(s)\|$. The lower bound divides each term of the sum by the number of pairs sharing that same i_c, thus being an underestimate (at most, there are $\|\mathcal{B}(i_c)\|$ perceptual matches associated with a given i_c). Thus, using (6.14) and (6.15), E_{IM} can be bounded by

$$\sum_{s\in\mathcal{S}} p(s) \sum_{\langle i_c, i_p\rangle \in \mathcal{M}} \frac{m_c(s,i_c)\, m_p(s,i_p)}{\|\mathcal{B}(i_c)\|} \leq E_{IM} \leq \sum_{s\in\mathcal{S}} p(s) \sum_{\langle i_c, i_p\rangle \in \mathcal{M}} m_c(s,i_c) m_p(s,i_p) \quad (6.17)$$

Observing that the joint probability in (6.5) can be also expressed in the form

$$P\left\{m_c(i_c), m_p(i_p)\right\} = \sum_{s\in\mathcal{S}} p(s)\, m_c(s,i_c)\, m_p(s,i_p) \quad (6.18)$$

the inequality (6.17) can be re-written, after exchanging the summation signs and employing expression (6.5), in the following form

$$\sum_{\langle i_c, i_p\rangle \in \mathcal{M}} \frac{\rho\left(\langle i_c, i_p\rangle\right)\; P\left\{m_p(i_p)\right\}}{\|\mathcal{B}(i_c)\|} \leq E_{IM} \leq \sum_{\langle i_c, i_p\rangle \in \mathcal{M}} \rho\left(\langle i_c, i_p\rangle\right)\; P\left\{m_p(i_p)\right\} \quad (6.19)$$

Consider now the conditional probability

$$\sigma\left(\langle i_c, i_p\rangle\right) \equiv P\left\{m_p(i_p) \mid m_c(i_c)\right\} \quad (6.20)$$

corresponding to the probability of a perceptual match of i_p indexing an i_c, given that there is a cognitive match of i_c, for an associated pair $\langle i_c, i_p\rangle$.

Noting that equation (6.9) can be re-written as follows, using the cardinalities of the base sets to cancel the repeated summations of pairs sharing the same i_c

$$\begin{aligned} E_{CM} &= \sum_{i_c \in \mathcal{M}_c} P\{m_c(i_c)\} \\ &= \sum_{\langle i_c, i_p \rangle \in \mathcal{M}} \frac{1}{\|\mathcal{B}(i_c)\|} P\{m_c(i_c)\} \end{aligned} \tag{6.21}$$

the bounds for E_{IM} can be expressed taking (6.17), and using the σ function defined above, as well as the equality $P\{m_c(i_c), m_p(i_p)\} = \sigma(\langle i_c, i_p \rangle) P\{m_c(i_c)\}$.

$$\sum_{\langle i_c, i_p \rangle \in \mathcal{M}} \frac{\sigma(\langle i_c, i_p \rangle)\ P\{m_c(i_c)\}}{\|\mathcal{B}(i_c)\|} \leq E_{IM} \leq \sum_{\langle i_c, i_p \rangle \in \mathcal{M}} \sigma(\langle i_c, i_p \rangle)\ P\{m_c(i_c)\} \tag{6.22}$$

Definition. A cognitive image $i_c \in \mathcal{I}_c$ is said to be *fully-indexed* iff for all stimuli $s \in \mathcal{S}$, a cognitive match ($m_c(s, i_c) = 1$) implies that all associated perceptual images also do match, *i.e.*, $\forall_{i_p \in \mathcal{B}_c(i_c)}\ m_p(s, i_p) = 1$. This also implies that $\forall_{i_p \in \mathcal{B}_c(i_c)}\ \sigma(\langle i_c, i_p \rangle) = 1$

If all cognitive images in memory are fully-indexed, then by using (6.21), the lower bound in (6.22) equals E_{CM}. The expected number of indexed cognitive images is greater than or equal to the expected number of cognitive matches, since the fully-indexed property implies that all matching cognitive images are in the active set.

In this generic setting it is hard to estimate the probability of a cognitive image matching a stimulus, given a set of indexing perceptual images. To understand why, consider a cognitive image i_c indexed by i_p^1 and i_p^2. Assuming that a stimulus matches both of these two perceptual images, what can be concluded about the probability of i_c matching that stimulus? Using the compact notation $C \equiv m_c(s, i_c)$ and $P_k = m_p(i_p^k), k \in \{1, 2\}$, the cognitive match probability can be written as $P\{C|P_1, P_2\}$. Expanding this probability using the Bayesian rule, one can obtain

$$P\{C|P_1, P_2\} = P\{P_1, P_2|C\} \frac{P\{C\}}{P\{P_1, P_2\}} \tag{6.23}$$

If i_c is fully-indexed by both perceptual images, then $P\{P_1, P_2|C\}$ is one, and the above equation can be developed into

$$P\{C|P_1, P_2\} = \rho\left(\langle i_c, i_p^1 \rangle\right) \frac{1}{P\{P_2|P_1\}} \tag{6.24}$$

Since $P\{P_2|P_1\}$ is lower than or equal to one (as well as non-strictly positive), the above probability is greater than or equal to $\rho\left(\langle i_c, i_p^1 \rangle\right)$, which corresponds to the probability of the cognitive match of i_c only knowing that i_p^1 matches. The probability $P\{P_2|P_1\}$, related with the statistical dependence between the matching probabilities of each one of the perceptual images involved, seems to have an effect of leveraging $\rho\left(\langle i_c, i_p^1\rangle\right)$. For instance, if statistical independence is assumed among P_1 and P_2, this leveraging corresponds to the inverse of $P\{P_2\}$.

6.2.3 Computational efficiency

The indexing mechanism aims at computational efficiency, obtained from reducing the quantity of candidate cognitive images, at the price of requiring a prior match of the stimulus with all perceptual images in memory. Designating by J_c and J_p the computational costs of performing a cognitive and a perceptual match, and assuming that they are measured in the same units, the cost of a full cognitive (f.c.) match[1] is

$$J_{fc} = J_c \,\|\mathcal{M}_c\| \tag{6.25}$$

while the cost of a full perceptual (f.p.) match[2] is

$$J_{fp} = J_p \,\|\mathcal{M}_p\| \tag{6.26}$$

Note that a full perceptual match is always required by the indexing mechanism, while the goal of this mechanism is to replace a full cognitive match. It is assumed that $J_c \gg J_p$, *i.e.,* a cognitive match is computationally much more expensive than a perceptual one. Using the indexing mechanism, the total computational cost becomes uncertain, since the active set $\|\mathcal{A}_c(s)\|$ depends on the stimulus. However, its expected value can be computed using

$$\begin{aligned} E\{J_I\} &= J_p \,\|\mathcal{M}_p\| + J_c \, E\{\|\mathcal{A}_c\|\} \\ &= J_p \,\|\mathcal{M}_p\| + J_c \, E_{IM} \end{aligned} \tag{6.27}$$

The (expected) efficiency gain can be assessed as a ratio between this value and J_{fc}.

$$\eta = \frac{E\{J_I\}}{J_{fc}} = \frac{J_p}{J_c}\frac{\|\mathcal{M}_p\|}{\|\mathcal{M}_c\|} + \frac{E_{IM}}{\|\mathcal{M}_c\|} \tag{6.28}$$

The lower *eta*, the higher the efficiency of the mechanism. The bounds found in (6.19) and (6.22) can be used here to find an upper bound for η. Namely,

$$\eta \leq \frac{J_p}{J_c}\frac{\|\mathcal{M}_p\|}{\|\mathcal{M}_c\|} + \frac{1}{\|\mathcal{M}_c\|} \sum_{\langle i_c, i_p\rangle \in \mathcal{M}} \rho\left(\langle i_c, i_p\rangle\right)\; P\{m_p(i_p)\} \tag{6.29}$$

---

[1]This means performing a cognitive match of a stimulus with all cognitive images in memory.

[2]The same as f.c., but regarding perceptual matching.

Note that the first term in (6.28) depends on the agent's computational resources and memory occupancy, while the second term depends also on the statistical properties of the images in the world.

6.2.4 Indexing errors

The indexing mechanism can be understood as an efficient replacement for a full cognitive match, *i.e.,* matching a given stimulus with all cognitive images from memory. However, "there ain't no such thing as a free lunch" [213], and the price to pay is the possibility of this mechanism excluding a matching cognitive image. This happens whenever there is a matching cognitive image in memory that is not associated with any matching perceptual image.

Consider now a stimulus $s \in \mathcal{S}$ subjected to the indexing mechanism. Focusing on a specific cognitive image, say $i_c \in \mathcal{M}_c$, together with the associated perceptual images $\mathcal{B}(i_c)$, let two Boolean events be defined as

$$\mathrm{CM} \equiv \{m_c(s, i_c) = 1\} \tag{6.30}$$

and

$$\mathrm{PM} \equiv \left\{\exists_{i_p \in \mathcal{B}(i_c)}\, m_p(s, i_p) = 1\right\} \tag{6.31}$$

These events correspond to the cognitive match of i_c, and to the presence of a perceptual image which simultaneously matches the given stimulus and indexes i_c. Figure 6.1 depicts the four possible combinations resulting from these two Boolean events. The cases marked with a '$\surd$' raise no concern, since

	¬PM	PM
¬CM	$\surd$	(a)
CM	(b)	$\surd$

Table 6.1: Four possible outcomes of the Boolean events CM and PM defined in the text.

they correspond to the desired behavior: at least one matching i_p indexing a matching i_c, or a non-matching i_c not indexed by any matching of i_p. The (a) case corresponds to the existence of at least one a perceptual match that yields a non-matching cognitive image, thus an inefficiency of the indexing. In other words, the indexing mechanism serves up an image for cognitive matching that will not be found to match. The only impact of such an occurrence is a wasted computation of m_c. On the contrary, the (b) case has a stronger impact, since it amounts to a matching cognitive image which is not indexed by any perceptual image. The indexing mechanism would not consider such i_c as relevant, thus not submitting it to the cognitive match process.

Designating $p_a(i_c)$ the probability of outcome (a), it is possible to develop an upper bound for it as follows:

$$\begin{aligned} p_a(i_c) &= P\{\neg\mathrm{MC} \wedge \mathrm{MP}\} \\ &= P\left\{\neg m_c(i_c) \wedge \bigvee_{i_p \in \mathcal{B}(i_c)} m_p(i_p)\right\} \\ &= P\left\{\bigvee_{i_p \in \mathcal{B}(i_c)} \neg m_c(i_c) \wedge m_p(i_p)\right\} \\ &\leq \sum_{i_p \in \mathcal{B}(i_c)} P\{\neg m_c(i_c) \wedge m_p(i_p)\} \end{aligned} \tag{6.32}$$

This last term can be expressed in function of ρ, since

$$\begin{aligned} P\{\neg m_c(i_c) \wedge m_p(i_p)\} &= P\{\neg m_c(i_c) | m_p(i_p)\} P\{m_p(i_p)\} \\ &= [1 - P\{m_c(i_c) | m_p(i_p)\}] P\{m_p(i_p)\} \\ &= [1 - \rho(\langle i_c, i_p\rangle)] P\{m_p(i_p)\} \end{aligned} \tag{6.33}$$

Therefore, an upper bound of $p_a(i_c)$ can be written as follows

$$p_a(i_c) \leq \sum_{i_p \in \mathcal{B}(i_c)} [1 - \rho(\langle i_c, i_p\rangle)] P\{m_p(i_p)\} \tag{6.34}$$

This upper bound accumulates the contributions from all perceptual images indexing i_c. Pairs with low ρ values contribute to make this kind of errors more probable. Such pairs correspond to perceptual images that index cognitive images with low conditional probability after perceptual match. Procedures to clean the memory from association pairs with low ρ help lowering the probability of this kind of error, thus contributing to the efficiency of the mechanism.

Considering now the probability of case (b), which is the one that raises more concerns with respect to indexing error. Let $p_b(i_c)$ be the probability of this case occurring, then

$$\begin{aligned} p_b(i_c) &= P\{\mathrm{MC} \wedge \neg\mathrm{MP}\} \\ &= P\left\{m_c(i_c) \wedge \bigwedge_{i_p \in \mathcal{B}(i_c)} \neg m_p(i_p)\right\} \\ &= P\left\{\bigwedge_{i_p \in \mathcal{B}(i_c)} m_c(i_c) \wedge \neg m_p(i_p)\right\} \end{aligned} \tag{6.35}$$

The probability of each one of the events in the last conjunction can be written as

$$\begin{aligned} P\{m_c(i_c) \wedge \neg m_p(i_p)\} &= P\{\neg m_p(i_p) | m_c(i_c)\} P\{m_c(i_c)\} \\ &= [1 - P\{m_p(i_p) | m_c(i_c)\}] P\{m_c(i_c)\} \\ &= \left[1 - \frac{\rho(\langle i_c, i_p\rangle) P\{m_p(i_p)\}}{P\{m_c(i_c)\}}\right] P\{m_c(i_c)\} \\ &= P\{m_c(i_c)\} - \rho(\langle i_c, i_p\rangle) P\{m_p(i_p)\} \end{aligned} \tag{6.36}$$

An upper bound of $p_b(i_c)$ can be found by observing that the probability of a conjunction of events is never higher than any of the individual probabilities of the contributing events. Hence, from (6.35)

$$\begin{aligned} p_b(i_c) &\leq \min_{i_p \in \mathcal{B}(i_c)} P\{m_c(i_c) \wedge \neg m_p(i_p)\} \\ &= P\{m_c(i_c)\} - \max_{i_p \in \mathcal{B}(i_c)} \rho(\langle i_c, i_p\rangle) P\{m_p(i_p)\} \end{aligned} \tag{6.37}$$

Unlike the sum in the bound of $p_a(i_c)$ previously developed, this one depends on a maximum over all indexing perceptual images. Thus, a single indexing perceptual image with high ρ may lower this bound (provided the probability of perceptual match is not too low), contributing to make this kind of error less probable.

Taking into account that the (b) kind of errors are much more prejudicial than the (a) ones, one can conclude that it is more important to seek for perceptual images with high values of ρ, than for instance getting rid of the low ones.

6.3 Metric analysis

6.3.1 Preliminaries

The probabilistic approach taken in the previous section simplified the problem assuming that two images of the same schema either match or do not match. So, no further structure was assumed within the images schemata. This section deals with the assumption of a metric relationship, thus giving the $\mathcal{I}_c$ and $\mathcal{I}_p$ a metric space nature. Each one of these two spaces is assumed to be equipped with its own metric function. These metric functions are designated d_c and d_p, mapping pairs of images of the same schema to the set $\mathbb{R}_0^+$ of non-negative real numbers

$$d_c : \mathcal{I}_c \times \mathcal{I}_c \longrightarrow \mathbb{R}_0^+ \tag{6.38}$$

$$d_p : \mathcal{I}_p \times \mathcal{I}_p \longrightarrow \mathbb{R}_0^+ \tag{6.39}$$

Both of these metrics ares assumed to satisfy the usual metric axioms:

(i). $d(x,x) = 0$

(ii). $d(x,y) \geq 0$

(iii). $d(x,y) = d(y,x)$

(iv). $d(x,y) + d(y,z) \geq d(x,z)$

for d being either d_c or d_p, and for any x, y, and z in the respective space ($\mathcal{I}_c$ or $\mathcal{I}_p$). Given a stimulus $s \in \mathcal{S}$, the following two functions stand for the processes of extracting the cognitive and perceptual images from it:

$$\begin{cases} p_c : \mathcal{S} \longrightarrow \mathcal{I}_c \\ i_c = p_c(s) \end{cases} \tag{6.40}$$

$$\begin{cases} p_p : \mathcal{S} \longrightarrow \mathcal{I}_p \\ i_p = p_p(s) \end{cases} \tag{6.41}$$

The setting differs from the one employed in section 6.2 in the following: the matching functions (m_c and m_p) are each one replaced by an image extraction function together with a metric function (p_c and d_c for the cognitive schema, p_p and d_p for the perceptual one).

One key idea behind the cognitive and the perceptual images is for the former to be a complex representation of a stimulus, while the latter is a simple representation of the same stimulus. In other words, the cognitive representation has a greater resolution power than the perceptual one. Motivated by this property, an assumption tenet of this section is raised: the d_c and d_p metrics are such that, for all $s_1, s_2 \in \mathcal{S}$,

$$d_c\left(p_c(s_1), p_c(s_2)\right) \geq d_p\left(p_p(s_1), p_p(s_2)\right) \tag{6.42}$$

Intuitively this means that, for any given two stimuli, their corresponding cognitive images are never closer than their perceptual counterparts.

The goal of the matching mechanism is to find, in the associations memory $\mathcal{M}$ (as defined in section 6.2.1), a pair where the cognitive image best matches a given stimulus received by the agent. The cognitive match, for a cognitive image i_c extracted from a given stimulus s, is here considered to correspond to the minimization of the cognitive distance:

$$\langle i_c^*, i_p^* \rangle = \mathrm{cm}^*(i_c) = \underset{\langle i_c^M, i_p^M \rangle \in \mathcal{M}}{\arg\min}\; d_c(i_c, i_c^M) \tag{6.43}$$

This corresponds to comparing a given cognitive image i_c with all cognitive images in memory, and selecting the one yielding the smallest distance. Note that this procedure is similar to the one termed full cognitive match in section 6.2.3, except that in this case, a single best match is sought.

6.3.2 Indexing

The first step performed by the indexing mechanism consists in narrowing the cognitive match process to a subset of memory pairs. A perceptual match is employed to obtain this subset. A simple strategy to select this subset, denoted $S_p(i_p)$, is considered first: for a given perceptual image i_p, the subset $S_p(i_p) \subseteq \mathcal{M}$ is defined by

$$S_p(i_p) = \left\{ \langle i_c^M, i_p^M \rangle \in \mathcal{M} \ : \ d_p(i_p, i_p^M) \leq T_p \right\} \tag{6.44}$$

where T_p is some threshold. This strategy is here designated *thresholding.* Having obtained $S_p(i_p)$, a cognitive match is then performed, restricted to this subset of pairs:

$$\langle i_c^+, i_p^+ \rangle = \mathrm{cm}^+(i_c) = \underset{\langle i_c^M, i_p^M \rangle \in S_p(i_p)}{\arg\min} \ d_c(i_c, i_c^M) \tag{6.45}$$

The efficiency gain obtained from restricting the cognitive match to $S_p(i_p)$ is higher the fewer pairs are contained in it. The price to pay is the need to evaluate the perceptual distance $d_p(i_p, i_p^M)$ for all memory pairs in $\mathcal{M}$. Therefore, if the perceptual representation is simple to process, in the sense that the perceptual metric d_p is fast to compute, the indexing mechanism is an efficient mechanism.

Using the assumption expressed in (6.42), the following theorem can be trivially proved:

Theorem 1 *Given the d_c and d_p metrics satisfying the condition (6.42), and $S_p(i_p) \subset \mathcal{M}$ as defined in (6.44), whenever $d_c(i_c, i_c^M) \leq T_p$, then $\langle i_c^M, i_p^M \rangle \in S_p(i_p)$.*

Proof. Observing that $d_p(i_p, i_p^M) \leq d_c(i_c, i_c^M) \leq T_p$ it follows that $\langle i_c^M, i_p^M \rangle \in S_p(i_p)$ by definition of $S_p(i_p)$ in (6.44).

An interesting consequence of this theorem is that, whenever the best cognitive match $\langle i_c^*, i_p^* \rangle = \mathrm{cm}^*(i_c)$ (from (6.43)) satisfies $d_c(i_c, i_c^*) \leq T_p$, then, $\langle i_c^*, i_p^* \rangle \in S_p(i_p)$ and $\langle i_c^+, i_p^+ \rangle = \mathrm{cm}^+(i_c)$, obtained from (6.45), are the same ($\langle i_c^*, i_p^* \rangle = \langle i_c^+, i_p^+ \rangle$). This is the same as to say that, under the above conditions, one gets an equally good cognitive match, using only the restricted set $S_p(i_p)$ for the cognitive match, thus preventing the computation of the cognitive distance d_c for all memory pairs in $\mathcal{M}$.

Pre-defining a value for the threshold T_p can be very problematic, since it depends on the metric properties of the domain, as well as on the cognitive and perceptual representations. On one hand, for a too high value of T_p, the $S_p(i_p)$ degenerates on $S_p(i_p) = \mathcal{M}$, if $d_p(i_p, i_p^M) \leq T_p$ for all $\langle i_c^M, i_p^M \rangle \in \mathcal{M}$. On the other hand, for a too low value of T_p, it may happen that the desired i_c^* is such that $d_c(i_c, i_c^*) > T_p$, not only breaking the premises of theorem 1, but also possibly leading to $\langle i_c^*, i_p^* \rangle \notin S_p(i_p)$. Note that this latter case corresponds to an indexing error, as defined in section 6.2.4.

To tackle the difficulty of pre-defining the threshold T_p, an alternative strategy is proposed, which is here designated *N-best:* instead of obtaining $S_p(i_p)$ from T_p, the idea is to include in $S_p(i_p)$ the N_p memory pairs with the lowest perceptual distance $d_p(i_p, i_p^M)$. This means that

$$T_p \in \mathbb{R}_0^+ \quad \text{such that} \quad \|S_p(i_p)\| = N_p \tag{6.46}$$

provided that $\|\mathcal{M}\| \geq N_p$. This results in an upper bound to the number of cognitive distances d_c to be computed[3]. Then, the best cognitive match $\langle i_c^+, i_p^+ \rangle$ in $S_p(i_p)$ can be found using (6.45). Suppose now that there is one memory pair $\langle i_c^S, i_p^S \rangle \in S_p(i_p)$ such that

$$d_p(i_p, i_p^S) \geq d_c(i_c, i_c^+) \tag{6.47}$$

Since $S_p(i_p)$ was defined in such a way that, for all $\langle i_c', i_p' \rangle \in \mathcal{M} \setminus S_p(i_p)$,

$$d_p(i_p, i_p') \geq d_p(i_p, i_p^S) \tag{6.48}$$

it follows that, for all $\langle i_c', i_p' \rangle \in \mathcal{M} \setminus S_p(i_p)$

$$\begin{array}{rcll} d_c(i_c, i_c') & \geq & d_p(i_p, i_p') & \text{by (6.42)} \\ & \geq & d_p(i_p, i_p^S) & \text{by (6.44)} \\ & \geq & d_c(i_c, i_c^+) & \text{by (6.47)} \end{array} \tag{6.49}$$

which means that, by transitivity, $d_c(i_c, i_c') \geq d_c(i_c, i_c^+)$. Together with definition (6.45), this inequality has an interesting consequence: *the $\langle i_c^+, i_p^+ \rangle$ pair minimizes the cognitive distance $d_c(i_c, i_c^M)$ over the whole memory $\mathcal{M}$, in other words, $\langle i_c^+, i_p^+ \rangle = \mathrm{cm}^+(i_c) = \langle i_c^*, i_p^* \rangle = \mathrm{cm}^*(i_c)$.* This result proves the following theorem:

[3] This assertion fails if there are several memory pairs with exactly the same perceptual distance to a given i_p. Choosing $S_p(i_p)$ with the N_p best perceptual matches does not yield a unique solution in this case. However, any of the possible solutions is *a priori* equally good.

Theorem 2 *For a subset $S_p(i_p) \subset \mathcal{M}$ as defined in (6.44), and the minimizations in (6.43) and (6.45), whenever $d_p(i_p, i_p^S) \geq d_c(i_c, i_c^+)$ for some pair $\langle i_c^S, i_p^S \rangle \in S_p(i_p)$, one has $\mathrm{cm}^+(i_c) = \mathrm{cm}^*(i_c)$.*

Both the thresholding and N-best strategies can be seen as stop criteria of the cognitive matching mechanism. Given a stimulus s, the agent computes the perceptual distances of the extracted i_p to the ones stored in memory $\mathcal{M}$. The goal of the thresholding and N-best strategies is to select a subset $S_p(i_p)$, which will be used to perform the computations of the cognitive distances and their minimization.

Theorem 2 can be used as a third stop criterion, in the following way. The $S_p(i_p)$ subset can be constructed incrementally: first considering the memory pair with the smallest perceptual distance $d_p(i_p, i_p^M)$, then with the second best, and so on. Whenever the hypothesis of theorem 2 is met, one has the guarantee that the cognitive match $\mathrm{cm}^+(i_c)$, in the subset $S_p(i_p)$ is the best one globally.

Looking at expression (6.42), the reader might be intrigued by the relationship between the two different metrics. Does it make sense to compare cognitive with perceptual distances? One could scale one of them arbitrarily, thus probably changing the inequality validity. Consider then replacing condition (6.42) with

$$\lambda\, d_c(i_{c1}, i_{c2}) \geq d_p(i_{p1}, i_{p2}) \tag{6.50}$$

for some positive λ, which is the same as scaling the cognitive metric by a scalar λ (or the perceptual one by $1/\lambda$). If condition (6.42) is not satisfied by some pair of metrics d_c and d_p, it may happen that, for some sufficiently large value of λ, condition (6.50) is. However, the hypothesis of theorem 2 can be re-written using the λ scaling value:

$$d_p(i_p, i_p^S) \geq \lambda d_c(i_c, i_c^+) \tag{6.51}$$

one can observe that, as λ increases, this condition becomes more difficult to satisfy. By difficult it is meant that more memory pairs need to be accumulated in $S_p(i_p)$, in order to satisfy (6.51).

However, this argument is reversible in the following sense: assuming that a pair of metrics already satisfy condition (6.42), it may be possible to choose a value of λ, between 0 and 1, such that not only condition (6.50) is true, but also the inequality (6.51) is more easily satisfied. In other words, by scaling the cognitive metric, a stop criterion based on theorem 2 may improve the efficiency of the indexing mechanism. The drawback of scaling the cognitive distance is that it may violate condition (6.42).

The following items summarize and discuss the three strategies of constructing $S_p(i_p)$ that were proposed above:

- The *thresholding* strategy allows adjusting a value up to which a cognitive distance is considered to be a successful match. After setting T_p to that distance, the subset $S_p(i_p)$ is constructed by the means of perceptual distances, and then, by theorem 1, one has the guarantee that, if the best cognitive match distance does not exceed T_p, that match is in $S_p(i_p)$. However, the choice of a threshold value T_p is very sensitive to the range of numerical values taken by the distances: too low or too high values of the threshold can lead to degenerated subsets $S_p(i_p)$ (empty or equal to $\mathcal{M}$);

- The *N-best* strategy addresses the above domain dependency problem by constructing $S_p(i_p)$ based on the N_p best perceptual matches, rather than on a threshold value. Moreover, depending on the computational resources and/or time available to process the stimulus, $S_p(i_p)$ can contain more or less pairs in $S_p(i_p)$ to perform the cognitive match, thus trading time to process for quality of the result (as in anytime algorithms [218]). The drawback is that there is no guarantee of finding the best cognitive match (theorem 1 is not applicable);

- The strategy based on theorem 2 can be used in conjunction with the previous one, in the sense that, when the conditions of the theorem are met, one has the guarantee that the best cognitive match was found; there is no benefit from considering further pairs from memory. However, depending on the magnitude range of the metrics, these conditions may never be met. One can scale the cognitive metric in order to facilitate the satisfaction of those conditions. The drawback of this scaling is that choosing an appropriate scale factor λ is domain dependent.

The following section presents an implementation devised to test empirically the above methodology. The results presented below show some interesting results which illustrate and quantify the efficiency gains that can be obtained by the indexing mechanism.

6.3.3 Illustrative example

The problem considered here consists of the classical hand-written digit recognition problem. Each digit consists of a binary image, being classified with the respective digit symbol: 0 to 9. The task is to perform recognition using the double-representation paradigm, comparing the performance of the pure cognitive match (exhaustive search comparing cognitive images) with the guidance provided by the indexing mechanism. A stimulus is considered successfully recognized whenever the digit corresponding to the stimulus is the same than the one from the memory match.

Implementation

The cognitive image corresponds to the binary image itself ($i_c = s$, *i.e.*, p_c is the identity function, and $\mathcal{I}_c = \mathcal{S}$). Considering W to be the width and H the height (in pixels) of the images, the stimuli and the cognitive images have the form:

$$i_c = s = \begin{bmatrix} b_{11} & \cdots & b_{1W} \\ \vdots & \ddots & \vdots \\ b_{H1} & \cdots & b_{HW} \end{bmatrix} \tag{6.52}$$

where $b_{kl} \in \{0, 1\}$ (for $k = 1, \ldots, H$ and $l = 1, \ldots, W$). The digit foreground is 1, and the background is 0. The perceptual image is a vector of size W (same as the images width) constructed by counting the number of foreground pixels for each column:

$$i_p = \begin{bmatrix} n_1 & \cdots & n_W \end{bmatrix} \tag{6.53}$$

where each $n_k \in \mathbb{N}_0$ (non-negative integers) is computed using the following expression:

$$n_k = \sum_{l=1}^{H} b_{lk}, \quad k = 1, \ldots, W \tag{6.54}$$

The perceptual metric d_p is a simple Euclidean distance between two vectors (the superscripts A and B distinguish each vector involved)

$$d_p(i_p^A, i_p^B) = \sqrt{\sum_{k=1}^{W} \left(n_k^A - n_k^B\right)^2} \tag{6.55}$$

while the cognitive metric corresponds to the Hamming distance between two binary images:

$$d_c(i_c^A, i_c^B) = \sum_{k=1}^{W} \sum_{l=1}^{H} \left| b_{lk}^A - b_{lk}^B \right| \tag{6.56}$$

These two metrics satisfy the metric axioms: the perceptual metric (6.55) is trivial, since it is an Euclidean norm; it is fairly easy to check that the cognitive one (6.56) also verifies them.

Given two stimuli s^A and s^B, the cognitive and perceptual images extracted are denoted $i_c^A = p_c(s^A)$, $i_p^A = p_p(s^A)$, $i_c^B = p_c(s^B)$, and $i_p^B = p_p(s^B)$. Defining X as follows

$$X = \left[d_c(i_c^A, i_c^B)\right]^2 - \left[d_p(i_p^A, i_p^B)\right]^2 \tag{6.57}$$

and using the definitions for d_c, d_p, and n_k above, one obtains

$$X = \left[\sum_{k=1}^{W}\sum_{l=1}^{H}\left|b_{lk}^A - b_{lk}^B\right|\right]^2 - \sum_{k=1}^{W}\left[\sum_{l=1}^{H}\left(b_{lk}^A - b_{lk}^B\right)\right]^2 \tag{6.58}$$

Since $|\sum_i x_i| \leq \sum_i |x_i|$, one obtains the following inequality

$$X \geq \left[\sum_{k=1}^{W}\left|\sum_{l=1}^{H}\left(b_{lk}^A - b_{lk}^B\right)\right|\right]^2 - \sum_{k=1}^{W}\left[\sum_{l=1}^{H}\left(b_{lk}^A - b_{lk}^B\right)\right]^2 = Y \tag{6.59}$$

Defining Y as above, and taking into account that $\left(\sum_i x_i\right)^2 \geq \sum_i x_i^2$ (because $x_i \geq 0$)[4], one can write

$$Y \geq \sum_{k=1}^{W}\left[\sum_{l=1}^{H}\left(b_{lk}^A - b_{lk}^B\right)\right]^2 - \sum_{k=1}^{W}\left[\sum_{l=1}^{H}\left(b_{lk}^A - b_{lk}^B\right)\right]^2 = 0 \tag{6.60}$$

Therefore $X \geq Y \geq 0$, which is the same to say that

$$\left[d_c(i_c^A, i_c^B)\right]^2 \geq \left[d_p(i_p^A, i_p^B)\right]^2 \tag{6.61}$$

Since the metrics satisfy the metric axiom (ii), it can be concluded that

$$d_c(i_c^A, i_c^B) \geq d_p(i_p^A, i_p^B) \tag{6.62}$$

which satisfies condition (6.42). This means that the theoretical results obtained in section 6.3.2 can be applied in this example.

Results

In the following experiments a test-set from an well known Machine Learning repository was employed[5]. This test-set consists of 1934 samples of

[4]Whenever $a, b \geq 0$, one can write $(a+b)^2 = a^2 + 2ab + b^2 \geq a^2 + b^2$. Therefore, by induction, $\left(\sum_i x_i\right)^2 \geq \sum_i x_i^2$, provided that $x_i \geq 0$ for all i.

[5]Optical Recognition of Handwritten Digits, from E. Alpaydin, C. Kayna, URL (in 2006): `ftp://ftp.ics.uci.edu/pub/machine-learning-databases/optdigits`

handwritten digits (0 to 9), scanned into binary images of 32 by 32 pixels ($W = H = 32$). Figure 6.1 shows four examples of these digits, together with the corresponding perceptual images extracted from each one of them. From these samples, a training and a test set were randomly picked up, forming two disjoint sets. An usual cross-validation procedure was employed to evaluate the recognition performance.

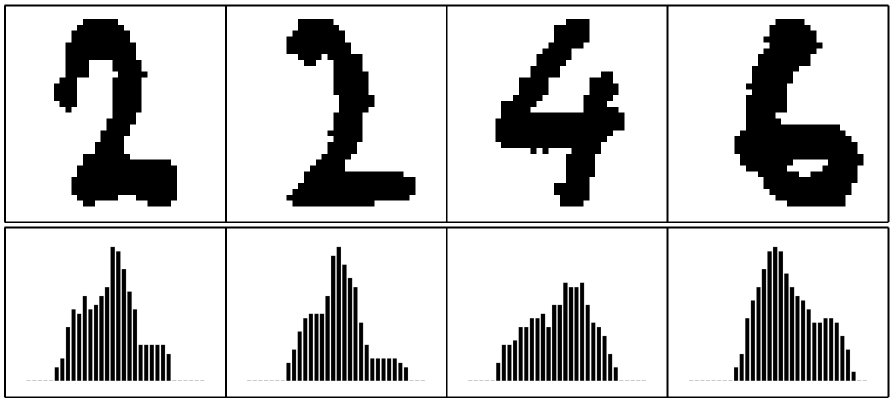


Figure 6.1: Four example digits from the data-set. The top row shows the bitmaps (corresponding to the stimuli and cognitive images), while the bottom row shows bar graphs of the extracted perceptual images. The size of each bar is proportional to the number of black pixels found on the corresponding column of the bitmap.

The training process consists in running through all elements of the training set, and for each one of them, storing in memory the pair of the cognitive and perceptual images extracted. In order to evaluate the recognition success ratio, the corresponding digit symbol was also attached to each pair.

The results shown here were obtained by averaging over 10 trials, each one using disjoint training and test sets, containing 1500 and 200 digits respectively, randomly chosen from the pool of 1934 patterns. Four matching mechanisms were evaluated:

1. Pure perceptual matching — the memory pair where the *perceptual image* is closest to the one extracted from the stimulus:
$$\langle i_c^{(1)}, i_p^{(1)} \rangle = \underset{\langle i_c^M, i_p^M \rangle \in \mathcal{M}}{\arg\min}\; d_p\left(i_p, i_p^M\right) \tag{6.63}$$

2. Pure cognitive matching — the memory pair where *cognitive image* is closest to the one extracted from the stimulus:
$$\langle i_c^{(2)}, i_p^{(2)} \rangle = \underset{\langle i_c^M, i_p^M \rangle \in \mathcal{M}}{\arg\min}\; d_c\left(i_c, i_c^M\right) \tag{6.64}$$

3. Guided cognitive matching (indexing), using *thresholding* — like the pure cognitive matching but where the set of memory pairs considered is restricted by a pure perceptual matching. This restriction is based on thresholding the perceptual distances, as discussed above:

$$\begin{cases} S_p(i_p) = \left\{ \langle i_c^M, i_p^M \rangle \in \mathcal{M} \ : \ d_p(i_p, i_p^M) \leq T_p \right\} \\ \langle i_c^{(3)}, i_p^{(3)} \rangle = \underset{\langle i_c^M, i_p^M \rangle \in S_p(i_p)}{\arg\min} \ d_c\left(i_c, i_c^M\right) \end{cases} \tag{6.65}$$

4. Guided cognitive matching (indexing), using *N-best* — like the previous one, but the restriction corresponds to choosing the N_p best perceptual matches, as previously discussed:

$$\begin{cases} T_p \in \mathbb{R}_0^+ \quad \text{such that} \quad \|S_p(i_p)\| = N_p \\ \langle i_c^{(4)}, i_p^{(4)} \rangle = \underset{\langle i_c^M, i_p^M \rangle \in S_p(i_p)}{\arg\min} \ d_c\left(i_c, i_c^M\right) \end{cases} \tag{6.66}$$

The implementation of the matching mechanisms (3) and (4) was based on the following algorithm: *for a $\langle i_c, i_p \rangle$ pair extracted from a given stimulus,*

1. Compute the perceptual distance $d_p(i_p, i_p^M)$ for all pairs in memory $\langle i_c^M, i_p^M \rangle \in \mathcal{M}$;

2. Sort the pairs in memory $\mathcal{M}$ by increasing order of its perceptual distance $d_p(i_p, i_p^M)$;

3. Scan the obtained list until the corresponding stop criterion is met: the threshold value in (3), or the number of best matching pairs in (4).

To illustrate the relationship between the cognitive and perceptual distances obtained after the sorting operation in step 2, figure 6.2 plots these two distances between a randomly chosen input stimulus, and images from the test set. Thus, the memory images bearing closer perceptual distances with the given stimulus are the leftmost ones. Note that the bottom left area of the upper cloud of points, are the ones corresponding to the memory images with both small cognitive and perceptual distances.

Table 6.2 shows the results for the matching mechanisms (1) and (2). There is a clear trade-off between an extremely slow cognitive matching, with a high success rate, and the fast perceptual matching leading by itself to poor results.

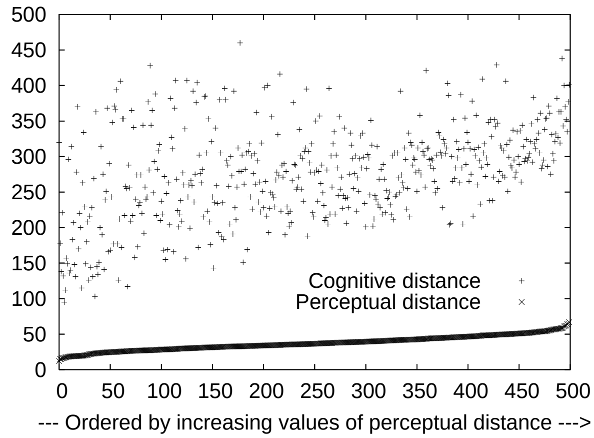


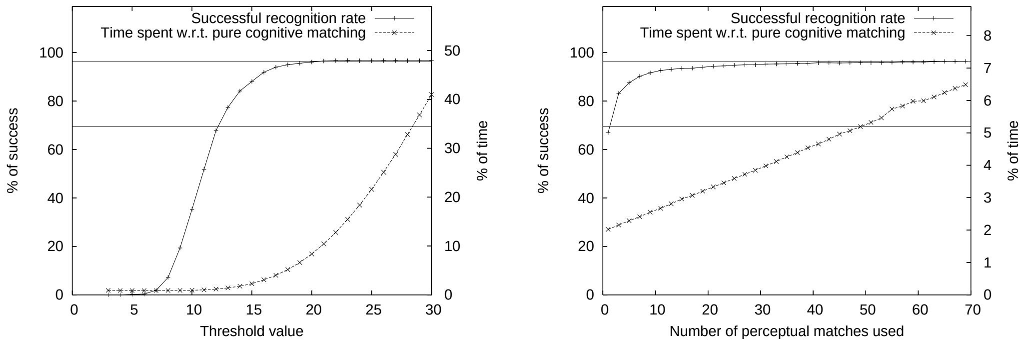
Figure 6.2: Cognitive and perceptual distances of a typical stimulus, with respect to the memory pairs in memory, sorted by increasing perceptual distances $d_p(i_p, i_p^M)$. The lower dots sketching the (visually) continuous line correspond to the perceptual distances.

mechanism	min (%)	avr (%)	max (%)	time
cognitive	94.0	96.45	100.0	130
perceptual	66.0	69.45	72.0	1

Table 6.2: Results for the pure cognitive and perceptual matching. The minimum, average, and maximum success rates for all trials, as well as the computational time (ratio, perceptual=1) are shown.

Regarding the other two mechanisms (3) and (4), evaluating the indexing mechanism in this domain, the plots in figure 6.3 show the success rates in function of the relevant parameter. Using thresholding (figure 6.3a), the parameter is the threshold value (T_p), and using N-best (figure 6.3b), the parameter is the number of perceptual matches (the closest ones) considered for indexing (N_p).

The two plots shown in figure 6.3 express basically the same outcome, since they both result from the indexing mechanism. What makes them different is the dependency on the parameter: in (a) the dependency on T_p is explicit, while in (b) the relationship is implicit. It is easy to realize that there is a non-linear monotonic relation between the horizontal axis, since each value of T_p leads to some number of pairs in $S_p(i_p)$. And this number of pairs increases monotonically with T_p.



(a) (b)

Figure 6.3: Success rates obtained using the indexing mechanism: (a) *thresholding*, as function of the threshold value; (b) *N-best*, as function of the number of perceptual matches used. The values for execution time are expressed as a percentage of the time taken by the pure cognitive match (first line of table 6.2). In both plots, the two horizontal lines denote the average success rate for the pure cognitive (higher) and pure perceptual (lower) matching mechanisms. Note the different scales in the rightmost axis (time) of each plot.

In the (b) plot, the relative execution time increases linearly with the N_p parameter, because it determines how many cognitive distances have to be computed. However, it is interesting to note that the success rate rises above 90% when just about 10 perceptual matches (in $S_p(i_p)$) are used. At this point, the cognitive match after indexing is using about 2.5% of the time taken by a pure cognitive match.

One possible use of theorem 2 is as a stop criterion for an incremental construction of $S_p(i_p)$. A direct implementation of it alone led to poor results: the subset $S_p(i_p)$ often degenerated to $\mathcal{M}$, because the assumptions of the theorem were rarely met. Therefore, a scaling of the cognitive metric as in (6.51) was attempted. In doing so, the condition (6.62) does not hold always. Figure 6.4 shows the obtained results in function of the scaling parameter λ.

This plot shows the sensitivity of the results with respect to the λ parameter. For too low values of λ, (6.62) is easily satisfied, leading to the same success rate as a pure perceptual match. For too high values of λ, the stop criterion tends to be never used, degenerating in a slow pure cognitive match (the processing time raises at a significant rate, in direction of 100%). However, for a range of λ values it is possible to obtain very good results, keeping the processing time at low levels. Note that in these results, the stop criterion

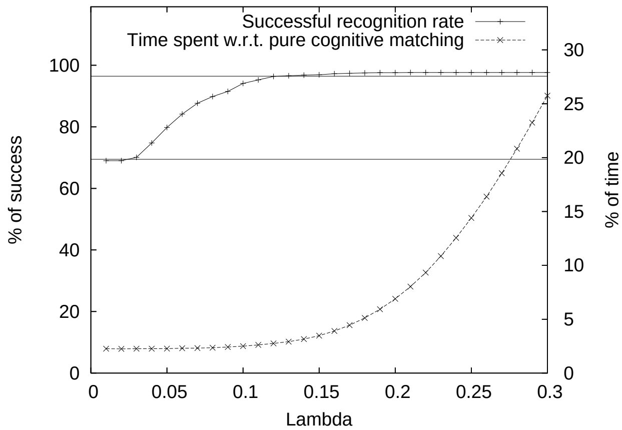


Figure 6.4: Results obtained using theorem 2 as a stop criterion (same plot format as in figure 6.3).

based on theorem 2 replaces entirely either the thresholding or the N-best strategies.

6.3.4 Discussion

The experimental results presented above show significant efficiency gains: for instance, identical recognition success rates (about 95%) resulted from using as few as 5% of the time taken by an exhaustive cognitive search. This corresponds to restricting the cognitive match to a few tenths of memory pairs, from a pool of 1500 pairs in memory.

It should be stressed that the goal of the experiments is not to obtain a good recognition rate. The recognition rates obtained here are the sole merit of the Hamming distance used in the cognitive metric[6]. The goal of the indexing mechanism is rather to approach the level of performance of the cognitive metric, *without the necessity of evaluating the cognitive metric for all memory pairs.* A good indexing mechanism should obtain results of similar quality as the ones from a pure cognitive match, with much fewer computations than exhaustive matching with all cognitive images in memory.

The theorems derived above do not demand the satisfaction of all metric axioms by the two distance functions. They are based on order relations

[6] In the test set used, the Hamming distance yields a good performance level. However, the Hamming distance is obviously unable to deal with invariance to translation, rotation, and scaling of the digits. A review of state-of-the-art methods for pattern recognition can be found in [101].

among distances. A *partially ordered set* (poset) is a binary relation $\leq$, together with a nonempty set $\mathcal{P}$, satisfying the following axioms:

(i) $a \leq a$

(ii) $a \leq b$ and $b \leq a$ imply $a = b$

(iii) $a \leq b$ and $b \leq c$ imply $a \leq c$

for every $a, b, c \in \mathcal{P}$. If in addition $a \leq b$ or $b \leq a$ for any $a, b \in \mathcal{P}$, it is called a *chain*, or a *totally ordered set* [61]. From an algebraic standpoint, any two functions $d_c : \mathcal{I}_c \times \mathcal{I}_c \to \mathcal{P}$ and $d_p : \mathcal{I}_p \times \mathcal{I}_p \to \mathcal{P}$, where $\mathcal{P}$ is a chain with order relation $\leq$, satisfy both theorems. This can be checked by examining that the formalism utilized to prove the theorems only makes use of properties of chains[7]. Therefore, the theoretical results obtained in this section can be generalized to domains beyond metric spaces.

6.4 Learning a perceptual metric

6.4.1 Motivation

The previous section presented a formulation of the indexing mechanism under the assumption that the matching of the cognitive and perceptual images are performed in metric spaces. The cognitive and the perceptual distance functions were assumed to be known *a priori*. However, for an agent to cope with unknown and dynamic environments, it may be desirable to be able to adapt the perceptual representation and metric. The research presented in this section concerns the following problem: how to construct a perceptual representation (and metric) with the goal of optimizing the indexing efficiency [208, 206, 207]. In other words, the ideal perceptual representation and metric are the ones that yield small perceptual distances iff the corresponding cognitive distances are also small. To do so, two strategies are explored. One corresponds to adapting a perceptual metric, via a set of parameters, such that cognitive proximity implies perceptual nearness:

$$d_c(i_c^1, i_c^2) < d_c(i_c^1, i_c^3) \Rightarrow d_p(i_p^1, i_p^2) < d_p(i_p^1, i_p^3) \tag{6.67}$$

for all arbitrary image pairs $\langle i_c^k, i_p^k \rangle$ $(k = 1, 2, 3)$. The second strategy addresses the improvement of the perceptual representation, in the following

[7] Care has to be taken with the operator arg min in (6.43) and (6.45), noting that any nonempty finite subset of a chain has a minimum (descending chain condition).

sense. Assuming that the perceptual representation is a vector of features extracted from stimuli, when these features are not sufficiently representative to satisfy (6.67), the goal is to upgrade the perceptual representation with new, more representative, features. Both of these strategies are approached here using Multidimensional Scaling techniques [49].

The Multidimensional Scaling (MDS) is a technique from statistics with the goal of recovering the coordinates of a set of points, provided that the all distances among them are known [49]. This technique proved an interesting inspiration to the research presented in this section. But before presenting it, the basic MDS techniques are reviewed first[8].

6.4.2 Multidimensional Scaling

Having its origins in the field of statistics, the Multidimensional Scaling (MDS) comprises a group of techniques sharing a common goal: given a set of n objects, together with a measure of dissimilarity among them, to assign point coordinates to each one of the objects, in a metric space, so that their distances approximate as much as possible the given dissimilarities. For a pair of objects r and s, the dissimilarity between them is a real value, being denoted by δ_{rs}. It is here assumed that two properties are satisfied: $\delta_{rr} = 0$ (identity) and $\delta_{rs} = \delta_{sr}$ (symmetry).

In the context of the MDS, the terms *dissimilarity* and *distance* have distinct and specific meanings: the dissimilarities set $\{\delta_{rs}\}$ is given beforehand, and may or may not constitute a metric, while the distances result from the metric space $\mathcal{E}$ where the points live, and whose coordinates are sought.

The MDS techniques distinguish between two major categories: metric and non-metric. The difference between them lies in the kind of constraints imposed upon the distances on $\mathcal{E}$.

Metric Multidimensional Scaling

In the metric MDS category, the goal is to find coordinates for the objects such that the distances among them, in $\mathcal{E}$, approximate as much as possible the given dissimilarities, according to a continuous monotonic function f

$$d_{rs} \approx f(\delta_{rs}) \tag{6.68}$$

where d_{rs} stands for the Euclidean distance in $\mathcal{E}$ between the objects r and s.

The simplest case corresponds to an identity function f (termed classic metric MDS), where a solution can be found (under certain restrictions)

[8] This following review is mainly based on the book [49].

using spectral decomposition techniques, as follows. Consider for now that the n objects are points in $\mathbb{R}^p$ with coordinates $\mathbf{x}_r = (x_{r1}, \ldots, x_{rp})^T$ for $r = 1, \ldots, n$. Let a matrix $\mathbf{B}$ denote the matrix of inner products among them

$$[\mathbf{B}]_{rs} = b_{rs} = \mathbf{x}_r^T \mathbf{x}_s \tag{6.69}$$

Assuming that the centroid of the point set is at the origin, *i.e.,* $\sum_{r=1}^{n} x_{ri} = 0$ for all values of i, the $\mathbf{B}$ matrix can be written in the following form, after some algebraic manipulation

$$\mathbf{B} = \mathbf{HAH} \tag{6.70}$$

where the matrices $\mathbf{H}$ and $\mathbf{A}$ are defined by

$$[\mathbf{A}]_{rs} = a_{rs} = -\frac{1}{2} d_{rs}^2 \tag{6.71}$$

$$\mathbf{H} = \mathbf{I} - n^{-1}\mathbf{1}\mathbf{1}^T \tag{6.72}$$

In this last equation, $\mathbf{I}$ stands for a $n \times n$ identity matrix, and $\mathbf{1}$ denotes a vector of n ones $(1, \ldots, 1)^T$.

Note that the matrix $\mathbf{B}$ is thus constructed from the distances d_{rs}. Next, a spectral decomposition of this matrix is performed

$$\mathbf{B} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T \tag{6.73}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with the eigenvalues

$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \tag{6.74}$$

and the corresponding eigenvectors in the columns of $\mathbf{V}$

$$\mathbf{V} = [\mathbf{v}_1 | \cdots | \mathbf{v}_n] \tag{6.75}$$

It can be mathematically shown that the rank of $\mathbf{B}$ is p, hence $n - p$ eigenvalues are zero. Using the non-zero ones, $\boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, as well as the corresponding eigenvectors $\mathbf{V}_1 = [\mathbf{v}_1 | \cdots | \mathbf{v}_p]$, $\mathbf{B}$ can be expressed in the following way

$$\mathbf{B} = \mathbf{V}_1 \boldsymbol{\Lambda}_1 \mathbf{V}_1^T \tag{6.76}$$

Taking (6.69) into consideration, the original coordinates can then be recovered (apart from an arbitrary translation and an uniform scaling) using

$$[\mathbf{x}_1 | \cdots | \mathbf{x}_n] = \mathbf{V}_1 \boldsymbol{\Lambda}^{\frac{1}{2}} \tag{6.77}$$

where $\boldsymbol{\Lambda}^{\frac{1}{2}} = \mathrm{diag}(\lambda_1^{\frac{1}{2}}, \ldots, \lambda_p^{\frac{1}{2}})$.

In the classic metric MDS, the distances d_{rs} above are set to the dissimilarities δ_{rs}. If the $\mathbf{B}$ matrix thus obtained is a positive semi-definite matrix of rank p, then a configuration in a p-dimensional Euclidean space can be found that faithfully replicates the given distances. Otherwise, a constant can be added to all dissimilarities (except δ_{rr}), in order to facilitate $\mathbf{B}$ to become positive semi-definite.

In the general case, if an Euclidean space with fewer than p dimensions is sought, then one can employ the eigenvectors with the highest eigenvalues to derive the point coordinates. Moreover, there are techniques to help determine the dimension of the Euclidean distance, for instance, by comparing the magnitude of the eigenvalues to select a subset of them considered more relevant.

One common applications of MDS techniques is the visualization of datasets. In this case, depending on whether a 2-D or a 3-D view is desired, the dimension of the Euclidean space is set to either 2 or 3.

Besides this classical approach, there is a least squares formulation, assuming a continuous monotonic function f mapping dissimilarities to distance values. The goal is then to find the objects coordinates satisfying (6.68). This can be accomplished by minimizing the cost function

$$S = \frac{\sum_{r \neq s} w_{rs} \left(d_{rs} - f(\delta_{rs})\right)^2}{\sum_{r \neq s} d_{rs}^2} \tag{6.78}$$

The weights w_{rs} can be appropriately chosen, for instance, such that large dissimilarities dominate lower ones (e.g., $w_{rs} = \delta_{rs}^{-1}$). The function f has to be given beforehand. For instance, a parametric function $f(\delta_{rs}) = \alpha + \beta\delta_{rs}$ is a straightforward choice. The above cost function has to be minimized numerically, with respect to the coordinates, as well as to the parametrization of f (α and β).

Nonmetric Multidimensional Scaling

The second category of MDS techniques is the one used in the research presented here, since it imposes looser constraints between the dissimilarities and the metric distances. Here, the functional relationship (6.68) is replaced by a monotonic one:

$$\delta_{rs} < \delta_{tu} \Rightarrow d_{rs} \leq d_{tu} \tag{6.79}$$

for any two pairs of objects (r, s) and (t, u). In other words, for any two pairs of objects, the distance between the two more dissimilar ones will not be less than the distance between the more similar ones. Here it is assumed that no two dissimilarities are equal (ties). However, the case of ties among the dissimilarities can be easily accounted for, as described at the end of this section.

The commonly used Kruskal approach to nonmetric MDS will be taken here [108, 109] (alternative approaches are reviewed by Trevor Cox *et al.* in [49]). This method introduces a third set of distances — $\{\hat{d}_{rs}\}$ — thus dividing the original problem in two: in the first one, the $\{\hat{d}_{rs}\}$ distances are adjusted such that they are monotonically coherent with the dissimilarities:

$$\delta_{rs} < \delta_{tu} \Rightarrow \hat{d}_{rs} \leq \hat{d}_{tu} \tag{6.80}$$

and in the second one, the point coordinates are adjusted such that the distances between them approximate, as much as possible, the distances $\{\hat{d}_{rs}\}$. The latter is performed by the minimization of a cost function, termed *stress* (S), assessing the degree of this approximation:

$$S = \sqrt{\frac{S^*}{T^*}} \qquad \begin{aligned} S^* &= \sum_{r,s} \left(d_{rs} - \hat{d}_{rs}\right)^2 \\ T^* &= \sum_{r,s} d_{rs}^2 \end{aligned} \tag{6.81}$$

These summations are performed for $r = 1, \ldots, (n-1)$ and $s = (r+1), \ldots, n$, since $\delta_{rr} = 0$ and $\delta_{rs} = \delta_{sr}$.

This formulation uses a Minkowski metric for the distances determination, defined by

$$d_{rs} = \left[\sum_{i=1}^{p} |x_{ri} - x_{si}|^{\lambda}\right]^{\frac{1}{\lambda}} \tag{6.82}$$

where $\lambda > 0$ (e.g., $\lambda = 2$ corresponds to the usual Euclidean metric), and p is the metric space dimension ($x_{ri} \in \mathbb{R}^p$).

Given a distances set $\{d_{rs}\}$, the set $\{\hat{d}_{rs}\}$ is obtained by an *isotonic regression* algorithm, which can be proved [49] to satisfy two conditions:

1. the set $\{\hat{d}_{rs}\}$ satisfies condition (6.80), and
2. it minimizes S^* (and thus S) with respect to the distances $\{\hat{d}_{rs}\}$, restricted to the previous condition.

For the description of the isotonic regression procedure, an alternative notation is required here, for clarity: all distances sets are renumbered, replacing the two indices by a single one, in such a way that $\delta_i < \delta_{i+1}$, for $i = 1, \ldots, n(n-1)/2$. This establishes a mapping from the rs indexing schema to a single index i, thus ordering all distance sets by ascending dissimilarities. Note that the same mapping is also used to re-index the distance sets $\{\hat{d}_{rs}\}$ and $\{d_{rs}\}$. Both notations will be used below, depending on convenience; they are easily distinguishable by the number of indices.

Using this ordering, let the cumulative sums of distances be defined by

$$D_i = \sum_{j=1}^{i} d_j \tag{6.83}$$

for $i > 0$, and $D_0 = 0$. As the distances are positive, D_i is monotonically increasing (with i). The isotonic regression is defined as the greatest convex minorant of these cumulative sums. This can be easily visualized by imagining a plot of the D_i points, and then stretching a string attached to these points. Figure 6.5 illustrates this idea with an example. The resulting graph (dashed line) is convex, coinciding with the graph D_i in as many points as possible, as well as in the first and last ones. Wherever convexity prohibits such coincidence, the points are obtained by linear interpolation.

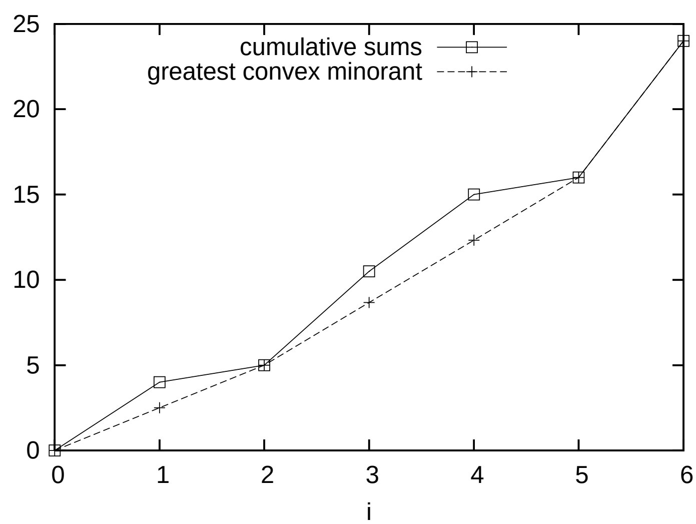


Figure 6.5: Illustration of a sample isotonic regression, corresponding to the greatest convex minorant, obtained from a given set of cumulative sums (adapted from figure 3.1, page 47 of [49]). In this example, 4 points are considered, yielding a total of 6 distances among them. Note that a point at the origin $(0, 0)$ corresponds to $D_0 = 0$..

The resulting points, here denoted $\{\hat{D}_i\}$, give rise to the set $\{\hat{d}_i\}$ using $\hat{d}_i = \hat{D}_i - \hat{D}_{i-1}$, for $i > 0$, *i.e.,* the slope of each segment of the $\hat{D}_i$ graph. Convexity of $\hat{D}_i$ implies that the $\{\hat{d}_i\}$ is (non-strictly) monotonically increasing, thus satisfying (6.80).

The isotonic regression algorithm works by partitioning the set of distances $\{d_i\}$ into blocks of consecutive ones (with respect to i), within which the values of $\hat{d}_i$ are the same (thus resulting in a linear interpolation). The process is iterative, starting with the most fine grained partitioning, *i.e.*, blocks of a single point (singletons). The idea is to find, iteratively, whether two consecutive blocks ought to be merged. Considering two consecutive partitions of indices $P_k = \{r, \ldots, s\}$ and $P_{k+1} = \{(s+1), \ldots, u\}$ (because of the equality of the $\hat{d}_i$ distances within each block, $\hat{d}_r = \cdots = \hat{d}_s$, and $\hat{d}_{s+1} = \cdots = \hat{d}_u$), whenever $\hat{d}_s > \hat{d}_{s+1}$, these two partitions are merged into a single one, *i.e.*, $P'_k = \{r, \ldots, u\}$. The distances $\hat{d}'_i$ within the new partition are all equal to the average of the distances d_i whose index belong to the new partition:

$$\hat{d}'_r = \cdots = \hat{d}'_u = \frac{1}{u - r + 1} \sum_{i=r}^{u} d_i \tag{6.84}$$

The algorithm iterates until no two blocks can be merged. Taking the example depicted in figure 6.5, the result of the isotonic regression would be

$$\{\{1, 2\}, \{3, 4, 5\}, \{6\}\} \tag{6.85}$$

Further computational aspects of the isotonic regression can be found in Kruskal's companion paper [109].

The stress can then be determined using (6.81). Since the goal consists in minimizing the stress with respect to the point coordinates, a gradient descent method can be employed, since S is differentiable. The stress gradient ∇S is obtainable by differentiating S with respect to the point coordinates. For a coordinate i of the point $\mathbf{x}_u$, the corresponding gradient components are then

$$\frac{\partial S}{\partial x_{ui}} = S \sum_{r,s} (\delta^{ru} - \delta^{su}) \left[\frac{d_{rs} - \hat{d}_{rs}}{S^*} - \frac{d_{rs}}{T^*} \right] \frac{|x_{ri} - x_{si}|^{\lambda - 1}}{d_{rs}^{\lambda - 1}} \operatorname{sgn}(x_{ri} - x_{si}) \tag{6.86}$$

where δ^{ij} is the usual Kronecker function (1 iff $i = j$, 0 otherwise), and $\operatorname{sgn}(x)$ is the *signum* function ($+1$ or -1 depending on whether x is positive or negative).

The complete nonmetric MDS algorithm consists of the following steps. Note that all distances are computed using the Minkowski metric given above (6.82).

1. Start with an initial configuration of points, e.g. randomly distributed with an uniform distribution;

2. Normalize[9] the configuration by translation and uniform scaling, such that the centroid is at the origin and the mean square distance from it is unitary;

3. Compute the distance set $\{d_{rs}\}$;

4. Perform the isotonic regression to obtain the intermediate set of distances $\{\hat{d}_{rs}\}$;

5. Find the gradient of the stress with respect to all the coordinates of all points (concatenated into a vector $\mathbf{x}$)

$$\nabla S = \frac{\partial S}{\partial \mathbf{x}} \tag{6.87}$$

   When the norm of the gradient is below some pre-defined threshold ϵ, stop the algorithm. Note that the stress values have no units, being commonly expressed as percentages.

6. Perform a step of the gradient descent method using

$$\mathbf{x}(t+1) = \mathbf{x}(t) - \eta \, \frac{\nabla S(t)}{\|\nabla S(t)\|} \tag{6.88}$$

   where η is the descent rate. Kruskal proposes this rate to vary along the descent, using a heuristic update rule;

7. Go to step 2.

In sum, this algorithm alternates between adjusting the $\hat{d}_i$ distances using the isotonic regression procedure, and descending through the stress gradient. As in many optimization problems, the descent path converges to a local minimum (but for pathological conditions, such as too large step size), but not necessarily to the global one. One way of mitigating this sub-optimality is to repeat the algorithm from several different (random) initial configurations, choosing the one yielding the least final stress.

The above algorithm works under the assumption that there are no ties among the dissimilarities set, *i.e.,* $\delta_i \neq \delta_j$ for all $i \neq j$. When this is not the case, one of the following two strategies can be employed:

1. whenever $\delta_i = \delta_j$, do not constrain the order relationship between $\hat{d}_i$ and $\hat{d}_j$. Implementation: for each subset of equal dissimilarities, rearrange the indices such that the distances d_i are in ascending order within that block;

[9] Note that the stress S is invariant to translation and uniform scaling.

2. whenever $\delta_i = \delta_j$, constrain $\hat{d}_i$ and $\hat{d}_j$ to be equal. Implementation: initialize isotonic regression with blocks of distances of equal dissimilarities (thus forcing the $\hat{d}_i$ to be equal within each of them), instead of singleton sets.

According to [49], the second approach is less satisfactory than the first one, since it imposes more restrictive constraints than the other.

6.4.3 Methodology

Recall that the goal of the indexing mechanism is to provide good candidates for cognitive matching, using the perceptual representation. Therefore, assuming that the matching is metric, a good perceptual representation is one which satisfies the implication (6.67) for all image pairs. Note that this goal is similar to the MDS one, once one considers the cognitive distances to be the dissimilarities, and the perceptual ones to be the distances among objects. However, there are differences. In the case of the MDS, the metric is given while the object coordinates are sought. In the case of the indexing, the object coordinates (perceptual images) are given, while the (perceptual) metric is subject to adaptation.

To do so, and in agreement with the goals set in section 6.4.1, a gradient descent is proposed, within the framework of MDS, with respect to a parametrization of the perceptual metric, instead of with respect to the point coordinates. Thus, the perceptual metric is assumed to depend on a vector of parameters. For instance, these parameters can assign a degree of relevance to each feature of the perceptual representation. Regarding the construction of additional perceptual features, it is proposed to append the perceptual image with a pre-determined amount of additional components. These components represent the values that the new features ought to take, for each of the perceptual images in the training set. Their values are randomly initialized, and subject to gradient descent as in the nonmetric MDS. However, nothing is said about how to obtain these values from new stimuli. The idea advanced here is to utilize the obtained values to construct a regression model. That regression model can then be used to obtain the new feature values for new stimuli.

The above assumes that a training set of cognitive and perceptual image pairs is employed. This training set can be seen as the agent memory after storing an amount of perceptual and cognitive image pairs. It is further assumed that the associations among the perceptual and cognitive images in $\mathcal{M}$ are one-to-one, *i.e.* neither two perceptual images index the same cognitive one, nor two cognitive images are indexed by the same perceptual one. The reason for this is to avoid complications arising from a dissimilarity/distance being zero, while the corresponding distance/dissimilarity is non-zero.

Metric adaptation

The adaptation of the metric parameters will be considered first. Let the perceptual images be made of vectors of N numeric features each, using the notation

$$i_p^r = (x_{r1}, \ldots, x_{rN})^T \tag{6.89}$$

for the r'th perceptual image, from a training set of T image pairs ($r = 1, \ldots, T$). The perceptual metric d_p is parametrized in the form $d_p(i_p^r, i_p^s; \Theta)$ with a vector of q parameters $\Theta = (\theta_1, \ldots, \theta_q)^T$. The gradient of S with respect to one of these parameters, say θ_k, is given by

$$\frac{\partial S}{\partial \theta_k} = S \sum_{r,s} \left(\frac{d_{rs} - \hat{d}_{rs}}{S^*} - \frac{d_{rs}}{T^*} \right) \frac{\partial d_{rs}}{\partial \theta_k} \tag{6.90}$$

Note that this gradient has q components (the number of parameters of the perceptual metric), while the number of components of the gradient employed by the nonmetric MDS equals the number of points times their dimension.

The cost function considered here is the sum of the MDS stress, as defined above, with a regularization term penalizing the absolute values of the metric parameters:

$$J = S + \xi \sum_{i=1}^{q} |\theta_i| \tag{6.91}$$

$$\frac{\partial J}{\partial \theta_k} = \frac{\partial S}{\partial \theta_k} + \xi \operatorname{sgn}(\theta_k) \tag{6.92}$$

The summation term in (6.91), weighted by ξ, in the cost function for two reasons. First, if the stress is invariant to a perceptual component, the stress gradient with respect to the corresponding weight would be zero, and therefore the initial parameter value would stay at the same value during the descent. The second reason is due to the quadratic contribution of the parameters to the stress. In order to prevent a slow asymptotic convergence to zero (and therefore never reaching zero exactly), the gradient of their absolute values forces them to approach zero at a faster pace[10] In sum, this term contributes to reduce the number of non-zero parameters θ_i. The usefulness

[10]Numerically this makes parameters close to zero to oscillate around zero, so, they are set to zero once they become negative. The implementation further forces them to stay at zero thereafter.

of this will become evident in the next section, where each parameter θ_i is used to weight a feature. Therefore, once one such parameter is zero, the corresponding feature can be deleted from the perceptual representation. The more metric parameters are zero, the lower the dimensionality of the perceptual space.

Weighting features

A simple option for the perceptual metric is to start with the traditional Euclidean distance, and weighting each dimension i with a parameter θ_i.

$$d_{rs} = \sqrt{\sum_{i=1}^{N} \theta_i^2 (x_{ri} - x_{si})^2} \tag{6.93}$$

This parametrization corresponds to assigning a weight (relevance) to each perceptual feature before computing the Euclidean metric. Moreover, when the algorithm assigns a zero weight to a feature, that feature can be deleted from the perceptual representation, since it is irrelevant.

The partial derivative of this distance with respect to each parameter is then

$$\frac{\partial d_{rs}}{\partial \theta_k} = \frac{(x_{rk} - x_{sk})^2}{d_{rs}} \, \theta_k \tag{6.94}$$

Discovering new dimensions

To address the second problem stated in section 6.4.1, the following approach was taken: each perceptual image in the training set is augmented with a pre-defined amount of new components, in such a way that the stress is reduced. This is accomplished within the framework of nonmetric MDS, by subjecting these components to the stress gradient descent. Doing so corresponds to solving the problem only partially, since a method to derive these components for new stimuli is still required. However, regression methods can be considered to estimate these values.

Using y_{ri} to designate the i-th appended component of the perceptual image i_p^r, each image takes thus the form

$$i_p^r = (x_{r1}, \ldots, x_{rN}, y_{r1}, \ldots, y_{rM})^T \tag{6.95}$$

for M added components. These components are initialized randomly and subject to the gradient descent method as in the nonmetric MDS. The outcome of this minimization is a set of values that these new features ought to take for each i_p, such that the stress is (locally) minimized.

The features values thus found concern solely the images in the training set. Nothing is said about the feature values for new stimuli. With the objective of determining a way to derive these values for new stimuli, a regression technique is proposed here. Given a set of training images, together with the y_{ri} values, a regression technique can create a model for the data. That model can then be used to estimate feature values for new stimuli. This requires not only the "output" values (the y_{ri}) but also the "input" data. Theoretically, any regression technique can be used here. An arbitrary regression model can be written as

$$y_{ri} = R(u_{r1}, \ldots, u_{rl}) \tag{6.96}$$

for l input variables $u_{r1}, \ldots, u_{rl}$. For instance, these variables can include features extracted from the stimulus.

The simple metric in (6.93) can be augmented to take into account the new features, in the following way

$$d_{rs} = \sqrt{\sum_{i=1}^{N} \theta_i^2 (x_{ri} - x_{si})^2 + \sum_{i=1}^{M} (y_{ri} - y_{si})^2} \tag{6.97}$$

The additional components are not weighted since doing so would just add redundant degrees of freedom.

To express the gradient of the stress one can consider a generic parameter vector Λ from which the distances d_{rs} depend:

$$\Lambda = [\lambda_1 \cdots \lambda_{N+TM}]^T = [\theta_1 \cdots \theta_N | y_{11} \cdots y_{TM}]^T \tag{6.98}$$

The partial derivative of the stress S with respect to a parameter λ_i is thus

$$\frac{\partial S}{\partial \lambda_i} = S \sum_{r,s} \left(\frac{d_{rs} - \hat{d}_{rs}}{S^*} - \frac{d_{rs}}{T^*} \right) \frac{\partial d_{rs}}{\partial \lambda_i} \tag{6.99}$$

If λ_i corresponds to a metric parameter θ_i, the expression (6.94) can be used. Otherwise, if it corresponds to a component y_{ui}, then

$$\frac{\partial d_{rs}}{\partial y_{ui}} = \frac{y_{ri} - y_{si}}{d_{rs}} \left(\delta^{ru} - \delta^{su} \right) \tag{6.100}$$

The gradient of the cost function J is then

$$\frac{\partial J}{\partial \lambda_k} = \begin{cases} \frac{\partial S}{\partial \lambda_k} + \xi \, \mathrm{sgn}(\lambda_k) & \text{if } \lambda_k \text{ corresponds to a metric parameter} \\ \frac{\partial S}{\partial \lambda_k} & \text{otherwise} \end{cases} \tag{6.101}$$

The dimension of the gradient thus obtained is therefore $N + TM$, recalling that N is the number of the metric parameters, T is the number of images in the training set, and M the number of new features added to each perceptual image.

Different possibilities for obtaining the input variables $u_{r1}, \ldots, u_{rl}$ using the regression model in (6.96) can be considered. One is to use a larger pool of features, selecting a subset of them for the base perceptual image vector, and to use a disjoint set to feed the regression model. Another is to extract more features from the stimulus. This extraction can be, for instance, guided by the y_{ri} values, e.g. a parametrized feature extraction algorithm, whose parameters can be tuned based on the values obtained for the training set.

A third possibility consists in using the features already present in the perceptual image. To understand why this may not be redundant, consider the following. An Euclidean distance between two vectors $\mathbf{x}_r$ and $\mathbf{x}_s$ can be written in matrix form as

$$d_{rs} = \sqrt{(\mathbf{x}_r - \mathbf{x}_s)^T(\mathbf{x}_r - \mathbf{x}_s)} \tag{6.102}$$

Consider now scaling each component $k = 1, \ldots, N$ with a positive weight θ_k. The vectors in the new space formed by the above scaling can be written as $\mathbf{x}'_r = \mathbf{\Theta}\mathbf{x}_r$, where the $\mathbf{\Theta} = \mathrm{diag}(\theta_1, \ldots, \theta_N)$ is a diagonal matrix. The Euclidean distance in this space is thus

$$d'_{rs} = \sqrt{(\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{\Theta}^2(\mathbf{x}_r - \mathbf{x}_s)} \tag{6.103}$$

using the notation $\mathbf{\Theta}^2 = \mathrm{diag}(\theta_1^2, \ldots, \theta_N^2)$. Note that this is another way of writing the metric (6.93), meaning that the metric parametrization is equivalent to scaling the space coordinates with $\mathbf{\Theta}$. If the regression model (6.96) is linear, and the input variables are the components of the perceptual vector, each vector of new features $\mathbf{y}_r = (y_{r1}, \ldots, y_{rM})^T$ is then

$$\mathbf{y}_r = \mathbf{R}\mathbf{x}_r \tag{6.104}$$

where the matrix $\mathbf{R}$ $(M \times N)$ represents the linear model. Considering a new vector formed by appending each weighted vector $\mathbf{x}'_r$ with these components

$$\mathbf{x}''_r = \begin{bmatrix} \mathbf{x}'_r \\ \mathbf{y}_r \end{bmatrix} = \begin{bmatrix} \mathbf{\Theta}\mathbf{x}_r \\ \mathbf{R}\mathbf{x}_r \end{bmatrix} \tag{6.105}$$

the Euclidean distance in this augmented space takes the form

$$\begin{aligned} d''_{rs} &= \sqrt{(\mathbf{x}''_r - \mathbf{x}''_s)^T(\mathbf{x}''_r - \mathbf{x}''_s)} \\ &= \sqrt{(\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{\Theta}^2(\mathbf{x}_r - \mathbf{x}_s) + (\mathbf{y}_r - \mathbf{y}_s)^T(\mathbf{y}_r - \mathbf{y}_s)} \\ &= \sqrt{(\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{\Theta}^2(\mathbf{x}_r - \mathbf{x}_s) + (\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{R}^T\mathbf{R}(\mathbf{x}_r - \mathbf{x}_s)} \\ &= \sqrt{(\mathbf{x}_r - \mathbf{x}_s)^T(\mathbf{\Theta}^2 + \mathbf{R}^T\mathbf{R})(\mathbf{x}_r - \mathbf{x}_s)} \end{aligned} \tag{6.106}$$

This expansion can be interpreted in the following way: first, note that the Euclidean metric on the space formed by (6.105) is equivalent to the metric (6.97), when a linear regression model is used, because it equals the second line of the expansion; and second, the last expression is a non-diagonal metric on the original space of perceptual images. In other words, this last metric is more general, since the matrix $\mathbf{\Theta}^2 + \mathbf{R}^T\mathbf{R}$ has more degrees of freedom than the diagonal one (6.103). For instance, a generic metric $\sqrt{(\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{G}(\mathbf{x}_r - \mathbf{x}_s)}$ can be replicated by d''_{rs}, as long as $\mathbf{G}$ can be factorized[11] into $\mathbf{R}^T\mathbf{R}$ and $\mathbf{\Theta}$ is zero.

In sum, the regression over already existing features in the perceptual image is capable of overcoming limitations of a perceptual metric. Alternatively, one could replace this metric by a more general one, from the start. However, this leads to a larger amount of parameters (as long as $M < N - 1$), thus complicating the stress minimization procedure: the diagonal metric matrix has N degrees of freedom, a general metric matrix $\mathbf{G}$ has N^2, while the one from (6.106) has $N(1 + M)$. Adding new features, one at at time, provides an incremental way of increasing the complexity of the perceptual metric.

Grouping images

A single metric parametrization may not be satisfactory for the whole stimulus space. So, it could be interesting to consider specializing the metric according to characteristics of the stimuli. That could be accomplished by first finding a partition of the training set, and by applying the gradient descent separately, to each one of the resulting subsets. The outcome would be a separate parametrization, found by a minimization of the stress within that subset.

For instance, imagine a situation where separating the gradient descent into two non-trivial[12] subsets would lead to very small stress values in each one of them, way below the global stress (when determined over the whole original set). It is possible in this case that the global stress (considering all images) did not drop below a certain value, because the gradient did counter-balance two trends towards two distinct parametrizations.

Using such a mechanism would require, first, a clustering technique to isolate subsets from the training set, in such a way that the sum of all the stress values would be minimal, and second, a classification engine to identify to which cluster a new stimulus would fit better (in terms of minimizing the cognitive distance in the end).

[11] According to the Cholesky decomposition, any real square symmetric positive-definite matrix $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ is a lower triangular matrix.

[12] Assuming the cardinality of the subsets is much larger than the number of features.

Several approaches in this direction were attempted, but no satisfactory results could be obtained. These approaches can be divided into two groups: (1) constructive approaches, staring from many small subsets, and proceeds by merging them according to a criterion, and (2) partitioning approaches, starting with the global training set, and trying to devise partitions according to a criterion, based for instance on the stress within candidate subsets. The major cause of impasse encountered consisted in the following: how can one determine whether a certain operation is profitable, in terms of stress reduction, without actually performing the gradient descent? A brute force approach of trying out all possible partitions would be intractable. Moreover, the stress, which is unit-less, is usually smaller for small subsets, since there is a lower number of constraints imposed by the dissimilarities, and therefore, there is an inherent bias towards partitioning into the degenerate set of singletons.

Algorithm

Taking into account the above considerations, and based on the standard nonmetric MDS algorithm [49], we propose the following one:

1. Start with an initial vector of variables Λ. For instance, the metric parameters θ_k ($k = 1, \ldots, N$) may be initialized to all ones, and the additional components $\{y_{ri}\}$ (for $r = 1, \ldots, T$, and $i = 1, \ldots, M$) may be drawn from a uniform distribution;

2. Normalize the metric parameter vector $\Theta = (\theta_1, \ldots, \theta_N)^T$ to unit norm, since the stress is invariant to scaling of this vector. The additional components $\{y_{ri}\}$ are, however, not normalized[13];

3. Compute the distances $\{d_{rs}\}$ using the parametrized perceptual metric (6.97);

4. Perform the isotonic regression on the cumulative sums (6.83), to obtain the set of distances $\{\hat{d}_{rs}\}$;

5. Compute the cost (6.91); if its value is below a threshold ϵ, stop the algorithm (stopping criterion);

6. Find the gradient of the cost function (6.91) w.r.t. the variables vector Λ;

[13]Normalizing them would constrain *a priori* the relative weights of the additional components w.r.t. the original features in (6.97). Normalizing the parameters vector prevents its norm from growing or shrinking because of numerical errors. Moreover, because of (6.97), the additional components do not grow/shrink arbitrarily.

7. Perform a step of the gradient descent method using

$$\Lambda(t+1) = \Lambda(t) - \eta \, \nabla J(t) \tag{6.107}$$

where η is the descent rate.

8. Go to step 2.

Further details concerning the gradient descent method can be found in section 6.4.5.

6.4.4 Related work

Methods related to the presented methodology include traditional dimension reduction algorithms such as Principal Component Analysis (PCA) [105], Independent Component Analysis (ICA) [100], Local Linear Embedding (LLE) [159], Non-negative Matrix Factorization (NMF) [115], and Latent Semantic Analysis (LSA) [64]. All of these methods fall into the category of unsupervised learning. They all have in common the analysis of a given (training) data set in order to obtain a reduced dimension space, capable of representing the original data set. For instance, the PCA selects a linear space spawned by the orthogonal axis along which the training points have greater variance, thus being able to reconstruct the original dataset with a smaller number of dimensions (the ones with higher eigenvalues); ICA is similar to PCA but uses a criterion of statistical independence instead; LLE relies on the identification of a lower-dimensional manifold capable of representing the original dataset. All of these methods employ information from the dataset itself. Their distinctions consist broadly in the criteria used to select the most relevant dimensions.

The proposed method, on the contrary, uses information originating from a second representation space. The algorithm relies on the cognitive distance among the instances. Thus, depending on how this metric is expressed, different weights can be attributed to components, and thus different relevance values are extracted.

Another method for extracting relevance, can be found in [183], which utilizes Rate Distortion Theory on an information theoretic framework. This method consists in coding a signal x such that it preserves as much relevant information as possible about another signal y. Relevant Component Analysis (RCA) employs side-information, in the form of equivalence relations, to learn a Mahalanobis metric [18]. Other applications of metric learning can be found in classification [66] and clustering [217].

6.4.5 Experimentation

Optimization technique

Plain gradient descent methods are slow to converge. Thus, to speed up convergence, a momentum term was added to (6.107), in a similar fashion to back-propagation techniques for neural networks [160]:

$$\Delta\lambda_k(t+1) = -\eta_k \frac{\partial S}{\partial \lambda_k} + \alpha_k\, \Delta\lambda_k(t) \tag{6.108}$$

$$\Delta\lambda_k(t+1) = \lambda_k(t+1) - \lambda_k(t) \tag{6.109}$$

where η_k and α_k are parameters that control the rate of the descent and the momentum inertial effect. This dependence on k (introduced here) expresses that two pairs of parameters were used, $\langle \eta_\theta, \alpha_\theta \rangle$ and $\langle \eta_y, \alpha_y \rangle$, depending on whether a given λ_k corresponds to a component of the parameter vector or a new component. For instance, if $\lambda_1 = \theta_1$, then $\eta_1 = \eta_\theta$ and $\alpha_1 = \alpha_\theta$. The momentum term accelerates convergence, since it tends to increase the descent rate whenever the descent is successively performed in the same direction.

Another issue requiring consideration is a stop criterion to detect whether to stop the descent. To do so, the total cost J is monitored along the descent. Two moving averages, for two contiguous windows, are computed in each step. If the percentual change is below a threshold, the descent is terminated. The purpose of the moving average is to filter out possible small oscillations that might occur during final stabilization (e.g., numerical round-off errors). Thus, even if the cost oscillates around a stable value, the detector is able to put a stop to the descent. The parameters for this criterion are the moving average window size, and the variation threshold (set to 20 steps and 0.1% respectively in the conducted experiments).

Since the gradient is differentiable, faster optimization methods can be considered, such as the Newton-Raphson method. This method requires the computation of the Hessian matrix in each step

$$[\mathbf{H}]_{ij} = \frac{\partial^2 J}{\partial \theta_i \partial \theta_j} \tag{6.110}$$

with the corresponding update rule

$$\mathbf{u}(t+1) = \mathbf{u}(t) - \gamma\, \mathbf{H}^{-1}(\mathbf{u}(t))\, \nabla J(\mathbf{u}(t)) \tag{6.111}$$

This is a second order procedure, since it uses the second derivative of the cost function, in contrast to the first order plain gradient descent. In experimental validation, this method failed to bring significant improvements, probably because the algorithm interleaves the gradient descent with the isotonic regression procedure, thus modifying the structure of the optimization problem at each step. Hence, the simpler gradient descent method was employed.

Performance metric

In order to evaluate the results, a measure of performance called *eval-order* was introduced, aimed at assessing how well an indexing mechanism would behave. This assessment is performed using a test set disjoint from the training set employed in the gradient descent (cross-validation). Inspired by the *N-best* indexing algorithm described in [204], the eval-order is defined in the following way: given a cognitive and perceptual images pair $\langle i_c, i_p \rangle$, determine all perceptual distances from it to images in the perceptual memory (*i.e.,* the training set); then, after sorting all these images with respect to the perceptual distances, determine which n-th image pair $\langle i_c^k, i_p^k \rangle$ on the resulting ordered list has the minimum cognitive distance to $\langle i_c, i_p \rangle$. In the ideal case, it corresponds to the first one, and thus an eval-order of one. Higher values correspond to worse performance.

This measure is admittedly reductionist, since it disregards what happens in sub-optimal situations, for instance, when the second best cognitive match corresponds to a relatively small perceptual distance. But since the eval-order is averaged over a large training set, this problem is assumed to be at least partially mitigated.

Normalization and initial conditions

The features in the perceptual images were all (training and test sets) normalized to zero mean and unit variance, prior to any experiment. Unless otherwise stated, the parametrization Θ of the perceptual metric was initialized to all ones. The additional components, when used, were initialized with a uniformly distributed random configuration, as in the nonmetric MDS algorithm.

Synthetic data set

To validate the proposed methodology, a simple test-bed was devised. Random points $\mathbf{x} \in \mathbb{R}^c$ (simulating stimuli) are uniformly drawn from an hypercube of unit side length. The cognitive images $i_c \in \mathbb{R}^c$ were set to the components of $\mathbf{x}$ multiplied by random coefficients $\{w_1, \ldots, w_c\}$ between 0 and 2 each

$$i_c = \mathrm{diag}(w_1, \ldots, w_c)\, \mathbf{x} = \mathbf{W}\mathbf{x} \tag{6.112}$$

These coefficients introduce different degrees of relevance of the components of i_c. The perceptual images were obtained by concatenating two vectors: the p first components of $\mathbf{x}$ multiplied by a second set of random weights $\{v_1, \ldots, v_c\}$ $(p \leq c)$; and n random numbers between 0 and 1 (noise, uniform distribution). Thus, the perceptual images have $p + n$ components.

$$i_p = \left(\frac{[\mathrm{diag}(v_1, \ldots, v_p)|\mathbf{0}]\ \mathbf{x}}{\mathbf{u}} \right) = \left(\frac{\mathbf{Vx}}{\mathbf{u}} \right) \tag{6.113}$$

where $\mathbf{0}$ stands for a matrix of zeros of appropriate dimension, and $\mathbf{u}$ for the vector of noise. The weights in $\mathbf{W}$ and $\mathbf{V}$, randomly drawn (uniform distribution) from the $[0; 2]$ interval, together with the numbers c, p, and n, define a *world,* represented by a tuple

$$\langle c, p, n, \mathbf{W}, \mathbf{V} \rangle \tag{6.114}$$

Figure 6.6 illustrates graphically the above computations.

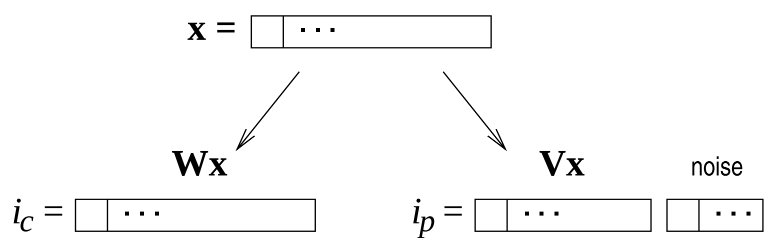


Figure 6.6: Illustration of how the cognitive and perceptual images are determined from a random vector $\mathbf{x}$.

The cognitive distances were computed by applying a non-linear monotonic function to the Euclidean distance among the images

$$\begin{aligned} d_c(i_c^r, i_c^s) &= f\left(\|i_c^r - i_c^s\|\right) \\ &= f\left((\mathbf{x}_r - \mathbf{x}_s)^T \mathbf{W}^2 (\mathbf{x}_r - \mathbf{x}_s)\right) \end{aligned} \tag{6.115}$$

while the perceptual ones employ the metric (6.93). By construction of the nonmetric MDS, the form of the f function is irrelevant[14], thus it was set to the identity function, except in one of the experiments.

Note that the cognitive distances in (6.115), assuming f to be the identity, are invariant to rotations after applying the weights (the $\mathbf{W}$ diagonal matrix)

[14]As long as it is strictly monotonic.

to the $\mathbf{x}$ points (*i.e.,* $i'_c = \mathbf{RWx}$), but the converse is not necessarily true (*i.e.,* $i''_c = \mathbf{WRx}$). This is so because

$$\begin{aligned} &\|i^r_c - i^s_c\| = \\ (\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{W}^2(\mathbf{x}_r - \mathbf{x}_s) &= (\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{W}^T\mathbf{R}^T\mathbf{RW}(\mathbf{x}_r - \mathbf{x}_s) \\ &= (\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{W}^T\mathbf{W}(\mathbf{x}_r - \mathbf{x}_s) \\ &\neq (\mathbf{x}_r - \mathbf{x}_s)^T\mathbf{R}^T\mathbf{W}^2\mathbf{R}(\mathbf{x}_r - \mathbf{x}_s) \end{aligned} \tag{6.116}$$

since rotation matrices are unitary ($\mathbf{R}^{-1} = \mathbf{R}^T$).

Results

The experimentation was conducted in two phases. In the first phase (experiments 1–6), no additional perceptual components were considered, leaving the experiments with additional ones to the second phase (experiments 7–10).

For the following experiments, a single world in the form (6.114) was generated with random parameters. The weights in $\mathbf{V}$ are irrelevant, because the normalization of the perceptual vectors cancels their scaling effect.

Experiment 1. The algorithm was run for 100 generated training sets with the same world parameters, each one containing 100 patterns (and thus 4950 distances among them). The world dimensions were $c = p = 10$ and $n = 3$. For each training set, a test set containing 100 patterns was also generated, for posterior eval-order assessment. The descent parameters for (6.108) were $\eta_\theta = 0.01$ and $\alpha_\theta = 0.8$.

Figure 6.7 shows the results: the bar graph (a) represents the weights in $\mathbf{W}$, while (b) represents the perceptual metric weights found by the algorithm. Note that the latter values faithfully represent the relative importance of the $\mathbf{x}$ coordinates in the cognitive metric. The observed extinguishing of the third weight is due to the combined effect of its diminished importance (*i.e.,* low value in $\mathbf{W}$), and the penalization of non-zero weights in (6.91). Moreover, the last three components (noise) were all zero, thus showing a successful capability of discarding irrelevant features.

Concerning the eval-order assessment, the results are shown in table 6.3. These are consolidated values, obtained in the following way: for each run, a training set and a test set were randomly generated, as mentioned above; then, the weights obtained in each run were tested against the test set (cross-validation), determining the mean, minimum, and maximum values of the obtained eval-orders for all images in the test set. The results shown here correspond to the mean of these means[15] (central tendency), the minimum

[15] This is equivalent to a mean over all image pairs, since all test sets have size.

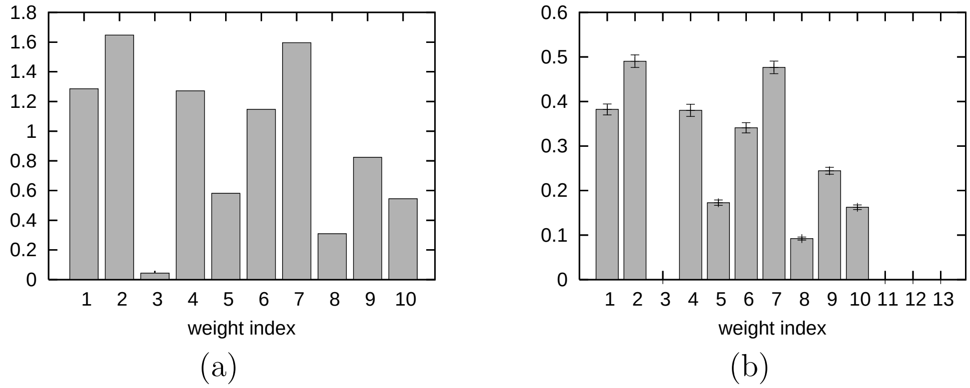


Figure 6.7: Results of Experiment 1 (100 training sets randomly drawn): (a) the weights **W** of the employed world, and (b) the metric weights found by the algorithm (error bars represent standard deviation values). Both vectors are normalized to unit norm, in order to be comparable.

of all minima, and the maximum of all maxima (worst case of eval-order). These results show a significant improvement in the eval-order performance after using the metric weights found by the algorithm. Namely, the worst case (maximal eval-order) went down from 92 to just 2. Note that the test set has 100 image pairs, therefore, the worst possible eval-order value is 100.

metric	mean	min	max
unweighted	11.44	1	92
weighted	1.003	1	2

Table 6.3: Evaluation of Experiment 1 in terms of the eval-order performance metric.

Experiment 2. All gradient descent methods are prone to local minima, unless the cost function is convex. In the present case, the cost function structure changes at each step, because the isotonic regression alters the $\{\hat{d}_{rs}\}$ distances. In order to determine the sensitivity of the solution with respect to local minima, the algorithm was run with initial metric weights other than all ones. In each run, the parameter were initialized with random values uniformly distributed between 0 and 2. The (a) plot of figure 6.8 shows a much higher variance in all weights, notably on the noise ones, when compared with figure 6.7. However, this effect was found to be a consequence of outlier runs. These outliers corresponded to runs where one of the weights, other than the noise ones (indexes 11 to 13) and the component index 3 (see above), were initialized close to zero, and were set to zero during the descent.

Because zero weights are kept at zero (as explained before), the algorithm was unable to find an appropriate result fitting the data. Leaving out these outliers, using the criterion by which a run is considered outlier whenever any of the noise weights is non-zero at the end, the results in plot (b) of the same figure approach significantly the ones of the previous experiment. This criterion considered 11% of the runs as outliers. Table 6.4 shows the results for the eval-order assessment, including the ones before and after removing the outlier runs.

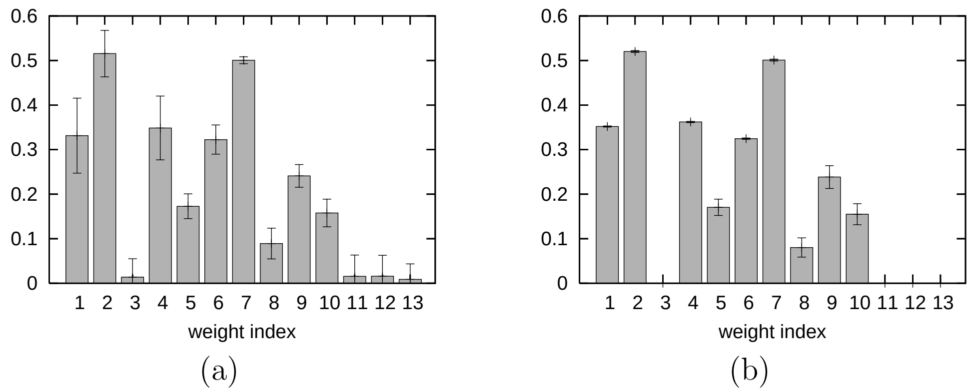


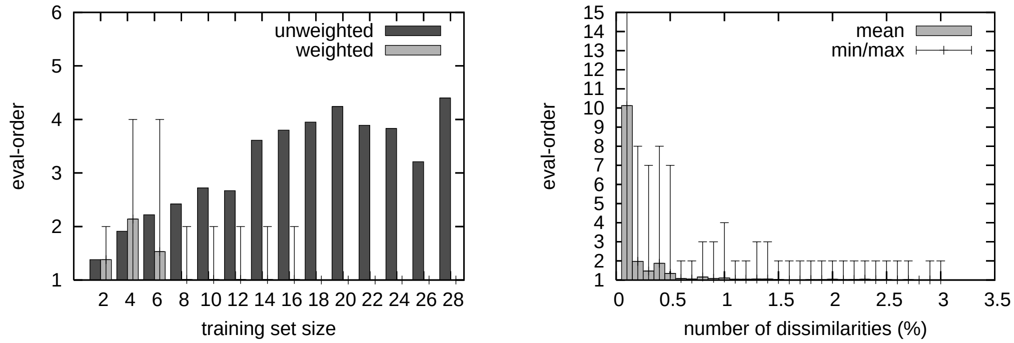
Figure 6.8: Obtained weights in Experiment 2 (random initial conditions): (a) before, and (b) after removing outlier runs (11%). Error bars denote standard deviations, as before.

metric	mean	min	max
unweighted	10.17	1	73
weighted	1.217	1	41
w/o outliers	1.017	1	9

Table 6.4: Results for the eval-order performance metric for Experiment 2.

Experiments 3 and 4. These experiments addressed the impact of reducing the training material. In Experiment 3, smaller training sets were used, while in Experiment 4, a subset of all dissimilarities were used (while keeping the training set size). The results are shown in terms of eval-order values with respect to either the training set size (Experiment 3, figure 6.9a), or the percentage of dissimilarities employed (Experiment 4, figure 6.9b).

Since the number of degrees of freedom of the perceptual metric is low (13 parameters), when compared with the full dimensionality of the training set, it can be expected that with a significantly smaller training set size, the correct values can still be obtained. The experimental results corroborate



(a) (b)

Figure 6.9: Results in terms of eval-order for Experiment 3 and 4, in plots (a) and (b) respectively.

this intuition: both with about 10 training patterns, or with about 0.5% of the total number of dissimilarities[16], the results were as good as with the original training set.

Experiment 5. The introduction of a strictly monotonic non-linear distortion function f in the cognitive distance computation was also tested. An arbitrary cubic polynomial was employed (plotted in figure 6.10).

$$f(x) = x^3 - 3x^2 + 3x \tag{6.117}$$

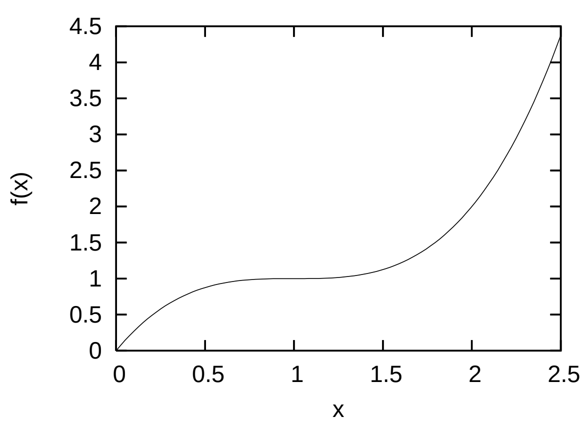


Figure 6.10: Plot of the cubic distortion function (6.117) employed in Experiment 5.

[16]Random sampling from the 4950 dissimilarities originated by the 100 patterns of the training set.

The results can be found in figure 6.11 for the weights, and in table 6.5 for the eval-order assessment. Comparing the former plot with the one in figure 6.7, it can be observed that the resulting weights are identical, as expected, by construction of the nonmetric MDS.

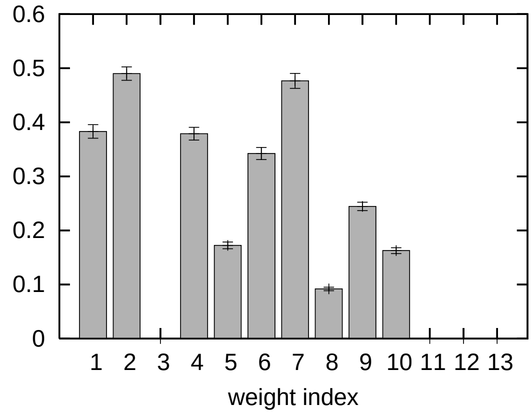


Figure 6.11: Weights obtained in Experiment 5 (using the distortion function (6.117)).

metric	mean	min	max
unweighted	11.47	1	97
weighted	1.004	1	2

Table 6.5: Results of the eval-order performance metric for Experiment 5.

Experiment 6. The relationship between the cost values and the eval-order is critical to the success of the approach. The algorithm seeks the reduction of the cost function (6.91), while the quality of the result is measured by the eval-order performance metric. For this synthetic world, the relationship between the cost and the eval-order during the gradient descent was examined. Figure 6.12 plots a sampling taken from 25 runs, by sampling randomly 1 out of 5 descent steps. This illustrates how, in this test-bed, smaller cost values lead systematically to better generalization in the test set. This kind of analysis can be useful to assess whether the method is appropriate for a given world, with respect to the generalization performance.

The second phase of the experimentation comprised the introduction of new components to the perceptual representation. To do so, the dimension of the cognitive images was made higher than the perceptual one, *i.e.,* $c > p$. Thus, the perceptual metric is performed with fewer components than the

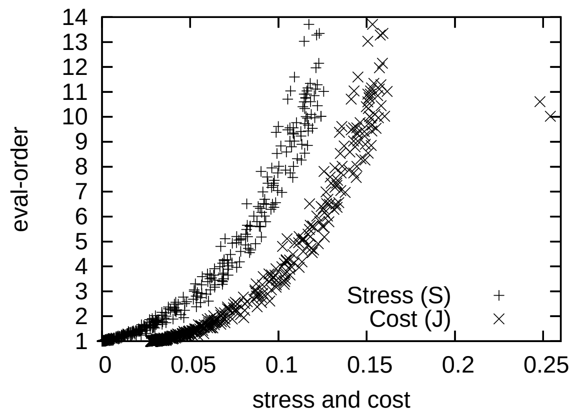


Figure 6.12: Samples of cost and mean eval-order pairs taken from 25 runs. The corresponding stress values are also shown.

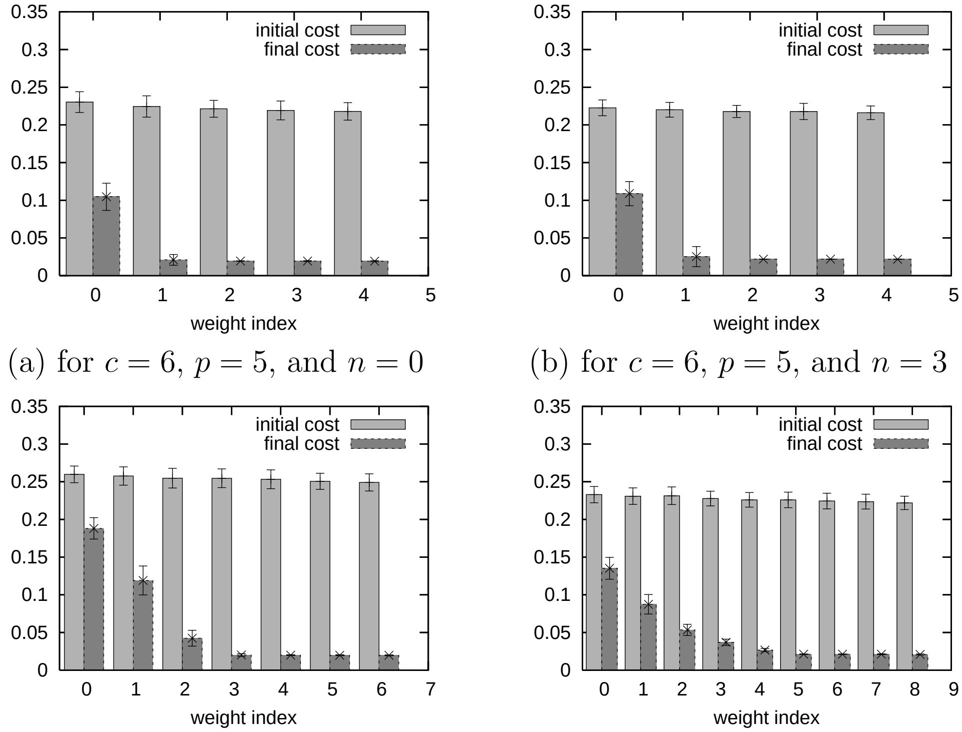
cognitive one. The first impact of this is that, without the introduction of new components in the perceptual representation, the final cost values are much higher than before, due to lack of fit (previous experiments reached final costs between 0.02 and 0.03).

In the following experiments, the descent parameters were set to $\eta_y = 0.04$ and $\alpha_y = 0.8$ (η_θ and α_θ remained at the same values as previous experiments).

Experiment 7. Figure 6.13 shows the obtained initial and final costs, after testing four different generated worlds. The algorithm was run for several numbers of new components for each one of the worlds. The plots display the mean and the standard deviation of the initial and final costs, after 100 runs performed in each world. The only difference among runs sharing the same world parameters is the initial values for the new dimension coordinates (initialized to random values, as explained above). The training set contained 20 patterns.

These plots corroborate the idea that, once the number of new components reaches $c - p$, the final cost stabilizes in values close to the ones found in previous experiments. This observation suggests a methodology for the estimation of how many new components are required for a given problem of unknown structure: to try successively higher amounts of new components, until the final cost value stabilizes.

Experiment 8. In this experiment, a single additional component to the cognitive representation was considered ($c = p + 1$). This setting (without additional perceptual components) results in a high final cost, as shown by the previous experiment, thus meaning a lack of fit. Then, a single new



(a) for $c = 6$, $p = 5$, and $n = 0$ (b) for $c = 6$, $p = 5$, and $n = 3$

(c) for $c = 8$, $p = 5$, and $n = 3$ (d) for $c = 10$, $p = 5$, and $n = 3$

Figure 6.13: Statistics of initial and final costs with respect to the number of new components, for various world parameters (indicated below each plot). Error bars denote the standard deviation of the cost values across the 100 runs.

perceptual component was added to the perceptual representation. It was observed that, not only a good fit was observed, but also that the values obtained for this new component were very close to the ones of the cognitive component missing in the perceptual image, apart from an affine transform.

This can be recognized by making the following observation: rewriting equation (6.97) for a single new dimension ($L = p$, the dimensionality of the perceptual representation)

$$d_{rs}^2 = \sum_{i=1}^{p} \theta_i^2 (x_{ri} - x_{si})^2 + (y_r - y_s)^2 \tag{6.118}$$

and considering that, if $d_{rs} = d_c(i_c^r, i_c^s)$ is a solution[17] (assuming f to be the

[17] It is here assumed that the metric parameters make both metrics numerically equal.

identity), then, one can write

$$\begin{aligned}\sum_{i=1}^{p} \theta_i^2 (x_{ri} - x_{si})^2 + (y_r - y_s)^2 &= (\mathbf{x}_r - \mathbf{x}_s)^T \mathbf{W}^2 (\mathbf{x}_r - \mathbf{x}_s) \\ &= \sum_{i=1}^{c} w_i^2 (x_{ri} - x_{si})^2 \qquad (6.119) \\ &= \sum_{i=1}^{p} w_i^2 (x_{ri} - x_{si})^2 + w_c^2 (x_{rc} - x_{sc})^2\end{aligned}$$

because $c = p + 1$. Letting $\theta_i = w_i$ for $i = 1, \ldots, p$, this equation is satisfied by any pair of points $\mathbf{x}_r$ and $\mathbf{x}_s$ if $y_r = w_c x_{rc}$. Thus, in this case, the new component is able to recover the values of original one, which is missing in the perceptual representation.

This reconstruction power was evaluated by measuring the signal-to-noise ratio (SNR) between the missing component and the recovered one, after normalizing them to zero mean and unit variance. The SNR was determined by dividing the energy of the reconstructed signal y_r (with respect to r) by the energy of the error $y_r - x_{rc}$, and then expressing it in decibel (dB) units

$$\mathrm{SNR} = 10 \log_{10} \frac{\sum_r y_r^2}{\sum_r (y_r - x_{rc})^2} \qquad (6.120)$$

Figure 6.14 shows the cognitive weights of the world, together with the weights obtained after running the algorithm. The world parameters were $c = 6$, $p = 5$, and $n = 3$. The trials consisted of 100 trials, with training sets of 20 patterns each. Note that the 6th component in (a) is the one that is hidden from the perceptual representation, and thus there is no corresponding weight in (b). The SNR results, then, measure how well the additional dimension reconstructs the values of this 6th component. The 6th to 8th weights in (b) correspond to the three noise components (and thus take negligible values in the end).

The SNR values (in dB) were collected in a histogram displayed in figure 6.15a. The shape of this histogram indicates the presence of a few outlier runs. To get rid of the outliers, the criterion SNR $>$ 20dB was used; figure 6.15b shows the histogram for the SNR after excluding these outliers (5%). Table 6.6 summarizes the statistics of the SNR, with and without outliers considered. Values of SNR between 40 and 50dB mean errors of about 0.3% to 1%, which are reasonably small.

Experiment 9. Regression methods can be employed to construct new features, based on the values found by the algorithm. As a proof of concept,

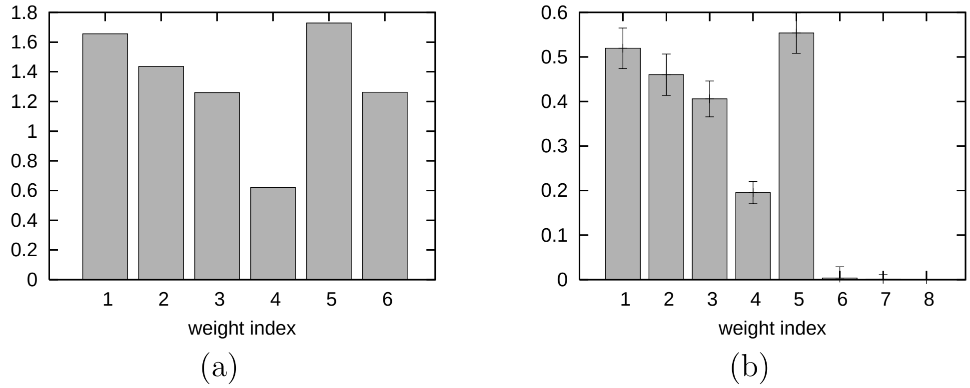


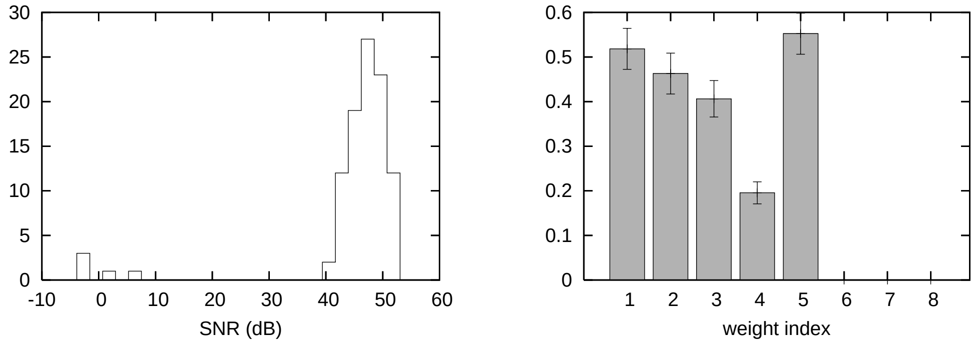
Figure 6.14: Weights obtained by Experiment 8: (a) world weights, (b) mean and standard deviation values of the obtained weights after 100 runs. See the main text for the explanation regarding the weights labeled 6–8.

outliers	mean	stdev	min	max
with	44.87	10.92	-3.855	53.05
without	47.27	2.920	40.46	53.05

Table 6.6: Results of Experiment 8 regarding the eval-order performance metric.

a set of components from the cognitive image was hidden from the perceptual representation, as in the previous experiment, and the algorithm was run for various amounts of additional components. After each run, a linear regression was performed in order to obtain a linear model mapping cognitive images to the new perceptual components. The perceptual images from the test set were augmented with the prediction coming from the regression model. Then, an eval-order assessment using the regression model over the test set was performed, as usual. Since the missing components belong to the cognitive images, the task is trivial: the regression model just needs to pick up the missing components from the cognitive image. The experimental parameters were: $c = 8$, $p = 5$, $n = 3$, 10 runs, 50 training patterns, 100 test patterns, and the number of new components ranged from 0 to 6. The results can be found in figure 6.16. The stabilization of the final cost values, for at least 3 new dimensions, can be verified in the plot (a) of this figure. Correspondingly, the eval-order drops to low levels, as it can be seen on plot (b) of the same figure.

Experiment 10. This experiment explored the idea advanced in section 6.4.3 regarding the possibility of overcoming the limitations of a given perceptual metric. Recall that the cognitive distances are invariant to a rota-



(a) (b)

Figure 6.15: Results of Experiment 8: (a) Histogram of the SNR reconstruction values for the 100 runs, and (b) weights after leaving out the outliers (5% of outliers, for SNR $\geq$ 20dB).

tion *after* applying the weighting, but not the converse. So, the latter case is beyond the degrees of freedom of the perceptual metric under use (6.93). To experiment whether the regression method over the perceptual images could overcome this limitation, this experiment used the same form of synthetic worlds, but with the cognitive images generated by

$$i_c = \mathbf{WRx} \tag{6.121}$$

instead of (6.112), where R is a random unitary matrix (a world parameter, and therefore shared by all points). Figure 6.17 shows the obtained results for several amounts of new components, both in terms of initial and final costs, and in terms of eval-order assessment.

The world parameters were set to $c = p = 5$, $n = 3$, while the results were collected after 10 runs with different worlds (and therefore with different **R** matrices), for each number of new components. The training and test sets contained 50 and 100 points respectively (and therefore, the maximum eval-order value was 50). Preliminary results had shown a tendency for the gradient descent to be stuck in local minima. So, in each run the gradient descent was performed 10 times, where only the results from the descent with minimum final cost were collected. These results are presented in figure 6.17.

The obtained results show a good capacity to overcome the perceptual metric limitation. This limitation is visible when no new components are added: although the average is relatively low (about 4), the maximum eval-order is high (48 with weights), which is about the same value as without weights (49). With as few as 3 new dimensions, the eval-order drops to values close to one.

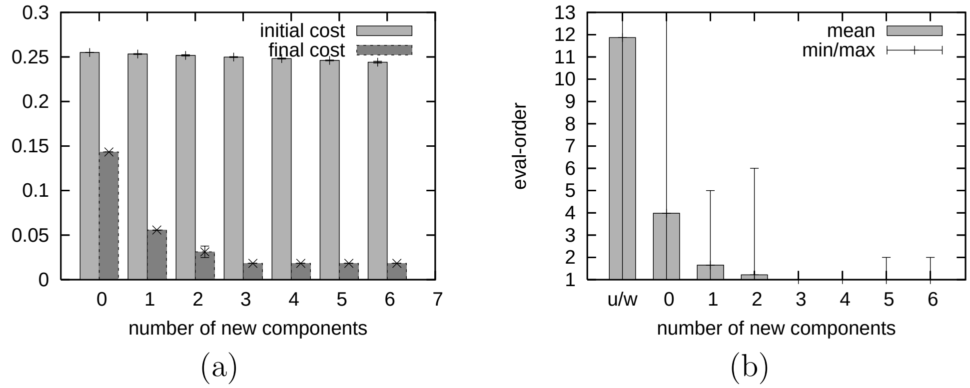
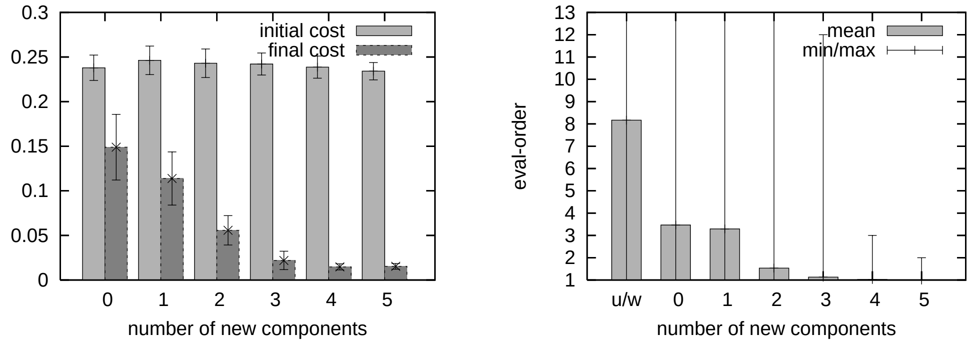


Figure 6.16: Results for Experiment 9: plot (a) displays the initial and final costs for various amounts of new components, while plot (b) shows the statistics for the eval-order, over all 10 runs. Error bars in (b) denote minimum and maximum eval-order values, over all test set points and runs. The bar labeled "u/w" denotes the eval-order results for the unweighted perceptual metric. The minimum eval-order for all bars (in (b)) is 1, and the maximum eval-order values for the first two bars were beyond the plot vertical range (42 and 17, respectively).

6.4.6 Concluding remarks

The experimental results presented above provide an illustration of the potentialities of the proposed method. The results were evaluated using the eval-order performance metric, designed to provide an assessment of the indexing efficiency. In the experiments it was observed that the algorithm, seeking the reduction of the cost function, led to good eval-order performance levels. This cost function is the sum of the stress, measuring the structural relationship between the cognitive and perceptual metrics, and a solution cost, aiming at the eradication of non-relevant perceptual components. Furthermore, the methodology for the improvement of the perceptual representation showed good results: the introduction of new components yielded a stress reduction. The way to derive these component values for new stimuli remained, however, an open issue.

At first glance, it might seem dubious to perform these experiments with cognitive and perceptual image spaces of approximately the same dimension, since the former was defined in chapter 4 as having a much higher dimensionality than the latter. However, note that the representation of cognitive images as vectors in $\mathbb{R}^c$ is purely instrumental for the generation of the training sets. Since these vectors are randomly distributed in $\mathbb{R}^c$, there is no way to find a perceptual representation with fewer dimensions that faithfully rep-



(a) (b)

Figure 6.17: Results of Experiment 10. See caption of figure 6.16 for a description of the plots. The minimum eval-order in (b) for all bars is 1, and the maximum eval-order values for the first four bars were: 49, 48, 49, and 14.

resents the $\mathbb{R}^c$ structure. Nowhere in the algorithm are these vectors actually used. Instead, distances are computed in this space, and used as cognitive distances. This is the generation law of the world. One can consider, for instance, a complex cognitive representation (e.g., vision images) from which a slow cognitive process extracts a lower dimension space spawning the world structure. At the same time, the perceptual representation is based on a fast feature extraction procedure.

During experimentation it was observed that one of the major limitations of the method was that, unless a relatively good fit is found, the weights assign relevance to irrelevant components. For instance, the weights corresponding to the noise components only converged to zero if a good fit was found[18]. Therefore, good results are to be expected only for worlds where the perceptual metric is able to approximate the cognitive one over the *full* range of distances, *i.e.,* not only the closer pairs, but also the more distant ones. This limitation is particularly serious for domains where one is concerned with good cognitive matches (small distances) only.

[18]Recall that one of the criteria utilized to leave out outlier runs was based on the non-nullity of those weights.

Chapter 7

Concluding discussion

7.1 Introduction

This thesis introduces the exploration of the possibilities arising from a specific emotion-based agent model. Thus, the work reported here should be taken as a snapshot of the state of this exploration. The starting point for this agent model consisted in a biologically inspired view of the role of emotions in decision-making. Contrary to common belief, and according to Damásio, the emotional mechanisms in the brain are crucial for appropriate decision making, even in domains of apparent rational dominance. Since this hypothesis concerns mechanisms, rather than specific emotions, the agent model is focused on mechanisms, and not on specific emotions. This option was taken from the initial formulation of the model, and has been followed throughout the developed research. This is why there are scarce references to emotions, apart from the initial chapters about the background and the review of the state-of-the-art. The conceptual architecture described in chapter 4 makes no explicit references to emotions, and the research that followed did not aim at the simulation of emotions. However, once can *a posteriori* discuss, for instance, whether it makes sense to relate emotions with the desirability vector.

7.2 Results

Two different methodologies where adopted in this work. The first one sought to experiment with complete agents, interacting with testbed environments. This research, presented in chapter 5, led to the development of two testbeds — an inverted pendulum system, and a symbolic POMDP — for which solutions concerning the anticipation of future outcomes were explored.

Constructing complete agents requires taking care of the various aspects involved, such as perception, action selection, memory management, and so on. The resulting behavior is therefore a non-trivial composition of all of these aspects. Thus, it becomes hard to assign credit to individual aspects of the agent architecture. Moreover, whenever the agent is able to cope with the environment, the question of whether the same methodology scales up to other domains remains unanswered. It is not trivial to separate the domain independent from the domain dependent parts of the architecture.

These difficulties suggested that a different methodology could provide more fruitful results. Thus, a back-to-basics approach was taken next, presented in chapter 6, where individual aspects of the architecture were selected and explored in depth. Among the mechanisms discussed in chapter 4, the indexing one was chosen, following the principle of the path-least-traveled. The marking mechanism, for instance, shows many similarities with the class of reinforcement learning techniques, which have already received extensive research attention [103] over recent years. Moreover, the indexing one seemed more interesting to explore, since it relates more closely to the differences between the two representation schemata of the model.

The indexing mechanism was then approached from two perspectives. The first one modeled the occurrence of stimuli as a probabilistic event, in order to draw conclusions with respect to the probabilities of matching. This approach encompassed analyses of efficiency and of the probabilities of incurring errors. A second approach assumed that the two representation schemata form distinct metric spaces. The consequences of the different resolutions of the cognitive and perceptual representations were explored, under an assumption relating the two metric spaces. An illustrative example — the handwritten digits classification domain — accompanies the theoretical considerations, assessing the benefits of the double-representation paradigm in that specific domain. It was interesting to verify that, even with features yielding poor information about the digits class, significant efficiency gains were obtained.

Following the metric space formulation, the problem of developing appropriate perceptual representations aiming at an efficient indexing of cognitive images was addressed. To tackle this problem, two complementary strategies were explored: the problem of identifying and adapting an appropriate perceptual metric, and the problem of improving the perceptual representation itself. Chapter 6 reports progress made on these two fronts. Experimental results were obtained using synthetic data sets consisting of random points in finite dimensional spaces. The results have corroborated the formulated theoretical hypotheses. However, it was not possible to obtain, so far, similarly good results with data sets from real domains. Possible explanations for this shortcoming were discussed at the end of the previous chapter.

7.3 Future research directions

From the two methodologies discussed above, the second one seems the most promising one. Providing the conceptual model with a solid theoretical basis is a fundamental endeavor. Nonetheless, one must always keep in mind that the *raison d'être* of the model is its integration into an fully functional agent architecture. Therefore, the first methodology should be resumed as soon as the theoretical framework of the agent model reaches a sufficient degree of maturity.

Regarding the perceptual metric and representation adaptation problem (section 6.4), it is crucial to orient experimental work towards domains beyond those of synthetic data points uniformly distributed in space. To do that, it is important to relax the need for a good data fit. It was observed experimentally that when the data does not fit the adapted metric (and/or representation), the results are not satisfactory. This was quite visible in the experimental results: dimensions not represented in the perceptual representation led the algorithm, for instance, to ascribe relevance to noise features (as non-zero metric weights). One way to relax the optimization is to allow for ambiguity in the perceptual matching, leaving the finer discrimination to the cognitive one. In this way, the problem of finding a good cognitive match is divided in two phases: the perceptual match, taking into account a perceptual representation of low dimensionality, yielding an intermediate result, on which the final match is performed by the cognitive layer.

Following the same methodological path, the remaining aspects of the conceptual model deserve a theoretical treatment. Associating cognitive and perceptual images with desirability vectors, for instance, is an essential mechanism for providing meaning to stimuli, grounding it in terms of desirability for the agent. This research crosses necessarily other areas, such as utility theory, as well as decision theory, dynamic programming, planning, among others.

Although not in a completely overt fashion, this thesis assumes that the cognitive and the perceptual images are internal representations of stimuli, *i.e.,* stimuli originating from the agent's sensors. It could be interesting to explore the idea of extending the proposed model beyond stimulus representation. For instance, one could consider a cognitive (complex) and a perceptual (simple) representation of actions. One could also consider more complex representations involving not only actions, but also interaction, as for instance in the case of manipulation tasks involving mechanical hands with visual feedback. Neuroscientific evidence has suggested the occurrence of shared representations across sensory and motor spaces. The discovery of mirror neurons in monkeys supports this hypothesis: it was found that certain neurons were active both when the subject were performing a task, and when the subject was observing another monkey performing the same task [86]. This discovery suggests that the brain uses such shared representations to perform imitation.

One key issue of the agent model is efficiency. Real-time systems require agents that perform adequately, taking into account that they cannot take an unreasonable amount of time to respond to a solicitation from the environment. This implies a trade-off between the optimality of a solution, and the ability to respond to the environment in due time, thus providing a satisfying, but useful, solution. Herbert Simon discussed this trade-off within the larger context of optimizing *vs.* satisfycing [171]. Anytime algorithms [218] address this issue by providing the best solution possible for any time constraint, where ideally, the longer the available time, the better the quality of the solution. Note that the indexing mechanism can be formulated as a anytime algorithm, once one considers that the cognitive match is performed in the order determined by the preceding perceptual match, and that the decision process could use the best cognitive match found within the allotted time.

The quest for real-life problems originates from the biological motivation behind the use of emotions: the biologically inspired hypothesis that one of the main contributions of emotional mechanisms resides in the capability shown by humans of dealing with complex, dynamic, and unpredictable environments. Even with sophisticated cognitive competences, when there is an impairment of the emotional mechanisms, subjects are no longer capable of adequately dealing with simple day-to-day decisions. Their cognitive capabilities are intact (as asserted by IQ tests), but the same cannot be said about their capability to deal with their daily lives. Therefore, and transporting this point to the domain of this work, it should be kept in mind that the ultimate goal is a sustained capability of dealing appropriately with real-life complex problems. It is then crucial to orient research towards this kind of problems.

The challenges posed by real-life problems should. however, not be confused with the sheer complexity of, for instance, combinatorial problems. Current state-of-the-art algorithms are capable of dealing efficiently with combinatorial problems of considerable complexity. However, there are many issues that prevent them to address real-life domains in a straightforward fashion. One of the key issues is the structure of the world that can be exploited, yielding efficient ways of coping with it. Exploring the world structure is closely related with the problem of representing it internally. Appropriate representations can have a huge impact on the computational costs to find a solution.

Bibliography

[1] John R. Anderson. *Cognitive Psychology and its Implications*. Worth Publishers, fifth edition edition, 2000.

[2] Virgil Andronache and Matthias Scheutz. APOC – a framework for complex agents. In Hod Lipson, Erik K. Antonsson, and John R. Koza, editors, *Computational Synthesis: From Basic Building Blocks to High Level Functionality, Papers from 2003 AAAI Spring Symposium, Technical Report SS-03-02*, pages 18–25, Menlo Park, California, 2003. The AAAI Press.

[3] Michael Arbib and Jean-Marc Fellous. Emotions: from brain to robot. *Trends in Cognitive Science*, 8(12):554–561, December 2004.

[4] Ronald C. Arkin. *Behavior-Based Robotics*. MIT Press, 1989.

[5] Ronald C. Arkin and Tucker Balch. AuRA: principles and practice in review. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2–3):175–189, April 1997.

[6] Zippora Arzi-Gonczarowski. Wisely non-rational — a categorical view of emotional cognitive artificial perceptions. In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-98-03*, pages 7–12, Menlo Park, California, 1998. The AAAI Press.

[7] Zippora Arzi-Gonczarowski. A categorization of autonomous action tendencies: The mathematics of emotions. In Robert Trappl, editor, *Cybernetics and Systems 2000*, volume 2, pages 683–688, Wien, April 2000.

[8] Zippora Arzi-Gonczarowski. Self, empathy, manipulativity: Mathematical connections between higher order perceptions, emotions, and social cognition. In Dolores Cañamero, editor, *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Papers from the AAAI*

Fall Symposium, Technical Report FS-01-02, pages 9–14, Menlo Park, California, 2001. The AAAI Press.

[9] Zippora Arzi-Gonczarowski and Daniel Lehmann. From environments to representations — a mathematical theory of artificial perceptions. *Artificial Intelligence*, 102(2):187–247, 1998.

[10] Zippora Arzi-Gonczarowski and Daniel Lehmann. Introducing the mathematical category of artificial perceptions. *Annals of Mathematics and Artificial Intelligence*, 23(3,4):276–298, November 1998.

[11] Michel Aubé. A commitment theory of emotions. In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-98-03*, pages 13–18, Menlo Park, California, 1998. The AAAI Press.

[12] Michel Aubé. Designing adaptive cooperating animats will require designing emotions: Expanding upon Toda's urge theory. In Dolores Cañamero, Chisato Numaoka, and Paolo Petta, editors, *Workshop: Grounding Emotions in Adaptive Systems*, pages 7–12. SAB'98: From Animals to Animats, August 1998.

[13] Ruth Aylett. Emotion in behavioural architectures. In Juan Velásquez, editor, *Workshop on Emotion-Based Agent Architectures (EBAA'99)*, pages 10–12, May 1999.

[14] Ruth Aylett and Carlos Delgado. Emotion and agent interaction. In Dolores Cañamero, editor, *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-01-02*, pages 15–16, Menlo Park, California, 2001. The AAAI Press.

[15] Bernard J. Baars. The consciousness access hypothesis: origins and recent evidence. *Trends in Cognitive Science*, 6(1):47–52, January 2002.

[16] Christian Balkenius and Jan Morén. A computational model of emotional conditioning in the brain. In Dolores Cañamero, Chisato Numaoka, and Paolo Petta, editors, *Workshop: Grounding Emotions in Adaptive Systems*, pages 13–18. SAB'98: From Animals to Animats, August 1998.

[17] Christian Balkenius and Jan Morén. Emotional learning: A computational model of the amygdala. *Cybernetics and Systems: An International Journal*, 32:611–636, 2001.

[18] Aharon Bar-Hillel, Tomer Hertz, Noam Schental, and Daphna Weinshall. Learning a Mahalanobis metric from equivalence constraints. *Journal of Machine Learning Research*, 6:937–965, 2005.

[19] Andrew G. Barto, Richard S. Sutton, and Charles W. Anderson. Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on System, Man, and Cybernetics*, SMC-13(5):834–846, September/October 1983.

[20] Joseph Bates. The role of emotion in believable agents. *Communications of the ACM, Special Issue on Agents*, 37(7):122–125, July 1994. Technical Report CMU-CS-94-136.

[21] Joseph Bates, A. Bryan Loyall, and W. Scott Reilly. An architecture for action, emotion, and social behavior. In *Proceedings of the Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World*, Decentralized AI Series. Elsevier/North Holland, July 1992.

[22] Luc Beaudoin. *Goal Processing in Autonomous Agents*. PhD thesis, School of Computer Science, University of Birmingham, August 1994.

[23] Antoine Bechara, Antonio R. Damásio, Hanna Damásio, and Steven W. Anderson. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50:7–15, 1994.

[24] Antoine Bechara, Hanna Damásio, Daniel Tranel, and Antonio R. Damásio. Deciding advantageously before knowing the advantageous strategy. *Science*, 275:1293–1295, February 1997.

[25] Antoine Bechara, Hanna Damásio, Daniel Tranel, and Antonio R. Damásio. The Iowa gambling task and the somatic marker hypothesis: some questions and answers. *Trends in Cognitive Science*, 9(4):159–162, April 2005.

[26] R. Peter Bonasso, R. James Firby, Erann Gat, David Kortenkamp, and Davis P. Miller. Experiences with an architecture for intelligent reactive agents. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2–3):237–256, April 1997.

[27] Luís Botelho and Helder Coelho. Machinery for artificial emotions. *Cybernetics and Systems*, 32(5):465–506, 2001.

[28] Luís Botelho, Pedro Ramos, and Pedro Figueiredo. Emotion eliciting in Salt & Pepper. In Robert Trappl, editor, *Cybernetics and Systems 2004*, pages 657–662, Wien, 2004.

[29] Luís Miguel Botelho and Helder Coelho. Adaptive agents: emotion learning. In Dolores Cañamero, Chisato Numaoka, and Paolo Petta, editors, *Workshop: Grounding Emotions in Adaptive Systems*, pages 19–24. SAB'98: From Animals to Animats, August 1998.

[30] Cynthia Breazeal. A motivational system for regulating human-robot interaction. In *Proceedings of AAAI-98*, pages 54–61. AAAI, 1998.

[31] Cynthia Breazeal. Robot in society: Friend or appliance? In Juan Velásquez, editor, *Workshop on Emotion-Based Agent Architectures (EBAA'99)*, pages 18–26, May 1999.

[32] Cynthia Breazeal. Regulation and entrainment in human-robot interaction. *The International Journal of Robotics Research*, 21(10):883–902, October 2002.

[33] James F. Brennan. *History and Systems of Psychology.* Prentice Hall, 6th edition edition, 2003.

[34] Rodney A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2(1):14–23, March 1989.

[35] Rodney A. Brooks. Intelligence without reason. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 569–595, Los Altos/Palo Alto/San Francisco, 1991. IJCAI, Morgan Kaufmann.

[36] Alastair Burt. Emotionally intelligent agents: The outline of a resource-oriented approach. In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-98-03*, pages 43–48, Menlo Park, California, 1998. The AAAI Press.

[37] Alastair Burt. Research into emotionally intelligent agents. In Dolores Cañamero, Chisato Numaoka, and Paolo Petta, editors, *Workshop: Grounding Emotions in Adaptive Systems*, page 79. SAB'98: From Animals to Animats, August 1998.

[38] Martin V. Butz, David E. Goldberg, and Wolfgang Stolzmann. The anticipatory classifier system and genetic generalization. *Natural computing*, 1(4):427–467, 2002.

[39] Murray Campbell, A. Joseph Hoane Jr., and Feng Hsiung Hsu. Deep blue. *Artificial Intelligence*, 134:57–83, 2002.

[40] Dolores Cañamero. Modeling motivations and emotions as a basis for intelligent behavior. In W. Lewis Johnson, editor, *Proceedings of the First International Conference on Autonomous Agents*, pages 148–155. ACM Press, 1997.

[41] Anthony R. Cassandra, Leslie Pack Kaelbling, and Michael L. Littman. Acting optimally in partially observable stochastic domains. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, pages 1023–1028. AAAI, AAAI Press, 1994.

[42] Christos G Cassandras, Stephane Safortune, and Stephane Lafortune. *Introduction to Discrete Event Systems*. Kluwer Academic Publishers, 1999.

[43] Patricia S. Churchland and Terrence J. Sejnowski. *The Computational Brain*. The MIT Press, 1992.

[44] Patricia Smith Churchland. *Neurophilosophy*. The MIT Press, 1989.

[45] Cristina Conati. Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence*, 16(7/8):555–575, August 2002.

[46] Cristina Conati and Xiaoming Zhou. A probabilistic framework for recognizing and affecting emotions. In Eva Hudlicka and Dolores Cañamero, editors, *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, Papers from 2004 AAAI Spring Symposium, Technical Report SS-04-02*, pages 13–16, Menlo Park, California, 2004. The AAAI Press.

[47] Samuel D. Conte and Carl de Boor. *Elementary Numerical Analysis*. McGraw-Hill, third edition edition, 1987.

[48] Silvia Coradeschi and Alessandro Saffiotti. Perceptual anchoring of symbols for action. In *Proceedings of the 17th International Conference on Artificial Intelligence (IJCAI-01)*, pages 407–412, Seattle, WA, 2001.

[49] Trevos F. Cox and Michael A. A. Cox. *Multidimensional Scaling*. Chapman & Hall, London, UK, 1994.

[50] Tim Dalgleish and Mick Power, editors. *Handbook of Cognition and Emotion*. Wiley, 1999.

[51] Bruno D. Damas and Luís M. Custódio. Emotion-based decision and learning using associative memory and statistical estimation. *Informatica*, 27:147–157, 2003.

[52] Bruno Duarte Damas. Arquitectura de aprendizagem e decisão de um agente inspirado em emoções. Master's thesis, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal, January 2006.

[53] Antonio R. Damásio. The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, 1:123–132, 1989.

[54] Antonio R. Damásio. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33(1–2):25–62, 1989.

[55] Antonio R. Damásio. *Descartes' Error: Emotion, Reason and the Human Brain*. Picador, 1994.

[56] Antonio R. Damásio. *The Feeling of What Happens: Body and emotion in the making of consciousness*. Harcourt Brace, 1999.

[57] Antonio R. Damásio. *Looking for Spinoza: Joy, sorrow, and the feeling brain*. Harvest Books, 2003.

[58] Antonio R. Damásio and Hanna Damásio. *Large-Scale Neuronal Theories of the Brain*, chapter Cortical Systems for Retrieval of Concrete Knowledge: The Convergence Zone Framework, pages 61–74. The MIT Press, 1994.

[59] Antonio R. Damásio, Daniel Tranel, and Hanna C. Damásio. *Frontal Lobe Function and Dysfunction*, chapter Somatic Markers and the Guidance of Behavior: Theory and Preliminary Testing, pages 217–229. Oxford University Press, NY, 1991.

[60] Hanna Damásio, T. J. Grabowski, Daniel Tranel, R. D. Hichwa, and Antonio Damásio. A neural basis for lexical retrieval. *Nature*, 380:499–505, 1996.

[61] B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 2nd edition edition, 2002.

[62] Peter Davies and Tony Coxon, editors. *Key Texts in Multidimensional Scaling*. Heinmann Educational Books, 1982.

[63] Li curgo B. de Almeida, Bruno C. da Silva, and Ana L. C. Bazzan. Towards a physiological model of emotions: first steps. In Eva Hudlicka and Dolores Cañamero, editors, *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, Papers from 2004 AAAI Spring Symposium, Technical Report SS-04-02*, pages 1–4, Menlo Park, California, 2004. The AAAI Press.

[64] Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391–407, 1990.

[65] Daniel C. Dennett. *Consciousness Explained*. Back Bay Books, 1991.

[66] Carlotta Domeniconi, Jing Peng, and Dimitrios Gunopulos. An adaptive metric machine for pattern classification. *Advances In Neural Information Processing Systems*, 13:458–464, 2001.

[67] Dietrich Dörner and Ulrike Starker. Should successful agents have emotions? the role of emotions in problem solving. In *Proceedings of the Sixth International Conference on Cognitive Modeling*, pages 344–345, Mahwah, NJ, 2004. Lawrence Earlbaum.

[68] Prashant Doshi and Piotr Gmytrasiewicz. Towards affect-based approximations to rational planning: A decision-theoretic perspective to emotions. In Eva Hudlicka and Dolores Cañamero, editors, *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, Papers from 2004 AAAI Spring Symposium, Technical Report SS-04-02*, pages 33–36, Menlo Park, California, 2004. The AAAI Press.

[69] Michael G. Dyer. Emotions and their computations: Three computer models. *Cognition and Emotion*, 1(3):323–347, 1987.

[70] Paul Ekman. An argument for basic emotions. *Cognition and Emotion*, 6(3/4):169–200, 1992.

[71] Paul Ekman. *Handbook of Cognition and Emotion*, chapter Basic Emotions, pages 45–60. Wiley, 1999.

[72] Clark Elliott. I picked up Catapia and other stories: a multimodal approach to expressivity for "emotionally intelligent" agents. In *Proceedings of the first international conference on Autonomous agents*, pages 451–457, Marina del Rey, California, 1997. ACM Press.

[73] Clark Elliott. Hunting for the holy grail with "emotionally intelligent" virtual actors. *ACM SIGART Bulletin*, 9(1):20–28, June 1998.

[74] Clark Davidson Elliott. *The Affective Reasoner: A process model of emotions in a multi-agent system*. PhD thesis, Northwestern University, Evanston, Illinois, June 1992.

[75] Jean-Marc Fellous. The neuromodulatory basis of emotion. *The Neuroscientist*, 5(5):283–294, 1999.

[76] Jean-Marc Fellous. From human emotions to robot emotions. In Eva Hudlicka and Dolores Cañamero, editors, *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, Papers from 2004 AAAI Spring Symposium, Technical Report SS-04-02*, pages 37–47, Menlo Park, California, 2004. The AAAI Press.

[77] Lesley K. Fellows and Martha J. Farah. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex*, 15:58–63, January 2005.

[78] Joseph P. Forgas. *Handbook of Cognition and Emotion*, chapter Network Theories and Beyond, pages 591–611. Wiley, 1999.

[79] Nico H. Frijda. *The Emotions*. Cambridge University Press, Editions de la Maison des Sciences de l'Homme, Paris, 1986.

[80] Nico H. Frijda. *Comparative Approaches to Cognitive Science*, chapter Emotions in Robots. MIT Press, 1995.

[81] Nico H. Frijda and Jaap Swagerman. Can computers feel? theory and design of an emotional system. *Cognition and Emotion*, 1(3):235–257, 1987.

[82] Sandra Clara Gadanho. *Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions*. PhD thesis, University of Edinburgh, 1999.

[83] Sandra Clara Gadanho. Emotional and cognitive adaptation in real environments. In Robert Trappl, editor, *Cybernetics and Systems 2000*, pages 762–767, Wien, 2002.

[84] Sandra Clara Gadanho and Luís Custódio. Learning behavior-selection in a multi-goal robot task. *Informatica*, 27:175–183, 2003.

[85] Sandra Clara Gadanho and John Hallam. Robot learning driven by emotions. *Adaptive Behavior*, 9(1):42–64, 2002.

[86] Vittorio Gallese, Luciano Fadiga, Leonardo Fogassi, and Giacomo Rizzolatti. Action recognition in the premotor cortex. *Brain*, 119:593–609, 1996.

[87] Michael S. Gazzaniga. *The Scientific American book of the brain*, chapter The Split Brain Revisited, pages 129–138. Lyons Press, 1999.

[88] Piotr J. Gmytrasiewicz and Christine L. Lisetti. Using decision theory to formalize emotions for multi-agent systems. In *Second ICMAS-2000 Workshop on Game Theoretic and Decision Theoretic Agents*, Boston, 2000.

[89] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley, 1988.

[90] Daniel Goleman. *Emotional Intelligence*. Bloomsbury, 1996.

[91] Ian H. Gotlib and Lynn Y. Abramson. *Handbook of Cognition and Emotion*, chapter Attributional Theories of Emotion, pages 613–636. Wiley, 1999.

[92] Jonathan Gratch. Why you should buy an emotional planner. In Juan Velásquez, editor, *Workshop on Emotion-Based Agent Architectures (EBAA'99)*, pages 53–60, May 1999.

[93] Jonathan Gratch. Émile: Marshalling passions in training and education. In *Proceedings of the 4th International Conference on Autonomous Agents*, Barcelona, Spain, June 2000.

[94] Jonathan Gratch and Stacy Marsella. Evaluating a general model of emotional appraisal and coping. In Eva Hudlicka and Dolores Cañamero, editors, *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, Papers from 2004 AAAI Spring Symposium, Technical Report SS-04-02*, pages 52–59, Menlo Park, California, 2004. The AAAI Press.

[95] Jennifer Healey, Rosalind Picard, and Frank Dabek. A new affect-perceiving interface and its application to personalized music selection. In *Proceedings of the 1998 Workshop on Perceptual User Interfaces*, San Francisco, CA, November 1998.

[96] Marcus J. Huber. JAM: a BDI-theoretic mobile agent architecture. In *Proceedings of the third annual conference on Autonomous Agents*, pages 236–243, Seattle WA USA, 1999.

[97] Eva Hudlicka. This time with feeling: Integrated model of trait and state effects on cognition and behavior. *Applied Artificial Intelligence*, 16(7/8):611–641, August 2002.

[98] Eva Hudlicka. Two sides of appraisal: Implementing appraisal and its consequences within a cognitive architecture. In Eva Hudlicka and Dolores Cañamero, editors, *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, Papers from 2004 AAAI Spring Symposium, Technical Report SS-04-02*, pages 70–76, Menlo Park, California, 2004. The AAAI Press.

[99] Eva Hudlicka. Modeling interaction between metacognition and emotion in cognitive architectures. In Mike Anderson and Tim Oates, editors, *Metacognition in Computation, Papers from 2005 AAAI Spring Symposium, Technical Report SS-05-04*, pages 55–61, Menlo Park, California, 2005. The AAAI Press.

[100] Aapo Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.

[101] Anil K. Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, January 2000.

[102] William James. What is an emotion? *Mind*, 9:188–205, 1884.

[103] Leslie Pack Kaelbling, Michael L. Littman, and Andrew W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[104] Ashish Kapoor, Selene Mota, and Rosalind W. Picard. Towards a learning companion that recognizes affect. In Dolores Cañamero, editor, *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-01-02*, pages 67–72, Menlo Park, California, 2001. The AAAI Press.

[105] M. Kirby and L. Sirovich. Application of the karhunen-loève procedure for the characterization of human faces. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 12(1):103–108, January 1990.

[106] Hiroaki Kitano, Minoru Asada, Yasuo Kuniyoshi, Itsuki Noda, and Eiichi Osawa. RoboCup: The robot world cup initiative. In *Proceedings of the First International Conference on Autonomous Agents (Agents'97)*, pages 340–347, New York, 1997. ACM Press.

[107] Teuvo Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.

[108] Joseph B. Kruskal. Multidimensional scaling by optimizing goodness-of-fit to a non-metric hypothesis. *Psychometrika*, 29:1–27, 1964. reprinted in [62].

[109] Joseph B. Kruskal. Non-metric multidimensional scaling: a numeric method. *Psychometrika*, 29:115–129, 1964. reprinted in [62].

[110] S. Mac Lane. *Categories for the Working Mathematician*. Springer-Verlag, 1971.

[111] Richard S. Lazarus. *Handbook of Cognition and Emotion*, chapter The cognition-emotion debate: a bit of history, pages 3–19. Wiley, 1999.

[112] Joseph LeDoux. *The Emotional Brain*. Simon & Schuster, 1996.

[113] Joseph LeDoux. Emotion circuits in the brain. *Annual Review of Neuroscience*, 23:155–184, March 2000.

[114] Joseph LeDoux. *Synaptic Self*. Penguin, 2002.

[115] Daniel D. Lee and H. Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, October 1999.

[116] Marc D. Lewis and Isabela Granic. *Handbook of Cognition and Emotion*, chapter Self-organization of Cognition-Emotion Interactions, pages 683–701. Wiley, 1999.

[117] Christine L. Lisetti. Emotion synthesis: Some research directions. In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-98-03*, pages 109–115, Menlo Park, California, 1998. The AAAI Press.

[118] George F. Loewenstein, Elke U. Weber, Christopher K. Hsee, and Ned Welch. Risk as feelings. *Psychological Bulletin*, 127(2):267–286, 2001.

[119] William Lyons. *Handbook of Cognition and Emotion*, chapter The philosophy of cognition and emotion, pages 21–44. Wiley, 1999.

[120] Márcia Maçãs, Paulo Couto, Rodrigo Ventura, Luis Custódio, and Carlos Pinto-Ferreira. Sociedades de agentes e implementação de inteligência sensória-motriz. Technical Report Internal Report RT-704-99, ISR, October 1999.

[121] Márcia Maçãs and Luís Custódio. Multiple emotion-based agents using an extension of DARE architecture. *Informatica*, 27:185–195, 2003.

[122] Márcia Maçãs, Rodrigo Ventura, Luis Custódio, and Carlos Pinto-Ferreira. Experiments with an emotion-based agent using the DARE architecture. In *Proceedings of the Symposium on Emotion, Cognition, and Affective Computing (AISB'01 Convention)*, UK, March 2001.

[123] Pattie Maes, Trevor Darrell, Bruce Blumberg, and Alex Pentland. The ALIVE system: wireless, full-body interaction with autonomous agents. *Multimedia Systems*, 5(2):105–112, March 1997.

[124] Tiago V. Maia and James L. McClelland. A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the United States of America*, 101(45):16075–16080, November 2004.

[125] Tiago V. Maia and James L. McClelland. The somatic marker hypothesis: still many questions but no answers. *Trends in Cognitive Science*, 9(4):162–164, April 2005.

[126] Stacy Marsella and Jonathan Gratch. Modeling coping behavior in virtual humans: Don't worry, be happy. In *Proceedings of the 2nd International Conference on Autonomous Agents and Multiagent Systems*, Melbourne, Australia, July 2003.

[127] Carlos Martinho and Ana Paiva. Developing pathematic agents. In Juan Velásquez, editor, *Workshop on Emotion-Based Agent Architectures (EBAA'99)*, pages 79–83, May 1999.

[128] Carlos Martinho, Ana Paiva, and Mário Gomes. Emotions for a motion: Rapid development of believable pathematic agents in intelligent virtual environments. *Applied Artificial Intelligence*, 14(1):33–68, 2000.

[129] John McCarthy. Ascribing mental qualities to machines. In Vladimir Lifschitz, editor, *Formalizing Common Sense: Papers by John McCarthy*, pages 93–118. Ablex Publishing Corporation, Norwood, New Jersey, 1990.

[130] Anna-Maria R. McGowan, David E. Cox, Barry S. Lazos, Martin R. Waszak, David L. Raney, Emilie J. Siochi, and S. Paul Pao. Biologically inspired technologies in NASA's morphing project. *Proceedings of SPIE*, 5051:1–13, July 2003.

[131] Marc Métivier and Claude Lattaud. *Learning Classifier Systems: 5th International Workshop, IWLCS 2002*, volume 2661 of *Lecture Notes in Computer Science*, chapter Anticipatory Classifier System Using Behavioral Sequences in Non-Markov Environments, pages 143–162. Springer, 2003. Revised Papers.

[132] Giorgio Metta, Giulio Sandini, David Vernon, Darwin Caldwell, Nikolaos Tsagarakis, Ricardo Beira, José Santos-Victor, Auke Ijspeert, Ludovic Righetti, Giovanni Cappiello, Giovanni Stellin, and Francesco Becchi. The RobotCub project: an open framework for research in embodied cognition. In *Proceedings of the 2005 IEEE-RAS International Conference on Humanoid Robots*, 2005.

[133] Marvin Minsky. *The Society of Mind*. Touchstone, 1988.

[134] Luís Morgado and Graça Gaspar. A generic agent model allowing a continuous characterization of emotion. In Robert Trappl, editor, *Cybernetics and Systems 2004*, pages 651–656, Wien, 2004.

[135] Luís Morgado and Graça Gaspar. Emotion based adaptive reasoning for resource bounded agents. In *Proceedings of AAMAS'05*, pages 921–928, Utrecht, Netherlands, July 2005.

[136] Lilia Moshkina and Ronald C. Arkin. On TAMEing robots. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 3949–3959. IEEE, October 2003.

[137] Lilia Moshkina and Ronald C. Arkin. Human perspective on affective robotic behavior. In *submitted to IROS2005*. IEEE, 2005.

[138] Ulric Neisser, Gwyneth Boodoo, Thomas J. Bouchard, A. Wade Boykin, Nathan Brody, Stephen J. Ceci, Diane F. Halpern, John C.

Loehlin, Robert Perloff, Robert J. Sternberg, and Susana Urbina. Intelligence: Knowns and unknowns. *American Psychologist*, 51(2):77–101, February 1996.

[139] Allen Newell and Herbert A. Simon. Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19(3):113–126, 1976.

[140] Chisato Numaoka. Personality development through interactions in virtual worlds. In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-98-03*, pages 135–140, Menlo Park, California, 1998. The AAAI Press.

[141] Katsuhiko Ogata. *Modern Control Engineering*. Prentice-Hall, second edition edition, 1990.

[142] A. Ortony, G. L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge, UK, 1988.

[143] Ivan P. Pavlov. Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. Classics in the History of Psychology web site, 1927. (Translated by G. V. Anrep).

[144] Paolo Petta. Principled generation of expressive behavior in an interactive exhibit. In Juan Velásquez, editor, *Workshop on Emotion-Based Agent Architectures (EBAA'99)*, pages 94–98, May 1999.

[145] Paolo Petta. Introducing emotions into the computational study of social norms: A first evaluation. *Journal of Artificial Societies and Social Simulation*, 4(1), 2001.

[146] Rolf Pfeifer. *Cognitive perspectives of emotion and motivation, Proc. NATO Advanced Research Workshop*, volume 44, chapter Artificial Intelligence Models of Emotion, pages 287–320. Kluwer Academic Publishers, Il Ciocco, Italy, 1988.

[147] Rolf Pfeifer. The fungus eater approach to emotion: a view from artificial intelligence. *Cognitive Studies*, 1:42–57, 1994.

[148] Rosalind Picard. *Affective Computing*. MIT Press, 1997.

[149] Rosalind Picard, Seymour Papert, Walter Bender, Bruce Blumberg, Cynthia Breazeal, David Cavallo, Tod Machover, Mitchel Resnick, Deb

Roy, and Carol Strohecker. Affective learning — a manifesto. *BT Technical Journal*, 22(4):253–269, October 2004.

[150] Rosalind W. Picard. Affective computing. Technical Report 321, M.I.T. Media Laboratory; Perceptual Computing Section, November 1995.

[151] Rosalind W. Picard and Jocelyn Scheirer. The galvactivator: A glove that senses and communicates skin conductivity. In *Proceedings from the 9th International Conference on Human-Computer Interaction*, pages 1538–1542, New Orleans, August 2001.

[152] Rosalind W. Picard, Elias Vyzas, and Jenifer Healey. Towards machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1175–1191, October 2001.

[153] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, 1993.

[154] Anand S. Rao and Michael P. Georgeff. BDI agents: From theory to practice. In *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, San Francisco, USA, June 1995.

[155] W. Scott Reilly. *Believable Social and Emotional Agents*. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, May 1996. Technical Report CMU-CS-96-138.

[156] W. Scott Reilly and Joseph Bates. Building emotional agents. Technical Report CMU-CS-92-143, CMU, School of Computer Science, Carnegie Mellon University, May 1992.

[157] Stanley J. Rosenschein and Leslie Pack Kaelbling. A situated view of representation and control. *Artificial Intelligence*, 73, 1995.

[158] Sheldon M. Ross. *Introduction to Probability and Statistics for Engineers and Scientists*. Wiley, 1987.

[159] Sam R. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, December 2000.

[160] David R. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: Foundations*, chapter Learning Internal Representations by Error Propagation, pages 318–362. MIT Press, 1986.

[161] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, second edition, 2003.

[162] Rui Sadio, Gonçalo Tavares, Rodrigo Ventura, and Luis Custódio. Aplicação de emoções artificiais num robot móvel. Technical Report Internal Report RT-704-00, ISR, 2000.

[163] Rui Sadio, Goncalo Tavares, Rodrigo Ventura, and Luis Custódio. An emotion-based agent architecture application with real robots. In Dolores Cañamero, editor, *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-01-02*, pages 117–122, Menlo Park, California, 2001. The AAAI Press.

[164] David Sander, Didier Grandjean, and Klaus R. Scherer. A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18:371–352, 2005.

[165] Klaus R. Scherer. *Handbook of Cognition and Emotion*, chapter Appraisal Theory, pages 637–663. Wiley, 1999.

[166] Matthias Scheutz. The evolution of simple affective states in multi-agent environments. In Dolores Cañamero, editor, *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-01-02*, pages 123–128, Menlo Park, California, 2001. The AAAI Press.

[167] Matthias Scheutz. A framework for evaluating affective control. In Robert Trappl, editor, *Cybernetics and Systems 2004*, pages 645–650, Wien, 2004.

[168] Matthias Scheutz. How to determine the utility of emotions. In Eva Hudlicka and Dolores Cañamero, editors, *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, Papers from 2004 AAAI Spring Symposium, Technical Report SS-04-02*, pages 122–127, Menlo Park, California, 2004. The AAAI Press.

[169] J. R. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3:417–457, 1980.

[170] Danbing Seto and Lui Sha. A case study on analytical analysis of the inverted pendulum real-time control system. Technical Report CMU/SEI-99-TR-023, Carnegie Mellon University, November 1999.

[171] Herbert A. Simon. Theories of decision making in economics and behavioural science. *American Economic Review*, 49:253–83, June 1959.

[172] Herbert A. Simon. Motivational and emotional controls of cognition. *Psychological Review*, 74(1):29–39, 1967.

[173] Herbert A. Simon. *The Sciences of the Artificial*. MIT Press, third edition edition, 1996.

[174] Aaron Sloman. chapter What sort of architecture is required for a human-like agent? In [215].

[175] Aaron Sloman. Damásio, Descartes, alarms and meta-management. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, pages 2652–2657, October 1998.

[176] Aaron Sloman and Monica Croucher. Why robots will have emotions. In A. Drinan, editor, *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI-81)*, Menlo Park, California, June 1981. The AAAI Press.

[177] Aaron Sloman and Matthias Scheutz. A framework for comparing agent architectures. In *Proceedings of UK Workshop on Computational Intelligence*, September 2002.

[178] Craig A. Smith and Leslie D. Kirby. *Feeling and Thinking: The role of affect in social cognition*, chapter Consequences Require Antecedents, pages 83–106. Cambridge University Press, New York, 2000.

[179] Alexander Staller and Paolo Petta. Towards a tractable appraisal-based architecture. In Dolores Cañamero, Chisato Numaoka, and Paolo Petta, editors, *Workshop: Grounding Emotions in Adaptive Systems*, pages 56–61. SAB'98: From Animals to Animats, August 1998.

[180] Wolfgang Stolzmann. Anticipatory classifier systems. In *Genetic Programming 1998: Proceedings of the Third Annual Conference*, pages 658–664, San Francisco, CA, July 1998. Morgan Kaufmann.

[181] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning*. MIT Press, 1998.

[182] John D. Teasdale. *Handbook of Cognition and Emotion*, chapter Multi-level Theories of Cognition-Emotion Relations, pages 665–681. Wiley, 1999.

[183] Naftali Tishby, Fernando C. Pereira, and William Bialek. The information bottleneck method. In *Proceedings of the 37-th Annual Allerton Conference on Communication, Control and Computing*, pages 368–377, 1999.

[184] Masanao Toda. *Man, robot and society*, chapter Emotional Fungus-eaters, pages 130–153. The Hague: Nijhoff, 1982.

[185] Bill Tomlinson and Bruce Blumberg. Social behavior, emotion and learning in a pack of virtual wolves. In Dolores Cañamero, editor, *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-01-02*, pages 135–140, Menlo Park, California, 2001. The AAAI Press.

[186] Bill Tomlinson, Marc Downie, Matt Berlin, Jesse Gray, Derek Lyons, Jennie Cochran, and Bruce Blumberg. Leashing the AlphaWolves: Mixing user direction with autonomous emotion in a pack of semi-autonomous virtual characters. In *Symposium on Computer Animation, Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 7–14, San Antonio, Texas, 2002. ACM Press.

[187] Daniel Tranel and Antonio R. Damásio. The covert learning of affective valence does not require structures in hippocampal system or amygdala. *Journal of Cognitive Neuroscience*, 5(1):79–88, 1993.

[188] M. S. Triantafyllou, G. S. Triantafyllou, and D. K. P. Yue. Hydrodynamics of fishlike swimming. *Annual Review of Fluid Mechanics*, 32:33–53, January 2000.

[189] Alan M. Turing. *Machine Intelligence*, volume 5, chapter Intelligent Machinery, pages 3–23. American Elsevier Publishing, 1970.

[190] Paul E. Utgoff, Neil C. Berkman, and Jeffery A. Clouse. Decision tree induction based on efficient tree restructuring. *Machine Learning*, 29(1):5–44, 1997.

[191] Pedro Vale and Luís Custódio. Learning individual basic skills using an emotion-based architecture. In *Proceedings of the Symposium on Emotion, Cognition, and Affective Computing, AISB'01 Convention*, 2001.

[192] Pedro Vale and Luís Custódio. Combining reinforcement learning with an emotion based architecture. In *Proceedings of International Conference on Artificial Intelligence and Applications 2002 (IASTED)*, pages 41–46, Malaga, Spain, 2002. ACTA Press. Special session on Automated Reasoning: Perception and Emotions.

[193] Juan D. Velásquez. Modeling emotions and other motivations in synthetic agents. In *Proceedings of the 14th National Conference on Artificial Intelligence*, pages 10–15, Cambridge/Menlo Park, 1997. The AAAI Press and The MIT Press.

[194] Juan D. Velásquez. Modeling emotion-based decision-making. In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-98-03*, pages 164–169, Menlo Park, California, 1998. The AAAI Press.

[195] Juan D. Velásquez. When robots weep: Emotional memories and decision-making. In *Proceedings AAAI-98*, pages 70–75. AAAI, AAAI Press and The MIT Press, 1998.

[196] Rodrigo Ventura. Emotion-based agents. Master's thesis, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal, March 2000.

[197] Rodrigo Ventura, Luis Custódio, and Carlos Pinto-Ferreira. Artificial emotions — goodbye Mr. Spock! In Helder Coelho, editor, *Progress in Artificial Inteligence, Proceedings of the 6th Iberoamerican Conference (IBERAMIA'98)*, pages 395–402. Colibri, 1998.

[198] Rodrigo Ventura, Luis Custódio, and Carlos Pinto-Ferreira. Emotions — the missing link? In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the 1998 Fall Symposium, Technical Report FS-98-03*, pages 170–175, Menlo Park, California, 1998. The AAAI Press.

[199] Rodrigo Ventura, Luis Custódio, and Carlos Pinto-Ferreira. Learning courses of action using the "movie-in-the-brain" paradigm. In Dolores Cañamero, editor, *Emotional and Intelligent II: The Tangled Knot of Social Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-01-02*, pages 147–152, Menlo Park, California, 2001. The AAAI Press.

[200] Rodrigo Ventura and Carlos Pinto-Ferreira. Emotion-based agents. In *Proceedings of the 15th National Conference on Artificial Intelligence*, page 1204, Cambridge/Menlo Park, 1998. AAAI Press and The MIT Press.

[201] Rodrigo Ventura and Carlos Pinto-Ferreira. Meaning engines — revisiting the chinese room. In Dolores Cañamero, Chisato Numaoka, and Paolo Petta, editors, *Workshop: Grounding Emotions in Adaptive Systems*, pages 68–70. SAB'98: From Animals to Animats, August 1998.

[202] Rodrigo Ventura and Carlos Pinto-Ferreira. Emotion-based agents: Three approaches to implementation (preliminary report). In Juan Velásquez, editor, *Workshop on Emotion-Based Agent Architectures (EBAA'99)*, pages 121–129, May 1999.

[203] Rodrigo Ventura and Carlos Pinto-Ferreira. Emotion-based control systems (preliminary report). In *Proceedings of the 1999 IEEE International Symposium on intelligent Control*, pages 64–66, Cambridge, MA, September 1999. IEEE.

[204] Rodrigo Ventura and Carlos Pinto-Ferreira. A formal indexing mechanism for an emotion-based agent. In *Proceedings of International Conference on Artificial Intelligence and Applications 2002 (IASTED)*, pages 34–40, Malaga, Spain, 2002. ACTA Press. Special session on Automated Reasoning: Perception and Emotions.

[205] Rodrigo Ventura and Carlos Pinto-Ferreira. From reactive to emotion-based agents: a prescriptive model. In Robert Trappl, editor, *Cybernetics and Systems 2000*, pages 757–761, Wien, 2002.

[206] Rodrigo Ventura and Carlos Pinto-Ferreira. Indexing by metric adaptation and representation upgrade in an emotion-based agent model. In *Proceedings of the 3rd International Conference on Natural Computation*, pages 108–112, 2007.

[207] Rodrigo Ventura and Carlos Pinto-Ferreira. Metric adaptation and representation upgrade in an emotion-based agent model. In Ana Paiva, Rui Prada, and Rosalind W. Picard, editors, *Affective Computing and Intelligent Interaction: Second International Conference (ACII2007)*, LNCS 4738, pages 731–732. Springer, 2007.

[208] Rodrigo Ventura and Carlos Pinto-Ferreira. A multidimensional scaling approach to indexing by metric adaptation and representation upgrade. In *Proceedings of the Twentieth International Florida Artificial Intelligence Research Society Conference (FLAIRS-07)*, pages 636–637, 2007.

[209] John von Neumann. *The Computer and the Brain.* Yale Nota Bene, second edition edition, 2000.

[210] Elias Vyzas and Rosalind W. Picard. Affective pattern classification. In Dolores Cañamero, editor, *Emotional and Intelligent: The Tangled Knot of Cognition, Papers from the AAAI Fall Symposium, Technical Report FS-98-03*, pages 176–182, Menlo Park, California, 1998. The AAAI Press.

[211] Ian Wilson. The artificial emotion engine, driving emotional behavior. In Wolff Dobson, editor, *Artificial Intelligence and Interactive Entertainment, Papers from 2000 AAAI Spring Symposium, Technical Report SS-00-02*, Menlo Park, California, 2000. The AAAI Press.

[212] Mark Witkowski. *Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories, and Systems*, volume 2684 of *Lecture Notes in Computer Science*, chapter Towards a Four Factor Theory of Anticipatory Learning, pages 66–85. Springer, 2004.

[213] David H. Wolpert and William G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, April 1997.

[214] Michael Wooldridge and Nicholas R. Jennings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152, 1995.

[215] Michael Wooldridge and Anand Rao, editors. *Foundations of Rational Agency.* Kluwer Academic Publishers, 1997.

[216] Ian Paul Wright. *Emotional Agents.* PhD thesis, Faculty of Science of the University of Birmingham, February 1997.

[217] Eric P. Xing, Andrew Y. Ng, Michael I. Jordan, and Stuart Russell. Distance metric learning, with application in clustering with side-information. *Advances In Neural Information Processing Systems*, 15:521–528, 2003.

[218] Shlomo Zilberstein. Using anytime algorithms in intelligent systems. *AI magazine*, 17(3):73–83, Fall 1996.