Università di Roma “La Sapienza” Scuola di Dottorato in Neuroscienze

Dottorato in Psicobiologia e Psicofarmacologia

Exploring the Psychobiology of Emotions and Motivations through Computational Models

Anno Accademico 2008/2009

Candidato Francesco Mannella

Relatore Prof. Stefano Puglisi Allegra

A Carmela,
per il suo amore
e la sua pazienza.

Contents

Acknowledgments

Francesco Mannella
Roma
November 27, 2009

The author has been supported by the EU-funded Integrated Project *ICEA - Integrating Cognition, Emotion and Autonomy*, contract no. FP6-IST-027819-IP.

The author is also grateful to the “Laboratory of computational Embodied Neuroscience” within the “Istituto di scienze e tecnologie della cognizione” of the Italian National Research Council, namely all the current and former members of the “Room 23”.

A particular mention is due to the group leader Dr. Gianluca Baldassarre who was a very precious teacher, a brilliant scientist and a true friend.

The author wants to thank all people and groups collaborating during these years, mostly within very interesting workshops and conferences. A special thank to Prof. Simona Cabib for the several stimulating scientific discussions.

The author is sincerely grateful to his advisor, Prof. Stefano Puglisi-Allegra, for his guide and introduction to the heart of the most recent techniques and studies.

Abstract

This thesis investigates emotions and motivations on the basis of an operational approach. This approach has both computational and psychobiological roots. Three main directions of research are followed: (*1*) investigation on the neural substrates of emotional systems though the exploration of the literature about comparative functional anatomy and physiology; (*2*) definition the relationship between emotion, cognition and behaviour through the exploration of the psychobiological literature about animal models; (*3*) building of computational models constrained by the sources of information 1 and 2; (*4*) testing the behaviour of such models within simulated robots acting in simulated environments. The main focus will be on the interaction between the emotional and motivational systems and high level cognitive processes behind adaptive behaviour. The whole study will be informed by the current psychobiological knowledge about the functioning of the neural systems pivoting on amygdala, given that this is considered to be one of the major nodes of interaction between the processing of internal values and the processing about the past, current and future world outside the organism in superior vertebrates.

Chapter 1

Introduction

This chapter outlines the main approaches to emotions in psychology through a brief historical survey. Then it presents an excursus over the major fields in psychobiological research on motivation and emotion. Then it describes the building blocks of our attempt to link the current psychological mainstream theories on emotions with the psychobiological data through a systemic approach based on embodied computational models. Finally it presents an outline of the structure of the thesis.

1.1 The Psychobiology of Emotions and Motivations

1.1.1 Psychological theories of emotions

The debate about the nature and the functions of emotions is a major thread of the study of psychology as a science since its beginning. Unfortunately, an operational unanimous definition of what emotions are and what they are for has not be found yet within psychological research. We will proceed through the main stages of the history of psychology of emotions, trying to find the elements of a possible operational hypothesis about the adaptive function of emotions themselves and their relation with physiological reactions and cognitive processes.

One of the first hypotheses on the origin and nature of emotions was developed independently by two 19th-century scholars, William James (James, 1884) and Carl Lange (Lange, 1912). Their theories state that the autonomic nervous system elicits physiological events such as muscular tension, a rise in heart rate, perspiration, and dryness of the mouth, in response to events in the outer world. Emotions are a result of these physiological changes, rather than being their cause. In this view, the cognitive experience of emotion follows the physiological expression of emotion.

Although the James-Lange theory became popular in the early twentieth century, it soon came under attack. In 1920, the american physiologist Walter Cannon published a paper containing several compelling criticisms of the James-Lange theory and went on to propose a new theory (Cannon, 1920). Cannon's theory was modified by Philip Bard (Bard, 1934), and the Cannon-Bard theory of emotion, as it came to be known, proposed that emotional experience can occur independently of emotional expression. Where James argued that emotional behaviour often precedes or defines the emotion, Cannon and Bard argued that the emotion arises first and then stimulates typical behaviour. One of the arguments Cannon used against James-Lange theory consisted in observations of animals with transection of the spinal cord. Those animals, which surgery had deprived of body sensation below the level of the cut, still exhibited sign of emotional experiences. Bard described this in a anatomophysiological way by claiming that subcortical structures such as hypothalamus and thalamus process information in order to both regulate the peripheral signs of emotions and to provide the cortex with the information needed for the cognitive processing of emotions. This claim was supported by

the physiological literature on decorticated animals.

In the meanwhile, behaviourist theories were taking place, dominating the scene during the first half of the twentieth century. Behaviourist theorizations treated emotions as an intervening motivational variable, often in the sense of a general drive state. Activation theory, by Duffy (1957), can be seen as a classical example of this trend, stating that only one major dimension of emotion exists (activation, arousal).

During the sixties, the criticism made by Magda Arnold and Richard Lazarus to the behaviouristic view of emotions, became the root of a cognitive theory of emotions based on appraisal. In their view emotions are the product of unconscious evaluation on a situation while feelings are the conscious reflection of the unconscious appraisal (Scherer, Schorr, and Johnstone, 2001). According to Arnold (1960), the initial appraisals start the emotional sequence and arouse both the appropriate actions and the emotional experience itself, so that the physiological changes, recognized as important, accompany, but do not initiate, the actions and experiences. Lazarus et al. (1970) identified two essential goals for the research on emotional processes: "first, what is the nature of the cognitions (or appraisals) which underlie separate emotional reactions (e.g. fear, guilt, grief, joy, etc.). Second, what are the determining antecedent conditions of these cognitions." He specified two major types of appraisal processes which sit at the crux of the appraisal method: (1) primary appraisal, directed at the establishment of the significance or meaning of the event to the organism, and (2) secondary appraisal, directed at the assessment of the ability of the organism to cope with the consequences of the event. These two processes go hand in hand as the first establishes the importance of the event while the second assesses the coping mechanisms. Lazarus divided the latters into two parts: direct actions and cognitive reappraisal processes.

The ideas of Arnold and Lazarus failed to have an immediate effect on psychology of emotion, even though Lazarus's theorizing strongly influenced stress research from the moment of its pubblication. Instead, a theory proposed by Stanley Schachter and his student Jerome Singer in the late fifties, became the representative theory of emotions for more than twenty years. The theory put together the James-Lange peripheral theory, the behaviouristic idea of a general, unspecific arousal, and the psychological hypothesis that human experience was largely based on one's self-observation of a contextual behaviour. The so called 'two-factor' theory of emotion claimed that "cognitions are used to interpret the meaning of physiological reactions to outside events." (Schachter and Singer, 1962). In the experiment by which they tested the theory, Schachter and Singer induced symptoms of sympathetic activation using epinephrine and manipulated emotion inference by confronting their subjects with the emotional behaviour of a stooge. The explanation of the results given by the authors was that (1) both cognitive and physiological factors contribute to emotion, (2) under certain circumstances cognition follows psychological arousal, (3) in part ,people assess their emotional state by observing how physiologically stirred up they are (Schachter and Singer, 1962). Despite consistent negetive evidence in the attempts to replicate the results, in the eighties this theory proved to be very influential (Scherer et al., 2001).

Currently, there are three major theoretical positions on the functioning of the emotional system. First, the work of Antonio Damasio (Damasio, 1994) within the neuropsychological research, partly recovering James-Lange tradition, distinguishes the physiological affective activation (emotions) from the cognitive processing biased by this activation (feelings). Damasio hypothesis is that

cognitive processes are biased by the emotions, described as the physiological affective states elicited by reinforcers. The amygdala and the orbitofrontal cortex lead the biasing of the cognitive elaboration occurring in the ventromedial prefrontal cortex. In particular, emotions act as markers to the cognitive processing of past, present and anticipated states of the world. They act as 'somatic markers', as their elicitation highly depends on the processing about the states of the internal milieu.

Second, the tradition coming from neo-behaviourism of Hull and Tolman defines emotions as behavioural states elicited by reinforcers. Cardinal, Parkinson, Hall, and Everitt (2002) state that "it is useful to consider under the umbrella of emotion those neural processes by which an animal judges and represents the *value* of something in the world, and responds accordingly". According to Rolls (2000) emotions are states elicited by rewards and punishments, including changes in rewards and punishments. Reinforcement association of stimuli, encoded in the orbitofrontal cortex and amygdala, is sufficient to elicit emotion-based learning and to affect behaviour via, for example, the orbitofrontal-striatal pathway (Rolls, 1999). Rolls defines several factors accounting for emotions: (*1*) the reinforcement contingency; (*2*) the intensity of reinforcement; (*3*) multiple associations of environmental stimuli with rewards or punishments; (*4*) the elicitation of different emotions by different primary reinforcers; (*5*) the differentiation of emotions elicited by different secondary reinforcing stimuli sharing a similar primary reinforcer; (*6*) the differentiation of emotions elicited on the base of an active or passive behavioural response being possible.

Finally all the current approaches within the framework of the appraisal theories share few major points (Scherer, Schorr, and Johnstone, 2001): *1*) emotions are differentiated by appraisals, each distinct emotion is elicited by a distinctive pattern of appraisal; *2*) differences in appraisal can account for individual and temporal differences in emotional response; *3*) all situations to witch the same appraisal pattern is assigned will evoke the same emotion; *4*) appraisals precede and elicit emotions, appraisals start the emotion process, initiating the physiological, expressive and behavioural reactions and other changes that comprise the resultant emotional state; *5*) the appraisal system has evolved to process information that predicts when particular emotional responses are likely to provide effective coping; *6*) conflicting, involuntary, or inappropriate appraisal may account for irrational aspects of emotions; *7*) changes in appraisals may account fo developmentally and clinically induced changes in emotions. A recent implementation of the appraisal theories is the component process model (CPM), by Sander, Grandjean, and Scherer (2005). The author claims this is a systemic approach to the study of appraisal mechanisms and is characterized by three features; (1) appraisal is a complex function: different systems stand behind appraisal at different levels. (2) the appraisal mechanisms have a correspondence in the neural systems that can be studied with psychobiological instruments. (3) The detailed specification of hypotheses constrained by neurophysiological data allows an analysis of them through their implementation in computational models.

These three approaches share some important features. In all the three approaches the physiological and behavioural responses are viewed as a consequence of a first evaluation (incentive salience, novelty) made via the interaction of internal states of the organism with external states (e.g. Bechara, Damasio, Damasio, and Lee, 1999; Rolls, 1999; Sander, Grandjean, and Scherer, 2005). Moreover, in all approaches the primary evaluations have a role in biasing higher level cognitive processes.

Some major differences can also be found in the three theorical frameworks. First, both Somatic-marker hypothesis and neo-behaviourism research focus on the interaction between emotions and

cognitive processes as choice, goal-oriented behaviours and decision making, while appraisal theories focus on the mechanisms that produce emotions starting from different patterns of evaluations. Second, somatic marker hypothesis is originally based on the claim that the expression of emotions is by itself the marker, while both neo-behaviourist tradition (see Balleine, 2005) and current appraisal theories (see Sander, Grandjean, and Scherer, 2005) state that the evaluation process stands behind the emotional expression. This last differentiation is fading down as far as research on both sides focuses on the underlying neural systems. The importance of the associative processing within amygdala in both the elicitation of emotions and in biasing decision making processes is leading to a common view that this associative system is behind, or can bypass, the very expression of emotions.

As current theories are getting more and more close to the psychobiological work of Balleine, Berridge, Dickinson, Everitt, LeDoux, and current psychobiological research using animal models, a systemic and unified explanation of all mechanisms behind emotions and motivations is emerging. One main point of this systemic explanation is that several different neural systems interactively determine the value of a state of the world at different levels of cognitive complexity. Furthermore, the great amount of interaction between the different neural components must be considered within a dynamical system framework (see the concept of circular causality in Lewis 2005).

1.1.2 Animal models of Conditioning and Affective Processes

The neo-behaviourist paradigm coming from Hull (1943) and Tolman (1932) granted to cognitive psychology the use of two fundamental classes of instruments in the experimental research: (*1*) instrumental and pavlovian behavioural paradigms; (*2*) the controlling of neurophysiological variables through lesioning, inactivation and microdialysis techniques. Through the use of animal models cognitive psychology is opening to the experimental study of the neural systems involved in low-and-higher-level cognitive processes. The study of emotions and motivations through animal models acquired a huge amount of new data in the last twenty years. The main fields can be distinguished as works on appetitive pavlovian conditioning, fear conditioning, instrumental conditioning and goal-oriented behaviours, incentive salience and hedonic values, and stress. Here we will rapidly go through these fields of study indicating the different contributions to the exploration of emotions and motivations.

In this section, and everywhere in the rest of the thesis we will mostly refer to research on rats. This choice is due to the huge amount of data that neuroanatomy, neurophysiology and psychophysiology has acquired about the rat nervous system, and, consequently, to the enormous collection of experimental works on rats in the last sixty years. Psychobiological studies on appetitive conditioning (e.g. Blundell, Hall, and Killcross, 2001; Cardinal, Parkinson, Lachenal, Halkerston, Rudarakanchana, Hall, Morrison, Howes, Robbins, and Everitt, 2002; Cardinal, Parkinson, Marbini, Toner, Bussey, Robbins, and Everitt, 2003; Gallagher, Graham, and Holland, 1990; Han, McMahan, Holland, and Gallagher, 1997; Hatfield, Han, Conley, Gallagher, and Holland, 1996; McDannald, Kerfoot, Gallagher, and Holland, 2005; Parkinson, Robbins, and Everitt, 2000; Petrovich and Gallagher, 2007; Petrovich, Holland, and Gallagher, 2005; Petrovich, Setlow, Holland, and Gallagher, 2002; Pickens, Saddoris, Setlow, Gallagher, Holland, and Schoenbaum, 2003; Setlow, Gallagher, and Holland, 2002) and goal-oriented behaviours (e.g. Balleine and Dickinson, 1998; Cador, Robbins, and Everitt, 1989; Corbit and Balleine, 2003; Corbit, Muir, and Balleine, 2001; Dickinson, 1985; Yin and Knowl-

ton, 2006) focus of the function of amygdala (AMG), ventromedial prefrontal cortex (vmPFC), orbitofrontal cortex (OFC) and nucleus accumbens (NAcc), and their interactions in appetitive pavlovian conditioning and instrumental conditioning processes.

A wide use of lesion techniques is made in which neural components are lesioned (typically injecting a N-methyl D-aspartate preparation into in the neural region) or inactivated (usually through the injection of a muscimol preparation) before or after one of the stages of the experimental scheduling.

This kind of research allows building a mapping of a neural system in which the absence of each different component can be associated with a specific alteration of the overall functioning. Possibly specific sub-functions can be associated with a single component or a sub-group of components. This area can be easily seen as a starting point for the psychobiology of motivations. Associative learning mechanisms allow organisms to transfer behaviours from reinforcers to neutral stimuli. Moreover, these learning mechanisms allow the previously neutral stimuli to bias cognitive processes after having acquired the value of predictors of reinforcers.

The object of the studies on fear conditioning (Armony, Servan-Schreiber, Romanski, Cohen, and LeDoux, 1997; Blair, Schafe, Bauer, Rodrigues, and LeDoux, 2001; Blair, Sotres-Bayon, Moita, and LeDoux, 2005; Calandreau, Desmedt, Decorte, and Jaffard, 2005; Davis, 1992a,b; LaLumiere, Buen, and McGaugh, 2003; Lanuza, Nader, and LeDoux, 2004; LeDoux, 1996, 2003; Nader, Majidishad, Amorapanth, and LeDoux, 2001; Paré, Quirk, and LeDoux, 2004) is the role of the amygdala in fear behaviour. Fear related behaviours are historically the first class of behaviours to be linked to the functioning of the amygdala in superior vertebrates. In vivo single-cell activity recording is used in parallel with lesion techniques in order to inspect changes in plasticity. Through the studies some important features of associative learning within the amygdala could be established. Being one of the first fields in which a psychobiological animal model of an emotional state was built, this research field assumed a pivotal role for the whole psychobiology of emotions. Moreover the hypotheses made at the level of the neural microstructure enlighten how plasticity processes lead to associative learning.

Studies on hedonic value versus incentive salience defined with animal models a behavioural differentiation between the neural systems through which incentive salience of a stimulus produces 'wanting' behaviours and those through which the hedonic value given to a stimulus elicit 'liking' behaviours (e.g. Berridge and Robinson, 1998, 2003; Kelley and Berridge, 2002; Peciña, Smith, and Berridge, 2006; Reynolds and Berridge, 2001, 2002; Smith and Berridge, 2005; Tindell, Smith, Peciña, Berridge, and Aldridge, 2006; Wyvell and Berridge, 2000). Methodologies in this research field are based on lesions or inactivations of neural regions too. Some important additional instruments are often used as the depletion of neurotransmitters in specific regions, pharmacological manipulation and even more sophisticated techniques as the combination of drug microinjections and C-Fos immunoreactivity mapping. This set of methodologies contributes to enlighten the functional neurophysiology at the level of specific regions and to associate the activity of different sub-populations of neurons in a region to different behavioural patterns (e.g. hedonic reactions versus eating behaviours in association to different neural populations within the ventral pallidum). From these works the vision emerges that emotional neural processes do not share a monolithic structure within neural modules. Instead, different functions as goal-oriented action and expression of the emotional state are processed by different not completely overlapping neural systems.

Finally, the research on stress (e.g. Amat, Baratta, Paul, Bland, Watkins, and Maier, 2005; Amat,

Paul, Watkins, and Maier, 2008; Amat, Paul, Zarza, Watkins, and Maier, 2006; Bland, Hargrave, Pepin, Amat, Watkins, and Maier, 2003; Cabib and Puglisi-Allegra, 1994, 1996; Maier, 1984; Maier and Watkins, 2005; Pascucci, Ventura, Latagliata, Cabib, and Puglisi-Allegra, 2007) look at the interaction of the catecholamines and other neurochemicals within animal models in controllable and uncontrollable contexts. The use of depletion and microdialysis is the main instrument used in these studies. The result is that the interaction of the activity of different neurochemicals distributed in different regions of the brain can be analysed. The psychobiology of stress has so acquired the notion that controllable and uncontrollable situations elicit different neural systems leading to completely different behaviours. Cortical processing is necessary to the evaluation that a situation is uncontrollable and to start a neural process that leads to passive coping mechanisms and to the acquisition of learned helplessness behaviours.

In summary, modern research has endowed the psychology of motivations and emotions of a wide range of instruments that permit to deeply explore the organization of the neural systems and their relation to behaviour and to acquire quantitative data on neurobiological constraints. In the next session we propose a theoretical framework to unify in an operational framework the evidence from the neuroscience of emotions and motivations.

1.2 Computational Embodied Neuroscience

This thesis presents a proposal to address the issues targeted by the modern psychology of motivations and emotions, with a specific attention to the progress made within psychobiology. The research approach followed might be termed 'CEN – Computational Embodied Neuroscience' (cf. Prescott, Gonzalez, Humphries, and Gurney, 2003, which propose a research method which shares some principles with the approach proposed here) . This method aims at providing general criteria for selecting models so as to produce *theoretical cumulativity* in the study of brain and behaviour. Indeed, the great amount of empirical data provided by neuroscience, psychology and the other related disciplines are seldom integrated by strong and general theoretical explanations, thus failing to produce a coherent picture of the phenomena under investigation. CEN aims to overcome these limits by relying upon the following principles.

Computational models. The investigation of brain and behaviour conducted on the basis of empirical experiments and observations (such as those of neuroscience, psychology and ethology) should be accompanied by the instantiation of theories into formal computational models, that is computer programs that simulate the mechanisms underlying brain processes and produce behaviour as an emergent outcome of their functioning. The rationale behind this principle is that brain and behaviour are complex systems, and theories should be expressed in a strict operational manner so to give truly generative accounts of these phenomena. Computational models both carry in themselves an operational implementation of a neural hypothesis and are able to furnish predictions about behaviour which make it possible to confirm or reject hypotheses with experimental research.

Constraints from behaviour. The computational models used to instantiate the theories have to be capable of reproducing the investigated behaviour, in line with what is proposed by 'artificial ethology'

(Holland and McFarland, 2001). Furthermore, the comparison between the model and the target behaviour should be done on a detailed basis (i.e., with reference to the outcomes of specific target experiments) and possibly in quantitative terms (i.e., not just with vague, qualitative comparisons).

Constraints from brain. Challenging models with the request to account for specific behavioural data is not enough as, given a behaviour, many alternative models capable of reproducing it can always be built. For this reason, a second fundamental source of constraints for models are the data on the anatomy and physiology of brain. These data should be used in two ways. First, for choosing the assumptions that drive the design of the architecture, functioning, and learning mechanisms of the models. Second, for testing the low-level predictions produced by the models (i.e. the predictions produced at the neural level). This principle comes from computational neuroscience (Sejnowski, Koch, and Churchland, 1988) urging computational models to account for data on brain.

Embodiment. In line with the ideas proposed by the 'animats' approach (Meyer and Wilson, 1991) and 'artificial life' (Langton, 1987), models should be capable of reproducing the addressed behaviours within 'whole' autonomous systems acting on the basis of circular interactions with the environment mediated by the body (in particular the sensors, the actuators and internal body systems). Indeed, the brain generates behaviour by forming a large dynamical complex system together with the body and the environment (Nolfi, 2006), so a full understanding of brain and behaviour needs to rely on models that take into consideration this fundamental fact. The principle has two implications. First, the computational models should involve the simulation of both organisms' brains and their body and environment, thus letting behaviours emerge from the interactions between those systems. Second, models should aim at being *scalable to realistic setups*, that is capable of functioning with realistic sensors (e.g. retinas), realistic actuators (e.g. bodies should be governed by realistic Newtonian dynamics), realistic scenarios (e.g. objects with complex textures, shapes, and dynamics), and noise (affecting both sensors and actuators).

Generality. Computational models should aim to reproduce and account for an increasing amount of data taken from an increasing number of different experiments. This principle is important as it is a strong drive towards the production of comprehensive accounts and general theories of brain and behaviour, against the tendency to generate many unrelated and mutually incompatible theories each accounting for only a limited set of empirical data. This principle is in line with the 'spirit' of both 'systems' computational neuroscience, that aims at explaining the functioning of whole brain systems rather than specific areas or physiological/chemical mechanisms, and the animat approach, that aims at identifying general principles and explanations of behaviour.

1.2.1 The features of the built models

Four are the main computational features of the models presented in this thesis: (*1*) Computational models are realized at the level of neural systems, with units that abstract the activity features of entire populations of neurons within each neuroanatomic component. This was eased by the absumption that investigation on the systemic interaction between functional neural subsystems can be reasonably isolated from the study of the detailed dynamic physiological processes at the level of the neural

microstructure. (*2*) The computational systems are tested as part of the control system of robots simulating the organisms of the target experiments within simulated environments. Thus activity within the models can be tested in an open loop with the environment reactions to the robot's actions. (*3*) Simulated environments reproduce the settings of real experimental environments. (*4*) The simulated robots are tested with the same schedulings as that of well known experimental paradigms. Thus simulations can be directly compared with data from psychobiology.

1.3 The structure of the thesis

The remaining chapters of the thesis are organized as follows. Chapter 2 presents a review of all psychobiological evidence on the role of amygdala in the several neural systems involved in emotional and motivational processing.

Chapter 3 to 5 describe three computational models implementing some of the systems described in chapter 2. In particular, Chapter 3 gives a computational hypothesis about how the associative learning processes within amygdala influence the activation of approaching behaviours torward biologically salient stimuli both via direct elicitation of ventral striatum subregions by the basolateral complex of amygdala and by a general sensitization to activation of the ventral striatum through the amplification of the dopaminergic efflux to it via the excitation of the mesolimbic dopaminergic pathway by the central nucleus of amygdala. Chapter 4 shows a model of how the internal associative learning mechanisms of the basolateral complex of amygdala and those of the central nucleus of amygdala have different roles in first-order and second-order conditioning of orienting appetitive responses. Chapter 5 a model presents of the interaction between associative amygdaloid mechanism and the striastocortical substrates of outcome-action association, in generating goal-directed behaviours.

The two appendices at the end of the thesis offer an outline of the computational mathematical instruments and methodology adopted throughout the thesis (Appendix A) and a table of acronyms used when referring to different neural components (Appendix B).

Chapter 2

The Roles of Amygdala in the Affective Regulation of Body, Brain and Behaviour

Abstract

Despite the great amount of data and theories produced by the neuroscientific literature on brain and behaviour on affective phenomena, current models tackling non-cognitive aspects of behavior are often bio-inspired but rarely bio-constrained. This paper presents a theoretical account of affective systems pivoting on amygdala, which aims to furnish a general framework and some specific pathways to implement models which are more closely related to biological evidence. Amygdala, which receives input from brain areas encoding internal states, innately relevant stimuli, and innately neutral stimuli, plays a fundamental role in emotions and motivations of organisms as it can implement two associative processes at the core of Pavlovian learning, plus it has the capacity of modulating them on the basis of internal states. These functionalities allow amygdala to have an important role in the regulation of three fundamental components of emotions (namely the regulation of body states, the regulation of brain states via neuromodulators, and the triggering of a number of basic behaviours fundamental for adaptation) and in the regulation of three high-level cognitive processes (namely the affective labeling of memories, the production of goal-directed behaviours, and the performance of planning/complex decision making). This analysis is conducted within a methodological approach which stresses the importance of understanding brain within an evolutionary/adaptive framework (i.e., stressing how systems involving amygdala increase survival and reproduction of organisms) and with the aim of isolating general principles capable to potentially account for the wider possible empirical evidence in a coherent fashion.

2.1 Introduction

Since the birth of the Cognitive Sciences in the 1950s, the study of *cognitive* functions (e.g. perception, attention, memory, planning, decision making...) has dominated the sciences of behavior, relegating research on the *non-cognitive* aspects of behavior (e.g. motivations, moods, emotions) to a marginal role. This is true in general for all the disciplines dedicated to the study of behavior: for the empirical sciences, from neuroscience to psychology, and for the 'sciences of the artificial' (Simon, 1996), from classical artificial intelligence to the new fields of connectionism, autonomous robotics, artificial life, and the simulation of adaptive behavior.

From the point of view of the sciences of the artificial, while classical artificial intelligence research was exclusively dedicated to the study of cognitive capacities, from their very beginning researchers in artificial life and new robotics presented pioneering works on the affective aspects of behavior (e.g., see Balkenius, 1993; Cecconi and Parisi, 1993; Pfeifer, 1993). The reason for this is related to the significant shift of attention, in the emerging embodied cognition framework, from high level cognitive processes to low level ones, and to the importance attributed to the link between behavior and its biological bases (body, brain, environment). One of the driving ideas of Embodied Cognition research is that of considering behavior and cognition from an adaptivist point of view, that

is on the basis of the advantages that they can give in terms of organisms' capacity to survive and reproduce. From this perspective, the motivational and emotional aspects of behavior are at least as important (but one could argue even more) than the cognitive ones.

The capacity of survival and reproduction of organisms depends on several different abilities, for example the ability to find food and water, the ability to prevent that one's own body gets damaged, and to recover when this happens, the ability to find a sexual partner willing to copulate and reproduce, the ability to escape from predators, the ability to find a suitable place for resting and sleeping, and so on. If an agent has to satisfy all these needs, a crucial 'meta-ability' raises, namely the ability to manage the interactions between all these activities. If in some moments an organism has to solve the problem of satisfying a certain need whereas in other moments it has to solve the problem of satisfying another need, in *each* moment it must solve the problem of establishing which need should be attended. Affective systems allow organisms to solve precisely such a problem, that is to choose which is the activity that has to be accomplished in each moment.

In sharp contrast with what happens in real organisms, artificial systems tend to be designed to accomplish only one or a very few well designed tasks, for example finding the food, *or* navigating in a complex environment, *or* categorising a certain object, *or* grasping and manipulating objects, *or* coordinating with other agents, and so on. In such kinds of agents the problem of selecting which activity to pursue in each moment does not raise because there is only one activity that they can and must pursue in every moment. This is the reason why even in the field of the simulation of adaptive behavior the study of motivations and emotions has always received little attention. In the last years, the realization of the extreme importance that the non-cognitive factors of behavior play in organisms' behavior (Arbib and Fellous, 2004; Cañamero, 2005; Parisi, 2004) has significantly boosted the number of researches dedicated to these aspects in the fields of artificial life and autonomous robotics (e.g. Avila-Garcia and Cañamero, 2004; Balkenius and Morén, 1999; Cañamero, 1997; Mirolli and Parisi, 2003; Montebelli et al., 2007; Murphy, 2002; Venditti et al., 2009).

The relationship between this kind of research and the empirical sciences is quite weak, when not completely absent. Generally speaking, the artificial systems developed in these fields are, at most, biologically inspired (*bio-inspired*) but not really biologically constrained (*bio-constrained*). In other words, the empirical knowledge on the behavior of natural organisms is at most occasionally used as a source of interesting ideas, but is not systematically used for constraining the design of artificial systems, nor for testing their empirical predictions. Such a state of affairs has both its motivations and its potential advantages. For example, a certain division of labour between empirical and artificial scientists is necessary. Furthermore, the freedom of not being constrained by available data and knowledge can lead to the development of new ways of framing old problems and of investigating them (i.e. to new 'research paradigms'), and to the discovery of new interesting specific problems and principles. Finally, it must also be considered that a significant proportion of artificial life research has technological rather than scientific aims, and, from a technological point of view, taking into account how natural organisms work is not a need but, at most, an opportunity.

On the other hand, at least from the scientific point of view, the current state of affairs has also important limits. The biological sciences, and the neural sciences in particular, have been producing a huge amount of knowledge on all the aspects that are relevant for understanding organisms' behavior. Furthermore, this empirical knowledge seems to be doomed to increase at an even higher pace in the

near future. For this reason, trying to incorporate this knowledge more systematically in the design of artificial systems is likely to produce a fundamental positive impact in our ability to build artificial systems with behavioral capacities more similar to those of natural organisms. This, in its turn, would no doubt considerably increase the impact that research on artificial system has on the behavioral and brain sciences. In fact, if it is undeniable that the latter disciplines are producing a great amount of relevant *data*, it is also true that integrative *theories* that are able to explain these data and predict new ones are quite scarce. Bio-constrained computational models represent very promising tools for developing such kind of theories.

With respect to the latter point, it is important to stress that the method followed here is centered on the principles of *Computational Embodied Neuroscience* (or 'CEN', cf. Mannella, Mirolli, and Baldassarre, 2009a). According to this approach, behaviour and brain are seen as means through which organisms adapt to the environment in order to increase their survival and reproduction chances, so a true understanding of brain passes through the comprehension of how it is structured, functions and learns in order to produce adaptive behaviour. Moreover, CEN stresses the importance of producing *general models* directed to capture fundamental principles underlying several different behaviours and brain phenomena instead of ad-hoc models which address only the outcome of specific behavioural or neuroscientific experiments. These two principles also guided the compilation of this review which is indeed supposed to furnish a general framework, but also some specific roadways, to design and implement models having a unifying theoretical scope.

In this paper we contribute to the study of non-cognitive aspects of behavior in artificial system by providing a theoretical framework on behavior that is based on the available empirical knowledge regarding one of the parts of the brain that is at the center of the motivational-emotional systems of higher organisms, namely the amygdala. In particular, we will propose a general brain architecture centered on the amygdala, and a number of specific hypotheses on the various functional roles that amygdala plays in the regulation of both affective and cognitive processes. The neuroscientific and behavioral data taken into consideration mainly refer to literature on rats, however the principles proposed and reviewed in the article can usually be extended to more complex mammals (in particular, non-human primates and humans) as they are very general and generated by parts of the central nervous system of rats which have homologies in all such animal species. Our general framework is intended to boost the design and implementation of biologically-constrained computational models, as the ones presented in previous works (Mannella, Mirolli, and Baldassarre, 2007; Mannella, Zappacosta, and Baldassarre, 2008).

The rest of the paper is structured as follows. Section 2.2 provides a general overview of the amygdala and of the various roles that it plays in the functional organization of adaptive behavior. Section 2.4 illustrates the three main functioning principles that characterize amygdala as the main locus of classical-conditioning associations. Section 2.5 presents the three basic functions that amygdala plays in the regulation of emotional responses. Section 2.6 shows the three higher-level functions that amygala plays by interacting with cognitive processes. Finally, section 2.7 concludes the paper. Note that the acronyms used throughout the paper are listed in the Appendix (table B.1).

2.2 Amygdala's roles in adaptive behavior: overview

Amygdala (Amg) is an almond-shaped group of nuclei located within each medial temporal lobe of the brain (figure 2.2). Amg is an important component of several brain subsystems involving the hypothalamus, insular cortex, brain stem, basal forebrain, hippocampus, basal ganglia, and prefrontal cortex, and it has been associated with a wide range of functions including emotional regulation, learning, action selection, memory, attention, and perception.

The fundamental hypothesis that underlies the framework proposed in this paper, and schematised in figure 2.1, is that amygdala is the place where most classical conditioning associations [1] are acquired on the basis of *three basic mechanisms*, which roughly correspond to the *three major sub-components* in which Amg can be divided, that is CEA, BLA, and MEA:

1. *CEA* associates neutral stimuli (conditioned stimuli, 'CS') directly to basic responses (unconditioned responses, 'UR') that are strictly related to organisms' survival and reproduction on the basis of the experienced co-occurrence of these neutral stimuli and the stimuli that are innately[2] linked to such basic responses by evolution (unconditioned stimuli, 'US'). The result of this process is the formation of CS-UR associations.

2. *BLA* associates neutral stimuli (CS) not directly to the basic responses (UR) but rather to the unconditioned stimuli (US) that are innately associated to those responses on the basis of the CS-US co-occurrences experienced during lifetime. The result of this process is the formation of CS-US associations.

3. *MEA* modulates CEA's and BLA's representations of stimuli and/or responses (in particular, URs and USs) on the basis of internal body states (i.e. on the basis of the current needs of the organism).

Amg performs these functions on the basis of *three main classes of inputs*:

1. *Body states information*, coming from visceral systems, that either constitute unconditioned stimuli or modulate the representations of unconditioned stimuli and responses.

2. *Innately relevant information*, coming from somatosensory, gustatory, and olfactory systems, that represent unconditioned stimuli.

3. *Innately neutral information*, coming from visual, auditory, polimodal, and associative areas, that represent stimuli that can be conditioned (i.e. associated to unconditioned stimuli and/or responses).

The basic unconditioned responses (UR) strictly related to survival and reproduction that amygdala is able to associate to innately neutral stimuli are of three different types, that, in our view, constitute the fundamental aspects underlying *affective* behavior:

[1] Other classical conditioning associations involving for example basic reflexes like eye blinking are known to be stored in the Cer (Thompson, Swain, Clark, and Shinkman, 2000). Of course, all classical conditioning processes involve also other parts of the brain beyond Amg and Cer, such as the brain stem nuclei and PFC.

[2] Note that, in the whole paper, we will use the expressions 'unlearned', 'unconditioned' or 'innate' to refer to responses that might be either innate or developed during the very first phases of life under strong genetic guidance and general environmental constraints (cf. Arias and Chotro, 2007).

1. *Regulation of body states*, accomplished through the links to the sympathetic, parasympathetic and hormonal systems.

2. *Diffuse brain modulation*, accomplished through the links to the four main neuromodulatory systems.

3. *Triggering of unlearned behaviors*, accomplished through the links to the various centers that control such basic behaviors.

Finally, amygdala has at least other three main outputs, through which it can make cognitive processes modulated by emotional states, thus allowing the emerge of new cognitive functionalities:

1. *Affective labeling*, accomplished through the reciprocal connections with the Hip, which is responsible for the encoding and consolidation of episodic memories: these connections allow Amg to include motivational and emotional elements in such memories and to enhance their encoding and recall.

2. *Goal-directed behavior*, accomplished through the connections targeting the NAccC-PL loop, which is responsible for the higher-level stages of action-selection: these connections allow the affective state of an organism to influence the selection of behaviors acquired through operant conditioning.

3. *Planning and decision-making*, accomplished through the reciprocal connections with PFC, which hosts many important cognitive processes such as working memory, attention, and prediction: these connections allow affective states to influence the processes taking place in PFC, thus guiding top-down attention, monitoring of action execution, complex decision making, and planning.

2.3 Anatomy of Amygdala

Three major groups of amygdaloid nuclei can be distinguished as the main loci where the various processes implemented by amygdala take place (Fig. 2.2): the basolateral amygdaloid complex (BLA), the central extended amygdala (CEA), and the medial extended (MEA)

2.3.1 BLA: afferent projections and internal connectivity.

LA is the principal input gateway of Amg (see figure 2.3). LA receives afferent connections directly from the thalamus, from various sensory and associative cortex areas, and from the brain-stem (Maren, 2005; McDonald, 1998; Paré, Quirk, and LeDoux, 2004; Pitkänen, Jolkkonen, and Kemppainen, 2000; Sah, Faber, Armentia, and Power, 2003). These connections allow LA to gather information about the distal world (e.g. through visual and auditory sensors), and information about the proximal world (e.g. through visceral, somatosensory, and gustatory sensors). A wide amount of data show that Ld is the place where USs (visceral, gustatory, somatosensory) and CSs (visual, uditive) information first converges within Amg (Maren, 2005; Pitkänen, Stefanacci, Farb, Go, LeDoux, and Amaral, 1995;

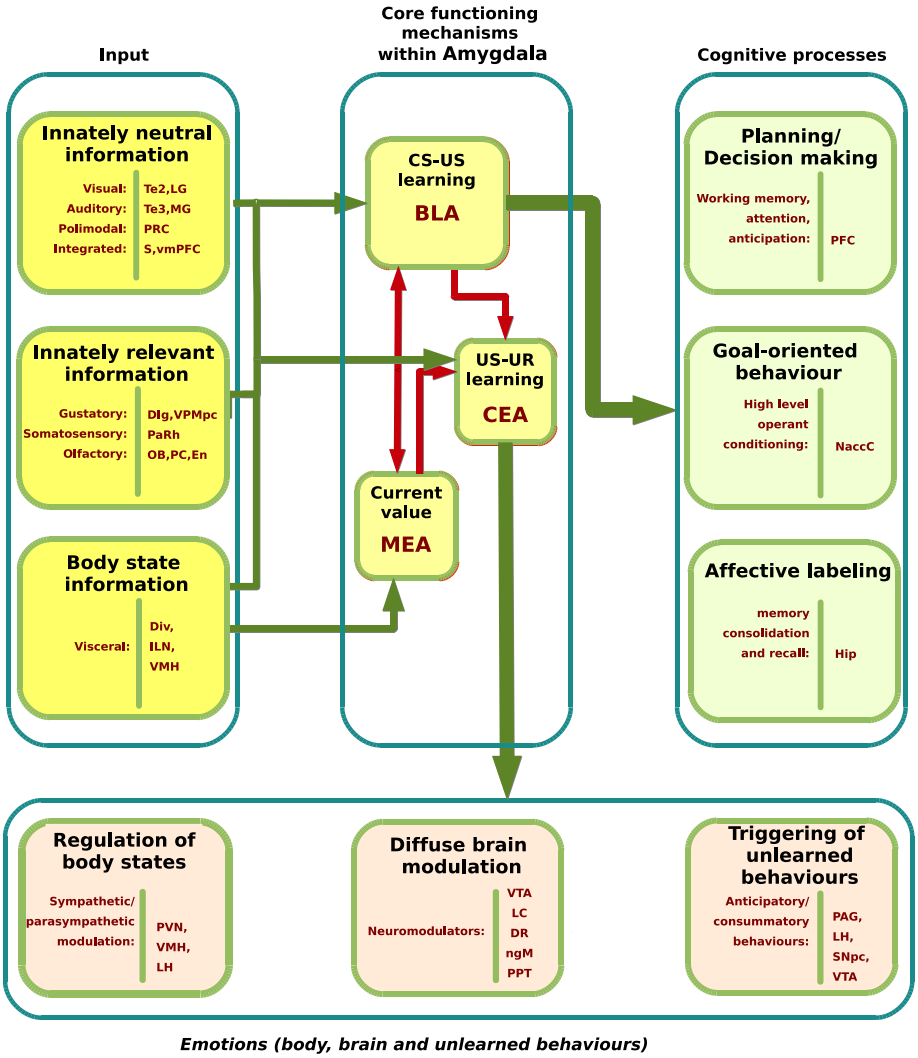


Figure 2.1: *Overview of the Amg: a scheme indicating the main functions played by Amg and the main brain anatomical areas through which it implements such functions. The graph indicates the three main classes of input received by Amg, the three basic mechanisms it implements, the three types of output through which it regulates the emotional systems, and the three main influences it exerts on higher cognitive processes.*

Romanski, Clugnet, Bordi, and LeDoux, 1993). Nevertheless, several data show that, within Ld in rats, gustatory, visceral and somatosensorial projections from cortex are restricted to Lda whereas visual and auditive projections terminate within Ldp (McDonald, 1998). There is also clear evidence about a separation of the projections from Lda and Ldp to Lv: Lda projects mainly to Lvl whereas Ldp projects mostly to Lvm (McDonald, 1998; Pitkänen, Stefanacci, Farb, Go, LeDoux, and Amaral, 1995). As a consequence, it seems plausible that at least part of the gustatory, visceral and somatosensory information remains relatively segregated from sensory information at the level of the lateral dorsal amygdaloid nucleus. Lv has efferent projections to both B and AB. both B and AB also get direct projections from gustatory, visceral, and somatosensory areas and projections from high-converging

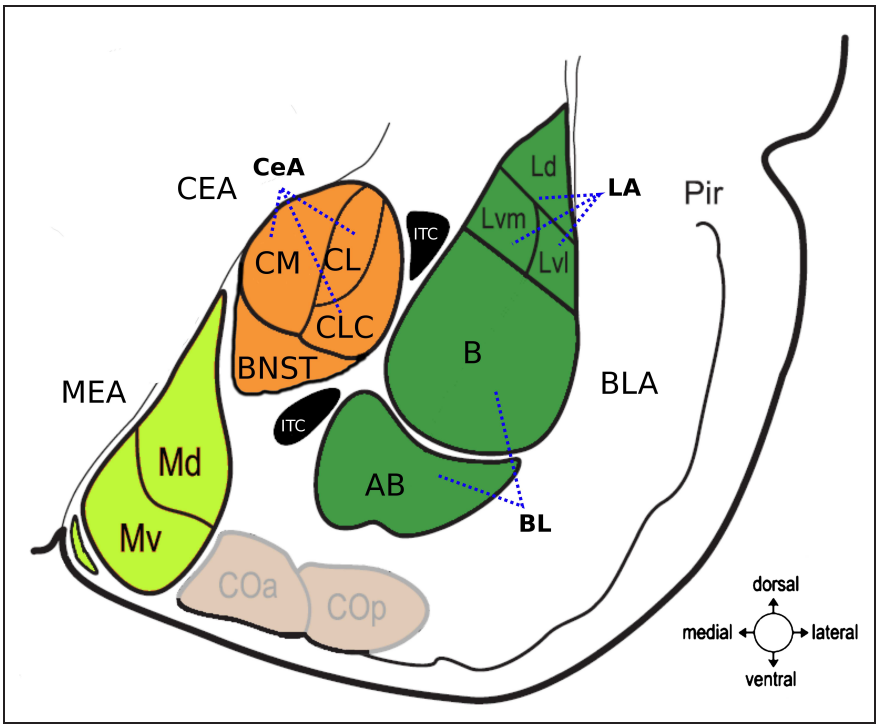


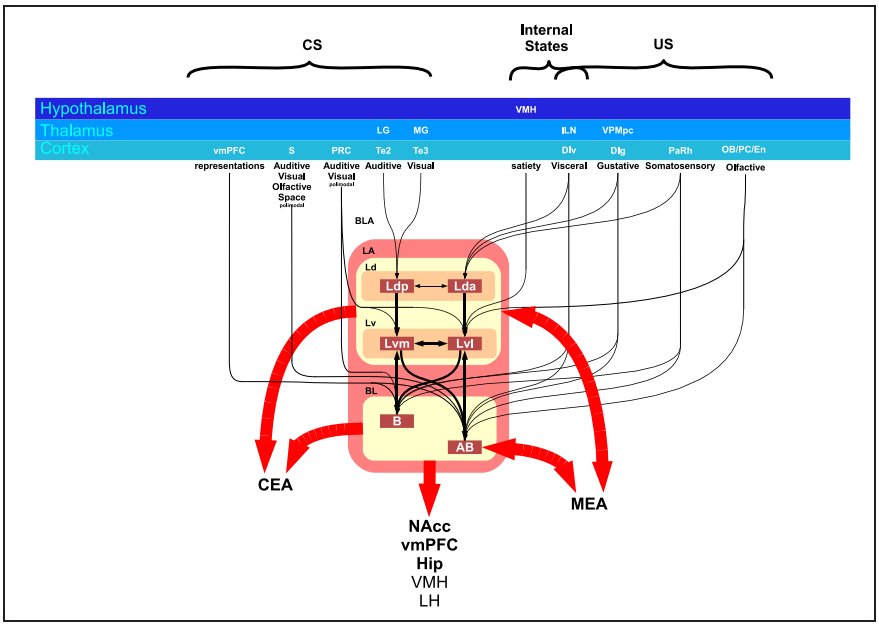
Figure 2.2: *Nuclear divisions and subdivisions of rat amygdala. Acronyms: AB (accessory basal amygdaloid nucleus), B (basal amygdaloid nucleus), BL (basolateral amygdaloid nucleus), BLA (basolateral amygdaloid complex), BNST (bed nucleus of the stria terminalis) CEA (central extended amygdala), CM (central medial amygdaloid nucleus), CL (central lateral amygdaloid nucleus), CLC (central amygdaloid nucleus, lateral capsular subdivision), ITC (intercalated nuclei), LA (lateral amygdaloid nucleus), Ld (lateral dorsal amygdaloid nucleus), Lvl (lateral ventrolateral amygdaloid nucleus), Lvm (lateral ventromedial amygdaloid nucleus), MEA (medial extended amygdala), Md (medial amygdaloid nucleus, dorsal part), Mv (medial amygdaloid nucleus, ventral part).*

areas such as vmPFC and Hip (McDonald, 1998; Pitkänen, Jolkkonen, and Kemppainen, 2000; Price, 2003; Sah, Faber, Armentia, and Power, 2003). One feature at this level is that olfactory information seems to be predominantly directed to AB whereas no inportant olfactory projections have a direct connection with B (McDonald, 1998; Sah, Faber, Armentia, and Power, 2003). Thus within BLA convergence between CSs and USs should take place in two sites organised in sequence: (a) at the level of Lv visceral, somatosensory and gustatory information (USs) converges with bimodal auditory-visual information; (b) at the level of BL information about USs converges with highly integrated polimodal information from hippocampal, cortical associative and cortical prefrontal areas. Notably, within the two rostro-ventral axex of BLA, the Lvl-AB axis is reached by the olfactory signals, whereas the Lvm-B axis has no direct olfactory information.

2.3.2 CeA: afferent projections and internal connectivity.

CeA is reached by afferent intra-amygdaloid projections from LA, BL and MeA (see figure 2.3). LA projections are both direct and undirect. The direct projections reach CLC. The undirect reach CM through a double inhibition (disinhibition) pathway based on two ITC (see figure 2.2 and figure 2.5).

CeA also receives afferent external projections coming from the same thalamic and cortical areas,



A)

Figure 2.3: *the general connectivity organization within BLA. For neuroanatomic data see McDonald (1998), Jolkkonen and Pitkänen (1998), and De Olmos, Beltramino, and Alheid (2004)*

unimodal and polymodal, which project to LA and BL (Jolkkonen and Pitkänen (1998); McDonald (1998); Sah, Faber, Armentia, and Power (2003), see figure 2.4). Moreover, gustatory, visceral and somatosensory information reaches CEA directly. Within CeA, there is a widespread convergence of signals carrying internal and external information, at different level of elaboration, coming from LA, BLA and from other brain areas and ending in CLC and CM (see figure 2.4). CLC is the main point through which BLA modulates CEA. MEA mainly projects to CLC and CL. Within CEA, projections from CLC and CL mainly converge within CM.

2.3.3 MEA: afferent projections and internal connectivity.

Afferent intra-amygdaloid projections to MEA come mainly from BLA, in particular from Lvm (De Olmos, Beltramino, and Alheid, 2004; Pitkänen, Jolkkonen, and Kemppainen, 2000; Pitkänen, Savander, and LeDoux, 1997; Sah, Faber, Armentia, and Power, 2003), Afferent projections of MEA from outside Amg come from two kind of brain areas. First, there are projections from Hyp, coming mainly from VMH, LH and PVN (De Olmos, Beltramino, and Alheid, 2004; Pitkänen, Jolkkonen, and Kemppainen, 2000). These projections reach both Mv and Md nuclei. Second, MEA gets projections from high associative areas such as vmPFC, Hip, and PC. These projections reach only its dorsal part, Md (De Olmos, Beltramino, and Alheid, 2004).

2.3.4 Amygdala: efferent projections to other functional systems.

While LA efferent connections are mainly directed to CeA and BL, these two latter components, together with MEA, have a wide range of targets within the nervous system. Such efferent connections are organised into four major pathways.

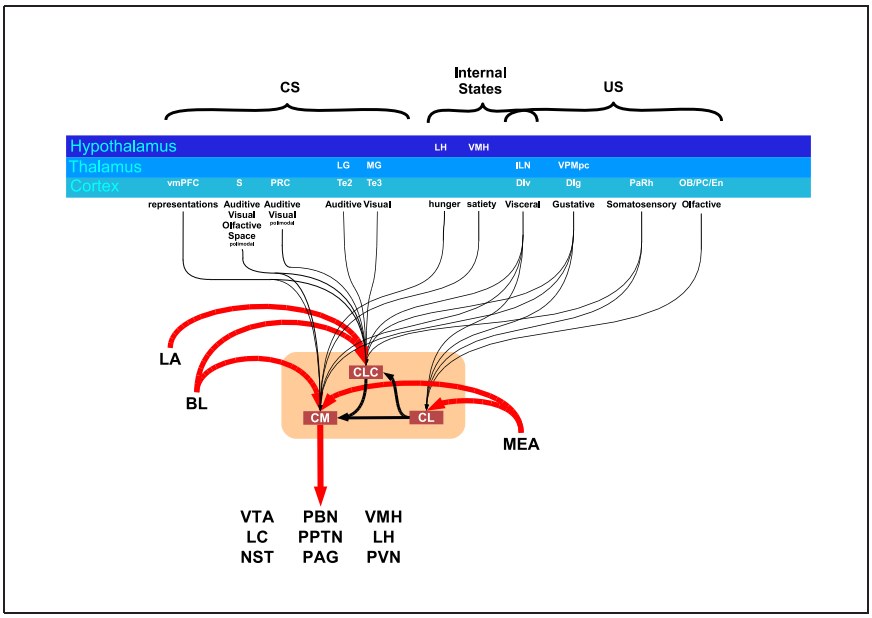


Figure 2.4: *The general connectivity organization within CeA . For neuroanatomic data see McDonald (1998), Jolkkonen and Pitkänen (1998), and De Olmos, Beltramino, and Alheid (2004)*

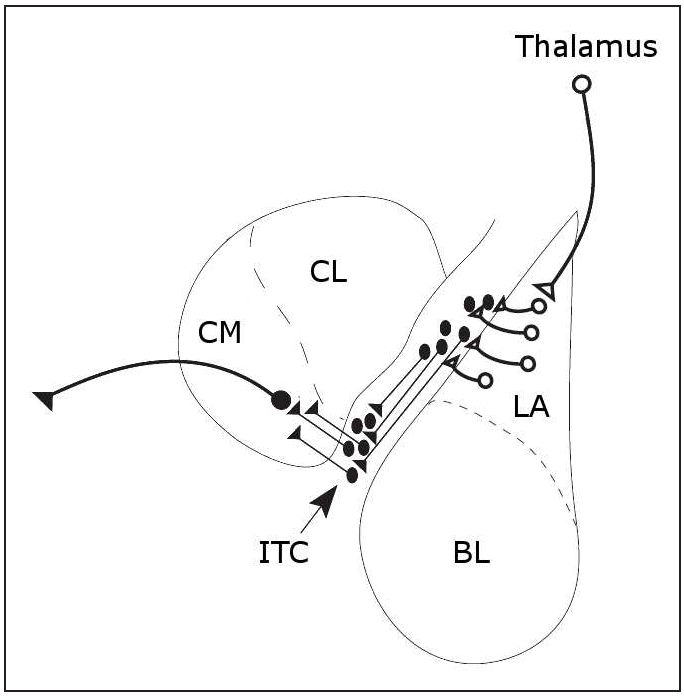


Figure 2.5: *The afferent indirect pathway of CeA based on ITC (modified from Paré, Quirk, and LeDoux 2004).*

CeA is the first major output gateway of Amg through which it regulates whole body and brain states and triggers some unlearned behaviours. In particular, CeA efferent connections are mainly directed to the brain stem, where they target regions such as the PAG controlling unlearned behavioural reactions (e.g. startle and freezing), and other brain systems such as Hyp controlling unlearned body

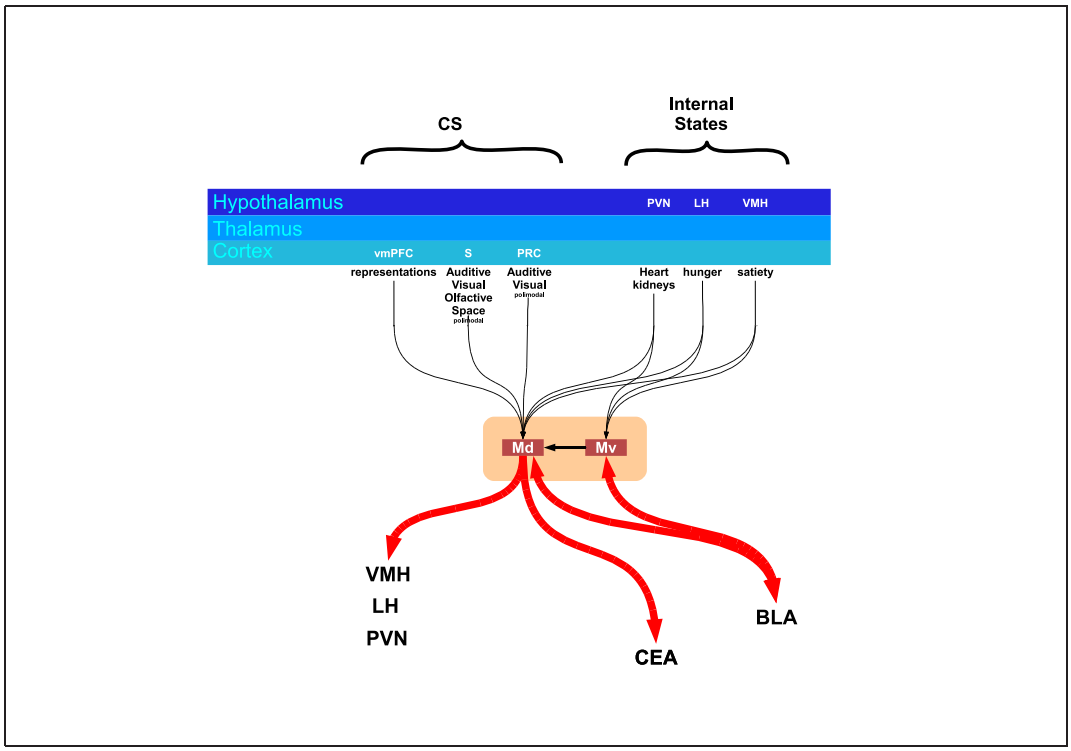


Figure 2.6: *the general connectivity organization within MeA . For neuroanatomic data see McDonald (1998), Jolkkonen and Pitkänen (1998), and De Olmos, Beltramino, and Alheid (2004)*

reactions (Davis and Whalen, 2001; Nader, Majidishad, Amorapanth, and LeDoux, 2001; Paré, Quirk, and LeDoux, 2004; Phelps and LeDoux, 2005). CeA has also efferent connections directed to sites that play a fundamental role in regulating diffuse brain activation, and also in enhancing important learning processes, such as LC, VTA, and Raphe (Fudge and Emiliano, 2003; Pitkänen, Jolkkonen, and Kemppainen, 2000; Rosen, 2004; Weidenfeld, Newman, Itzik, Gur, and Feldman, 2002). These are in fact key control centres of, respectively, the neuromodulators norepinephrine, dopamine and serotonin (5-HT), which play an important role in regulating the sleep/vigilance cycle, the overall brain arousal, attentional processes, and molecular reactions underlying learning.

BL has a major efferent pathway directed to NAcc (Baxter and Murray, 2002; Cador, Robbins, and Everitt, 1989; Pitkänen, Jolkkonen, and Kemppainen, 2000; Sah, Faber, Armentia, and Power, 2003; Setlow, Holland, and Gallagher, 2002). This is the second Amg's major output pathways through which Amg can exert an indirect influence on striato-cortical loops sub-serving habit (S-R) and goal-oriented action selection by transferring Pavlovian incentive values to instrumental behaviours.

BL has also an important strong connection to vmPFC, implemented directly through reciprocal connections with OFC and indirectly through reciprocal connections with Hip, on its turn connected with vmPFC (McDonald, 1998; Richter-Levin, 2004; Rolls, 2000; Saddoris, Gallagher, and Schoenbaum, 2005). This constitutes the third major Amg's output pathway. Through this Amg can influence the behavioural regulatory processes taking place in vmPFC (and downstream pre-motor, motor areas and loops with dorsomedial striatum), including working memory, reasoning, decision making, planning, and goal-oriented action selection. In this regards, the connections between Amg and PFC can

be considered as a fundamental bridge between sub-cortical emotional processes involving overall brain and body internal states and cortical high-level cognitive processes.

BL has also efferent connections directed to CeA, and this allows BL to exert control on the whole range of behavioural reactions and body/brain regulatory processes on which CeA has control (Pitkänen, Jolkkonen, and Kemppainen, 2000).

MeA efferent intra-amygdaloid projections reach Lv and BL nuclei (Pitkänen, Savander, and LeDoux, 1997), while amost all MeA projections outside Amg have Hyp as target (in particular VMH, LH and PVN) (De Olmos, Beltramino, and Alheid, 2004).

2.4 The roles of amygdala in classical conditioning

Individual learning plays a fundamental role in the adaptive behavior of organisms, especially in the most sophisticated ones like mammals. For this reason, animal psychology has devoted great efforts to the study of learning processes. In particular, in the last century a huge body of empirical data have been collected around the two main experimental paradigms of 'classical conditioning' and 'instrumental conditioning'.

Classical conditioning (or *Pavlovian conditioning*) refers to the experimental paradigm in which a certain basic behaviour such as salivation or approaching (UR), that is innately linked to a biologically salient stimulus such as food ingestion (US), gets associated to a neutral stimulus like the sound of a bell (CS), after the neutral stimulus is repeatedly presented before the appearance of the salient stimulus. Such acquired associations, as mentioned in section 2.2, are briefly referred to as 'CS-US' or 'CS-UR' associations (Lieberman, 1993; Pavlov, 1927, see below).

Instrumental conditioning (or *operant conditioning*) refers to an experimental paradigm in which an animal, given a certain stimulus, such as a lever in a cage (the stimulus, 'S'), learns to produce a particular action such as pressing the lever (the response, 'R'), if the performed action consistently leads to a rewarding outcome, such as the access to food. In this case, the acquired associations are briefly referred to as 'S-R' associations (Domjan, 2006; Skinner, 1938; Thorndike, 1911).

The current most influential models of conditioning phenomena, those based on *temporal-difference reward prediction error* (Schultz, 2002; Schultz, Dayan, and Montague, 1997; Schultz and Dickinson, 2000; Sutton and Barto, 1998), suffer of various limitations (cf. Berridge, 2007; Dayan, 2002; Mannella et al., 2007; Redgrave and Gurney, 2006; Redgrave et al., 1999). For example, they tend to conflate classical and instrumental conditioning, and they do not take into account the influences of internal states on the acquisition and expression of conditioned responses. On of the reasons of these limits is that such models have been developed within the machine learning framework, with the aim of building artificial intelligent systems capable of autonomously learning to perform actions useful for the user. As a result, they are more suitable for investigating *instrumental conditioning* phenomena but less adequate to explain Pavlovian ones (Dayan and Balleine, 2002; O'Reilly, Frank, Hazy, and Watz, 2007).

From the scientific point of view, the available empirical knowledge indicates that the basal ganglia represent the main neural substrate of the S-R associations acquired through instrumental conditioning (Bar-Gad, Morris, and Bergman, 2003; Barto, 1995; Yin and Knowlton, 2006), while amygdala represents the main neural substrate where the associations acquired through Pavlovian conditioning

are stored (Baxter and Murray, 2002; Cardinal, Parkinson, Hall, and Everitt, 2002).

A crucial question on classical conditioning regards the nature of the acquired association between the CS and the UR: is this association direct (CS-UR), as Hull (Hull, 1943) suggested, or does it pass through the unconditioned stimuli (CS-US-UR), as Pavlov himself seemed to claim (Pavlov, 1927)? The long-lasting debate on this topic (Lieberman, 1993) seems now to have settled in favor of both hypotheses: there is in fact strong empirical evidence supporting the co-existence of both CS-UR and CS-US associations (Dayan and Balleine, 2002). In particular, the available empirical evidence suggests that CEA stores CS-UR associations, while BLA stores CS-US associations (Cardinal, Parkinson, Hall, and Everitt, 2002; Mannella, Zappacosta, and Baldassarre, 2008). The rest of this section describes our hypotheses on the specific mechanisms that Amg might exploit to implement these two basic functionalities and to modulate them on the basis of the current internal states.

2.4.1 CEA as the locus of US-UR associations

All animals are genetically endowed with a set of basic responses that have a high direct relevance for their survival and reproduction. These responses belong to three classes: (a) internal responses directed to regulate the states of the body of the organism (discussed in section 2.5.1); (b) neuromodulatory responses that influence the general states of the brain or the relative activity of different parts of it (discussed in section 2.5.2); (c) basic behavioral responses (discussed in section 2.5.3). These responses are innately linked to specific stimuli so that when a given stimulus is perceived, the appropriate responses are automatically triggered. For example, when an animal perceives the odour of a predator its heart-rate speeds up (body), its general alertness increases (brain), and its body might freeze (behavior).

In the case of complex animals living in a complex and dynamic world it is not possible for evolution to a-priori associate the appropriate responses to all the possible stimuli that the animals can encounter during life. The solution that evolution found to this problem is endowing animals with a learning system that associates the basic (unconditioned) responses to the (conditioned) stimuli that are systematically experienced in conjunction to (as predecessors of) the relative basic (unconditioned) stimuli. In our view, CEA is the part of the brain that learns and stores these CS-UR associations. In fact, CEA has been shown to be necessary for the acquisition and expression of both aversive and appetitive conditioned reactions (e.g., startle and freezing behaviors in fear conditioning: see Lanuza, Nader, and LeDoux, 2004; Nader, Majidishad, Amorapanth, and LeDoux, 2001; Shi and Davis, 1999; and orienting and approaching behavior in appetitive conditioning: see Hatfield, Han, Conley, Gallagher, and Holland, 1996; Parkinson, Robbins, and Everitt, 2000). For example, Hatfield, Han, Conley, Gallagher, and Holland (1996) showed that CEA lesions impede the capacity of rats to acquire the association between an unconditioned response (orienting) and a conditioned stimulus (light), while lesions of BLA do not affect this capacity.

CEA is able to make these associations thanks to its pattern of connectivity (see figure 2.7). From the efferent side, CEA constitutes the main output gateway of Amg, sending projections to several brain areas that control all three kinds of basic, unlearned responses (affecting the body, the brain, and basic behaviors, see section 2.5). On the afferent side, CEA receives external projections from both the brain areas having information about unconditioned stimuli (i.e. visceral, somatosensory, olfactory and gustatory) and from those having information about conditioned stimuli (i.e. visual, auditory,

polimodal, and associative) (Jolkkonen and Pitkänen, 1998; McDonald, 1998; Sah, Faber, Armentia, and Power, 2003). Furthermore, both these kinds of information arrive to CEA also indirectly, via its afferent projections from BLA, in particular from LA, which constitutes the principal input gateway of the whole Amg: in fact, also LA receives information on both innately relevant and on neutral stimuli required for classical conditioning associations (Maren, 2005; McDonald, 1998; Paré, Quirk, and LeDoux, 2004; Pitkänen, Jolkkonen, and Kemppainen, 2000; Sah, Faber, Armentia, and Power, 2003). CS-UR associations seem to involve both the internal (from BLA) and the external (from the rest of the brain) afferent projections to CEA since LA lesions sometimes impede these associations to take place (Blair, Sotres-Bayon, Moita, and LeDoux, 2005; Lanuza, Nader, and LeDoux, 2004), while in other cases they do not (Hatfield, Han, Conley, Gallagher, and Holland, 1996).

Figure 2.7 provides a schematization of how CR-US associations can take place within CEA through the modification of the afferent connections going from conditioned stimuli (CS), represented both within LA and outside Amg, to the unlearned responses (UR), thanks to the experienced co-occurrence of innately relevant stimuli (US) and such unlearned responses (the scheme is both a simplification and an elaboration of the computational model that we used for simulating experiments on second-order conditioning in normal and BLA lesioned rats, cf. Mannella, Zappacosta, and Baldassarre, 2008, and also section 2.5.2).

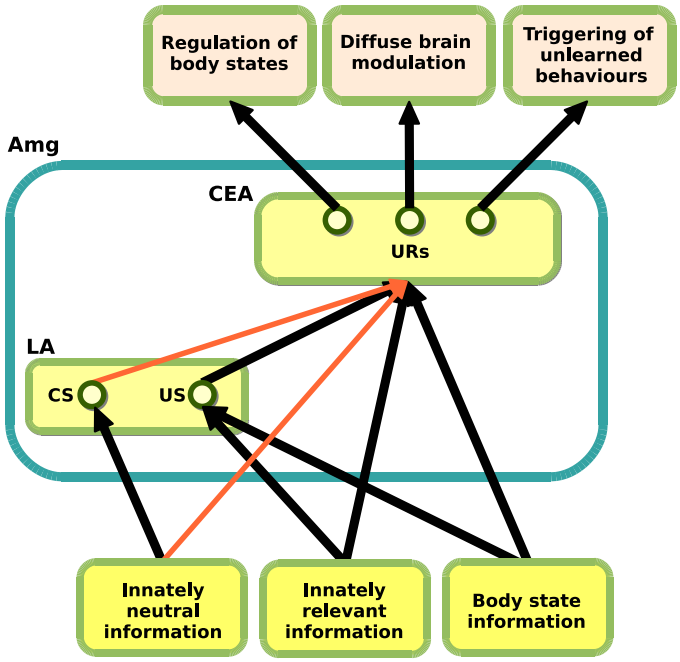


Figure 2.7: *CEA: schematization of the learning of CS-UR associations (thin arrows) on the basis of the pre-existing unlearned US-UR associations (thick arrows). Ancronyms: Amg (amygdala), CEA (central extended amygdala) LA (lateral amygdaloid nucleus).*

2.4.2 BLA as the locus of CS-US associations

Direct CS-UR associations have a clear adaptive advantage, but has two limits. First, among the unconditioned responses that can be triggered by CEA (and hence can be associated to conditioned

stimuli through CS-UR learning) there is not the production of learning signals, for example the production of phasic dopaminergic bursts (Schultz, Dayan, and Montague, 1997; Schultz and Dickinson, 2000) or noradrenergic bursts (Berridge and Waterhouse, 2003), which are usually at the basis of learning. Various facts indicate that this is the case. First, CEA has inhibitory GABAergic efferent connections McDonald (1998), for example reaching VTA producing dopamine (DA) or LC producing norepinephrine (NE). When this type of connections reach target areas having neurons producing neuromodulators, they tend to produce modulatory tonic signals instead of phasic signals as those usually used to trigger learning (cf. section 2.5.2). Second, while CEA lesion disrupts the capacity to show CS-UR associations, it does not disrupt the capacity of conditioned stimuli to be used as reinforcements in second-order conditioning (Hatfield, Han, Conley, Gallagher, and Holland, 1996). Third, the direct association of stimuli with basic reactions would not allow using of conditioned stimuli for influencing more complex cognitive processes. In order to overcome these limits the brain needs a mechanism to *link neutral stimuli to unconditioned stimuli*, so that the presentation of a CS can recall the associated US and both trigger the phasic bursts of neuromodulators driving learning and modulate high-level cognitive processes.

There is plenty of evidence that BLA is the part of the brain that learns and stores CS-US associations. In fact, BLA has been shown to be necessary: (a) for having second-order conditioning phenomena, where a conditioned stimulus (e.g. a light) is used as a secondary reward in extinction (i.e. without first order reward) to condition a second neutral stimulus (e.g. a tone) (Hatfield, Han, Conley, Gallagher, and Holland, 1996): this can rely on BLA glutamatergic connections directly or indirectly targeting neurons of neuromodulators, suitable for causing phasic responses (e.g., see figure 2.8 for DA); (b) for influencing high-level cognitive processes, as described in detail in section 2.6.

BLA is able to make these associations thanks to its connectivity (figure 2.8). As discussed above, LA (which is part of BLA) is the main input gateway of the whole Amg, receiving information both regarding USs (from visceral, gustatory, olfactory, and somatosensory areas) and regarding CSs (from visual, auditory, polimodal and associative areas). Furthermore, the areas of BLA that receive these two kinds of information are reciprocally interconnected, thus permitting the associations between CSs and USs to take place.

Interestingly, the internal connectivity within BLA seems to suggest that the convergence between CSs and USs takes place in two sites organised in sequence: (a) at the level of Lv (which is a part of LA) visceral, somatosensory and gustatory information (USs) converges with auditory and visual information (Maren, 2005; Pitkänen, Stefanacci, Farb, Go, LeDoux, and Amaral, 1995; Romanski, Clugnet, Bordi, and LeDoux, 1993); (b) at the level of BL, information about USs converges with highly integrated polimodal information from hippocampal, cortical associative and cortical prefrontal areas (McDonald, 1998; Pitkänen, Jolkkonen, and Kemppainen, 2000; Price, 2003; Sah, Faber, Armentia, and Power, 2003). This hierarchy in BLA's internal connectivity suggests that USs can be associated with stimuli of different levels of complexity: from the simplest, unimodal stimuli that are typically used in classical conditioning experiments (e.g. lights or tones), to complex objects, context, or places, like in conditioned place preference experiments (Hiroi and White, 1991; McDonald and White, 1993).

Finally, the representations of USs (that can be recalled by associated CSs) can control three different classes of systems thanks to different sets of BLA efferent projections (see figure 2.8):

(a) projections to Hip (McDonald, 1998; Richter-Levin, 2004), NAcc (Cador, Robbins, and Everitt, 1989; Pitkänen, Jolkkonen, and Kemppainen, 2000), and PFC (Rolls, 2000; Sah, Faber, Armentia, and Power, 2003) allow conditioned stimuli to influence cognitive functions (for details, section 2.6); (b) projections to neuromodulatory systems (e.g., VTA and SNpc for DA, reached by BLA through LH and PPT (McDonald, 1998; Pitkänen, Jolkkonen, and Kemppainen, 2000)) allow conditioned stimuli to act as second-order reinforcements by producing the activity bursts that are supposed to drive learning; (c) intra-amygdaloid projections to CEA (Sah, Faber, Armentia, and Power, 2003) allow CSs to trigger all the URs normally triggered by the associated USs.

Figure 2.8 represents a schematization of the BLA functioning: CS-US associations are learned through the modification of the collateral connections that, within BLA link the representations of the unconditioned stimuli (innately linked to their respective unconditioned responses) and the conditioned ones (the scheme is both a simplification and an elaboration of the computational models that we used for simulating real experiments on both second-order conditioning, Mannella, Zappacosta, and Baldassarre, 2008, and devaluation, see Mannella, Mirolli, and Baldassarre, 2007, 2009a, and section 2.6.2).

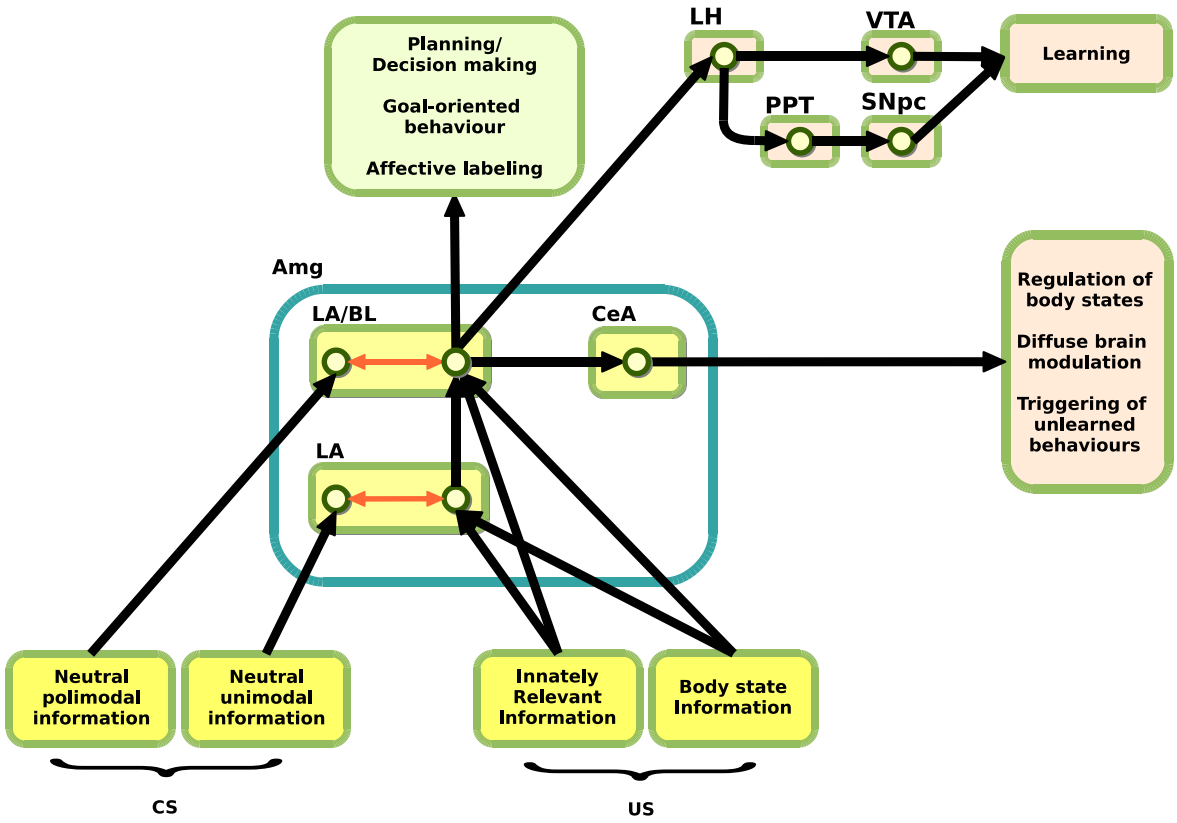


Figure 2.8: *BLA: schematization of the learning of CS-US associations (thin arrows) on the basis of the pre-existing unlearned US-UR associations (thick arrows). Ancronyms: Amg (amygdala), BL (basolateral amygdaloid nucleus), CeA (central amygdaloid nucleus), LH (lateral hypothalamus), PPT (pedunculopontine tegmental nucleus), SNpc (substantia nigra, pars compacta), VTA (ventral tegmental area).*

Finally, it is important to mentioned that all the CS-US associative properties and other properties

discussed so far in relation to BLA are likely implemented by a wider whole system formed by BLA *and* OFC, a region of PFC with which BLA exchanges dense reciprocal interconnections. In this respect, it would be more correct to say that such associative functions are produced by the whole BLA-OFC 'system' instead of by BLA alone. Experiments involving lesioning either BLA or OFC show in fact that it is very difficult to dissociate the functions of BLA and OFC (Pickens, Saddoris, Gallagher, and Holland, 2005; Roesch and Schoenbaum, 2006; Schoenbaum, Saddoris, and Stalnaker, 2007), although recent investigations are starting to show that OFC is more closely involved with working memory processes whereas BLA is more closely related to learning CS-US associations (Schoenbaum et al., 2003, cf. also section 2.6.3).

2.4.3 MEA as the locus of the modulation of USs and URs by internal states

The mechanisms for which an organism can learn to associate innately neutral stimuli to innately specific responses strictly linked to survival and reproduction is really useful only if there is a way to modulate these associations according to the current internal state of the organism. For example, let's consider feeding behavior. Even in presence of the stimuli that have been repeatedly experienced as predictive of food, it is useful to trigger all the responses related to feeding (e.g. orienting, approaching, salivating, etc.) only when the energy level of the organism is low (i.e. when it is hungry), but not when the organism is satiated. Otherwise when encountering a place where there is plenty of food an animal would indefinetely continue to produce feeding related responses, thus risking, for example, to die of thirst. As discussed in section 2.1, regulating which kind of activity an organism pursues in each moment is exactly the function of a well designed motivational system. The regulation of organisms' activities on the basis of its current internal state is what makes organisms' behavior *proactive* (i.e. controlled by their needs) rather than *reactive* (i.e. completely determined by external stimuli).

The need to flexibly and efficiently modulate basic unconditioned responses on the basis of the current state of the body might even represent one of the most important reasons why the CS-US system in BLA has evolved to supplement the probably more basic CS-UR system in CEA. In order to understand why, let's consider the case of food devaluation. There can be two types of devaluation: 'temporary', for example when the organism is satiated, and 'permanent', for example when a food resulted to be toxic (e.g., its ingesting was followed by nausea or stomachache).

Temporary devaluation could in principle be faced even with only a CS-UR system: if the current state of the body modulates directly the unconditioned responses related to feeding (e.g., orienting, approaching, and salivating) then these responses could be temporarily blocked regardless of the stimulus that would trigger it (be it unconditioned or conditioned). But the same solution is not viable for permanent devaluation: an animal cannot permanently block all feeding responses, otherwise it would die of starvation. With a direct CS-UR system, even the solution for which the permanent devaluation is done at the level of the US is satisfying. In fact, such solution could not prevent the execution of preparatory feeding responses eventually triggered by CSs linked to the food, with the result of an inefficient activity directed to an devaluated, even dangerous food.

A CS-US system allows preventing these drawbacks. The reason is that in such a system devaluation can be done at the level of the US. The devaluated US can thus inhibit the URs that are innately associated to it without preventing other stimuli to trigger those responses when neither the devalued US nor the CSs linked to it are present. This would work equally well for both temporary

and permanent devaluation.

While a considerable amount of empirical research has been dedicated to understanding the roles of CEA and BLA in CS-(US)-UR associations, much less work has been done for clarifying the exact neural mechanisms through which unconditioned responses are modulated by the internal states of organisms. The available empirical evidence suggests that this is exactly the function of the third main group of Amg nuclei, namely MEA. First of all, there is evidence that MEA does indeed play a role in regulating the triggering of basic behaviors on the basis of the state of the body: for example, lesions to MEA have been shown to produce disturbances to feeding behavior that lead to obesity (King, 2006), which depends on the incapacity of regulating the triggering of an unconditioned behavior (e.g. feeding) on the basis of the current state of the body (e.g. the level of hunger). Second, MEA has just the right kind of connectivity for supporting this modulatory function (see figure 2.9). In fact, MEA is reciprocally connected to Hyp (in particular VMH, LH, and PVN: De Olmos, Beltramino, and Alheid, 2004; Pitkänen, Jolkkonen, and Kemppainen, 2000), which is the main center of information regarding the current states of the body. Moreover, MEA sends efferent *inhibitory* GABAergic projections to both CEA and BLA (De Olmos, Beltramino, and Alheid, 2004; Pitkänen, Savander, and LeDoux, 1997), and receives excitatory connection from BLA.

Figure 2.9 represents a schematization of how MEA could modulate both US and UR representations in BLA and CEA on the basis of the current body states. Once a representation of US in BLA gets activated (either directly, or via the activation of an associated CS), it tends to activate the respective representation in MEA. If the parts of the brain representing the state of the body (e.g. the Hyp) inform MEA that that US is devalued, the corresponding unit in MEA gets fully activated and can inhibit both the representation of the stimulus in BLA and the representations of the corresponding URs in CEA. For the effectiveness of temporary devaluation (e.g., caused by free feeding and satiation), it is necessary that inhibitory connections from MEA to CEA/BLA have both fast learning and fast forgetting, so that, for example, when the organism is satiated they grow up and inhibit the related US and UR whereas when its hungry they decrease so permitting feeding.

This schema might also explain a last important phenomenon, known as *incentive learning* (Balleine and Dickinson, 1998; Balleine and Killcross, 2006), showed in experiments where the current value of a US (say 'USa') is transferred to another US (say 'USb') only if the animal can experience USb after the devaluation of USa. In fact, if USb is not re-experienced after devaluation of USa the connections from its representation in MEA and the one in BLA (and the relative URs in CEA) has not grown up, thus not inhibiting the responses to the associated CSs. On the other hand, as soon as USb is re-experienced when the animal is in a sated condition, the inhibitory connections immediately grow thus preventing the associated CS to trigger the unconditioned responses.

2.5 The roles of amygdala in emotional processes

According to the framework presented here, Amg has evolved to efficiently associate all the innate responses (URs) that are directly important for organisms' survival and reproduction to innately neutral stimuli (CSs) that are repeatedly experienced as predictors of those stimuli (USs) that trigger those responses. This section illustrates in detail the operation of this fundamental function of Amg with respect to the three classes of unconditioned responses: regulation of body states (section 2.5.1), dif-

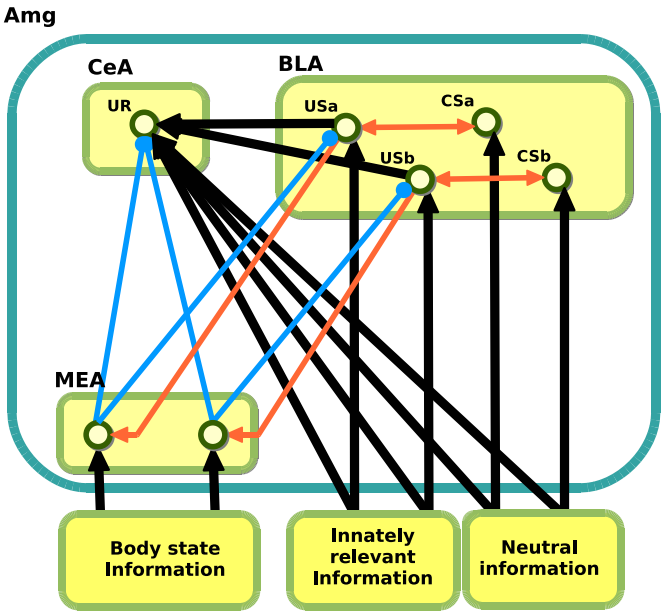


Figure 2.9: *MEA: schematization of the modulation of Pavlovian associations based on internal states, plastic connections (thin arrows) and innate connections (thick arrows). Circle edges denote inhibitory connections whereas arrow edges denote excitatory connections. Ancronyms: Amg (amygdala), BLA (basolateral amygdaloid complex), CeA (central amygdaloid nucleus), MeA (medial amygdaloid nucleus).*

fuse brain modulation (section 2.5.2), and triggering of unlearned behaviors (section 2.5.3). Recall (see section 2.2) that the processes regulating these three kinds of basic responses are here assumed to be essential components of emotions.

2.5.1 Regulation of body states

The regulation of body states based on external events is a fundamental functionality for complex organisms which have a several needs to satisfy. For example, if an organism is going to eat, it will be useful for it to prepare digestion with salivation and an increase of blood flow to the guts. But if a predator suddenly arrives, the same organism has to prepare its body for fighting or flighting, for example by suddenly redirecting the blood flow to the muscles, increasing the heart rate, increasing glucose release, etc.

Thanks to its associative properties, Amg can transfer all these body regulations from stimuli that innately trigger them to stimuli that are learned to be predictors of them. The adaptive advantages rendered by these processes are evident: body states can be regulated in advance with respect to the events that make them useful. These processes are captured in the laboratory by the classical experiments of Pavlov, in which a dog learns to prepare its body to digestion by salivating in advance when it hears a bell that has been systematically associated with the delivery of food.

Many of these body regulations take place via the influence of the ‘autonomic nervous system’ (‘ANS’, working aside the Central Nervous System, ‘CNS’), which includes the sympathetic and parasympathetic nervous systems (‘SNS’ and ‘PSNS’, respectively). The SNS is always active at a basal level (‘sympathetic tone’) and becomes more active during times of stress. With stress the

SNS prepares the body to *fight-or-flight* responses in that it boasts arousal and energy generation and inhibits digestion. In particular, it diverts blood flow away from the gastro-intestinal tract and skin via vasoconstriction, enhances blood flow to skeletal muscles and lungs, dilates bronchioles of lungs, increases heart rate, dilates pupils, inhibits peristalsis (Davis and Whalen, 2001; Iversen, Iversen, and Saper, 2000). The PSNS has a complementary function with respect to the SNS: in general, it can be said to prepare the body to a *rest-and-digest* mode of behaviour in that it promotes calm action and digestion. In particular, in absence of salient stimuli and compelling needs PSNS dilates blood vessels leading to the gastro-intestinal tract, constricts the bronchiolar diameter in lungs, diminish heart rate, causes constriction of pupils, stimulates salivary gland secretion, accelerates peristalsis, and cause erection of genitals (Iversen, Iversen, and Saper, 2000).

Amg influences the SNS and the PSNS mainly via CeA (Davis and Whalen, 2001): in particular, through its efferent connections directed to various nuclei of Hyp, mainly LH, PO, and PVN (Jolkkonen and Pitkänen 1998; Knapska, Radwanska, Werka, and Kaczmarek 2007; see figure 2.10), and through efferent connections to the brain-stem and the spinal-cord (Davis and Whalen, 2001). Through the connections to LH, CeA can influence thirst and hunger (that is, the perception of the internal lack of water and food); through the connections to PO it can modulate urination, heart rate, and blood pressure; and through the connections to PVN CeA can influence gastric reflexes, blood pressure, and temperature regulation.

The innervations to PVN are also very important as they allow CeA to control the *hypothalamic-pituitary-adrenal axis*, which, via the Pituitary gland (or 'hypophysis'), has a major role in the regulation of the network of body hormones (Iversen, Iversen, and Saper, 2000). Hence, through this axis, CeA can influence virtually all internal processes, including water retention, blood pressure, temperature regulation, male aggression, uterine contractions and lactation, the production of extrogens, analgesy and metabolism of nutrients (Iversen, Iversen, and Saper, 2000).

2.5.2 Diffuse brain modulation

Like the regulation of body, the regulation of diffuse brain states plays a central role for organisms that have to satisfy several different needs. In fact, the performance of different activities and actions requires the differential involvement of different brain areas and the functioning of such areas with different modes. The modulation of brain activity is accomplished in two ways: (a) indirectly, via the body, through the activation of endocrine glands that release hormones in the blood (hormones regulate both the body and brain states); (b) directly, via the activation of ancient nuclei of neurons that release the four principal neuromodulators: the monoamine *serotonin* (5-HT), and the three catecholamine *dopamine* (DA), *norepinephrine* (NE; also named 'noradrenaline'), and *acetylcholine* (ACh).

The neuromodulators are produced in two main ways, that tend to have different effects on target neurons:

1. *Tonic production* involves a prolonged populational activation of the neuromodulatory neurons, typically induced via their diffused GABAergic disinhibition, which leads to the accumulation of the neuromodulator in the extrasynaptic space. The main effect of tonic production of neuromodulators is the general modulation of the targeted areas.

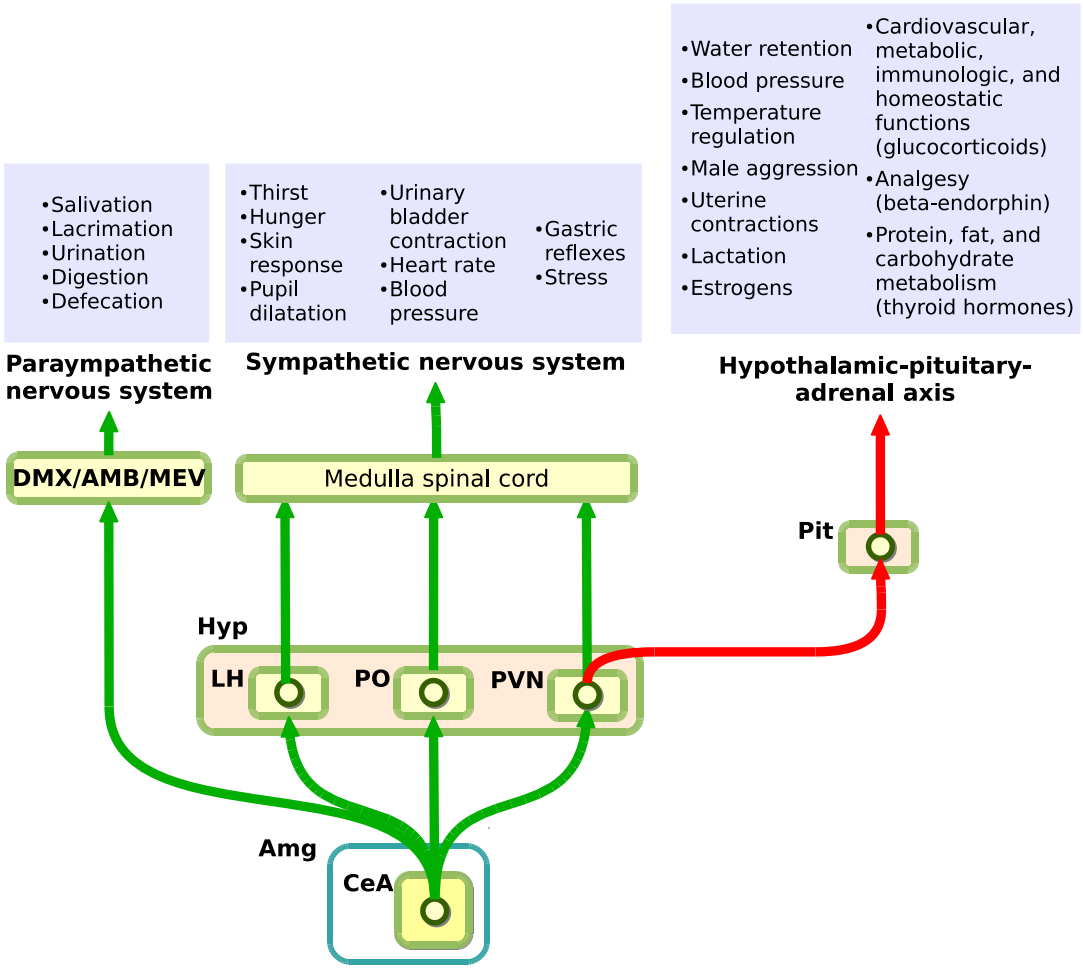


Figure 2.10: *Body states regulation: schematization of how amygdala contributes to regulate the body states via the sympathetic, parasympathetic and hormonal systems. Ancronyms: AMB (nucleus ambiguus), CEA (central extended amygdala), DMX (dorsal motor nucleus of the vagus nerve), Hyp (hypothalamus), LH (lateral hypothalamus), MEV (midbrain trigeminal nucleus), PO (preoptic nucleus of hypothalamus), PVN (paraventricular nucleus of hypothalamus), Pit (pituitary gland).*

2. *Phasic production* involves a high but short activation of the neuromodulatory neurons, typically induced via their glutammaergic direct activation, which leads to the fast but temporary high increase of neuromodulator in the intra-synaptic space. Phasic production of neuromodulators is supposed to have an important effect for learning (see the case of DA, below) or for quick regulation of brain states when speed is paramount (e.g. to face a predator).

Even with respect to the brain modulation, the core function of Amg is based on its capacity to transfer the effects originally associated to stimuli which have been genetically established as salient by evolution (US) to previously neutral stimuli (CS). So, for example, the increased levels of stress and alertness innately associated to the perception of a predator, can be transferred to the type of noises which preceded the attack, or to the sight of the place where the attack took place.

The Amg exerts brain modulations mainly via CEA (Davis and Whalen, 2001) which is connected

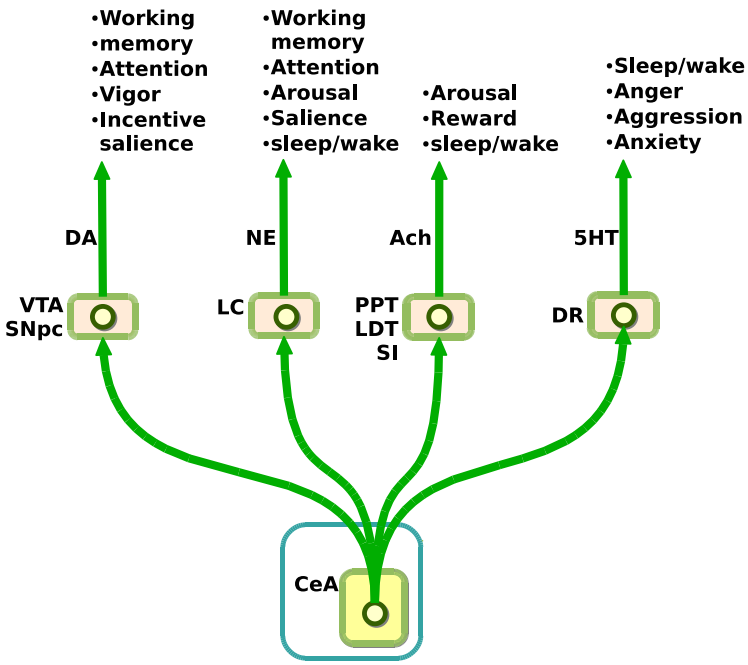


Figure 2.11: *Brain states regulation: schematization of how amygdala contributes to regulate brain states via the diffused action of neuromodulators. Ancronyms: VTA (ventral tegmental area), SNpc (substantia nigra pars compacta), LC (locus coeruleus), PPT (pedunculopontine tegmental nucleus), LDT (laterodorsal tegmental nucleus), SI (substantia innominata), DR (dorsal raphe), ACh (acetylcholine), DA (dopamine), NE (norepinephrine), 5HT (serotonine).*

to the main brain nuclei producing the neuromodulators. One important exception is the modulation by BLA of the burst firing of the dopamine neurons via glutamatergic projections to LH (Petrovich, Holland, and Gallagher 2005; Petrovich, Setlow, Holland, and Gallagher 2002; see also section 2.4.2).

Amg modulates the production of DA by influencing the two main centers of dopaminergic neurons: VTA, which reaches NAcc and PFC (Fudge and Emiliano, 2003; Fudge and Haber, 2000), and SNpc, which sends projections principally to BG, especially DLS and DMS (Han, McMahan, Holland, and Gallagher, 1997; Lee, Groshek, Petrovich, Cantalini, Gallagher, and Holland, 2005). Tonic DA enhances the general level of processing of PFC, so enhancing working memory and attention (Phillips, Vacca, and Ahn, 2008), and, via NAcc, the vigor of performance of selected actions (Floresco, 2007; Niv, Daw, Joel, and Dayan, 2006). Phasic DA signals the positive/negative salience of stimuli which is at the basis of some important learning processes within BG (Schultz, 2002; Surmeier, Ding, Day, Wang, and Shen, 2007) and vmPFC (Otani, Daniel, Roisin, and Crepel, 2003).

Amg modulates the production of NE through LC, which innervates virtually the whole cortex, the BG, Th, Hyp, Hip, Cer, and the spinal cord (Aston Jones and Cohen, 2005; Berridge and Waterhouse, 2003) (note that noradrenergic neurons play an important function also within the sympathetic system; NE is also released as an hormone in the blood by adrenal medulla). NE plays an important function in the regulation of the sleep/wake cycle, and in increasing attention, arousal, and working memory on the basis of the general saliency of stimuli (that is on the basis of their novelty) (Berridge and Waterhouse, 2003).

Amg regulates the production of Ach mainly via PPT, LDT (Knapska, Radwanska, Werka, and

Kaczmarek, 2007; Semba and Fibiger, 1992), and SI (Jolkkonen, Miettinen, Pikkarainen, and Pitkänen, 2002), which innervate the brainstem, Amg, Hip, and PFC. In the central nervous system, Ach is known to modulate the sleep/wake cycle, synaptic plasticity (LTP), general excitability, arousal, and reward (Chen, Nakamura, Kawamura, Takahashi, and Nakahara, 2006) (note that in the peripheral nervous system, Ach is used to activate muscles).

Both directly and via LH and PAG (Bandler, Keay, Floyd, and Price, 2000; Peyron, Petit, Rampon, Jouvet, and Luppi, 1998), Amg regulates the production of 5-HT by the DR, which innervates BG (including NAcc), Th, Hyp, Hip, Amg, and virtually the whole cortex (Barnes and Sharp, 1999). 5-HT modulates mood, anger, aggression, stress, sleep, body temperature, and metabolism (Grahn, Maswood, McQueen, Watkins, and Maier, 1999; Maier and Watkins, 2005; Nelson and Trainor, 2007; Sørensen, Bjorvatn, and Ursin, 2000)(note that 5-HT is also a peripheral signal mediator, in particular within the guts autonomic system).

2.5.3 Triggering of unlearned behaviors

In probably all animals, evolution has led to the emergence of a number of stereotyped unlearned basic behaviours that are triggered when specific stimuli are perceived. For example, these behaviours lead a hungry rat to approach food as soon as this is perceived (e.g. smelt), and, once it is close to the mouth, to ingest it. Similarly, a rat will regularly perform a rearing behaviour directed to looking for predators. In case the rat spots one, it will freeze if the predator is far or startle and then engage in flight or fight behaviors if the predators is close.

Amg plays an important function in the selection of these behaviors. First, it allows the anticipatory execution of these behaviours: that is, in correspondence to previously neutral stimuli that predict the appearence of the stimuli that innately trigger the behaviours. For example, the sight of a landmark previously associated with food might trigger an approaching behaviour directed to it and this might allow obtaining the food, or a particular smell associated with a predator might trigger a startle reflex and then a flight behaviour. Second, it allows triggering the behaviours only in the presence of suitable internal states. For example, a rat can stop executing a feeding behaviour if it becomes satiated, or can decide whether to fight or flight on the basis of its self-perceived internal state.

The Amg exerts a control on unlearned behaviours on the basis of a complex network of connections that CeA has with various nuclei (figure 2.12). So, for example, CeA can trigger freezing, flight or fight behaviours via PAG (Bandler, Keay, Floyd, and Price, 2000; Davis and Whalen, 2001), the startle reflex via NRPC (Davis and Whalen, 2001). Furthermore, Cea might also exploit more indirect mechanisms based on DA to modulate the triggering and execution of feeding, rearing and approaching behaviours. In particular, CeA might enhance feeding behaviours via the dopaminergic modulation of NAccS-VP-LH pathway through VTA (Ahn and Phillips, 2002; Smith and Berridge, 2005; Tindell, Smith, Peciña, Berridge, and Aldridge, 2006; Wyvell and Berridge, 2000). Similarly, rearing seems to be performed on the basis of a striato-cortical loop passing through DLS-PMC-MC and might be modulated by CeA via a DA influence passing through SNpc (Han, McMahan, Holland, and Gallagher, 1997). In the same way, the fundamental behaviour of approaching, which plays a central role in the adaptation of organisms as it allows them to get in contact with the needed resources scattered in the environment, is performed via a second striato-cortical loop involving NAccC and AC, that can be influenced by CeA through DA produced via a connection to VTA (Cardinal,

Parkinson, Marbini, Toner, Bussey, Robbins, and Everitt, 2003; Parkinson, Willoughby, Robbins, and Everitt, 2000). Note how these mechanisms differ from direct triggering, e.g. performed via PAG, as they imply an existing tendency to perform the behaviour, e.g. to move towards a seen object, and a modulation of Amg of this tendency, performed on the basis of the the VTA-NAcc or the nigro-striatal dopaminergic connections. This difference seems to be a general feature when comparing the neural substrates of fear-conditioning responses with those of appetitive-conditioning responses.

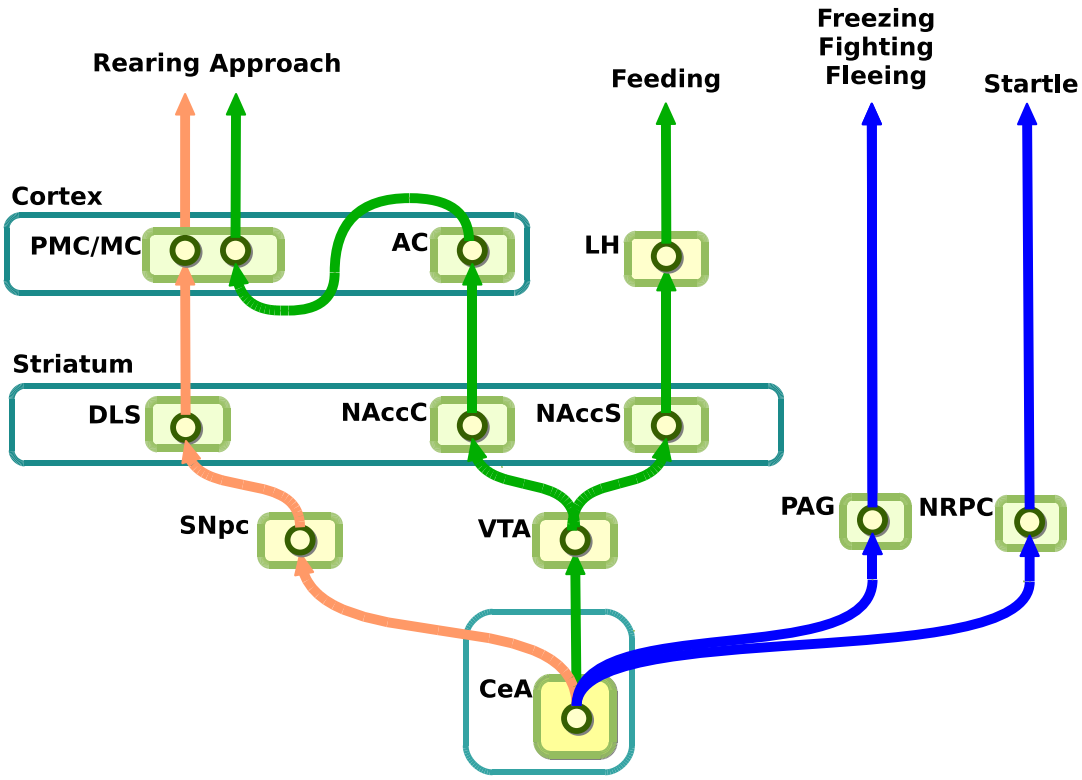


Figure 2.12: *Triggering of unlearned behaviors: schematization of how amygdala contributes to the triggering of unlearned behaviors via different sub-cortical and cortical brain areas. Acronyms: AC (anterior cingulate cortex), CeA (central amygdaloid nucleus), DLS (dorsolateral striatum), LH (lateral hypothalamus), MC (motor cortex), NAccC (nucleus accumbens core), NAccS (nucleus accumbens shell), NRPC (nucleus reticularis pontis caudalis), PAG (periaqueductal gray), PMC (premotor cortex), SNpc (substantia nigra, pars compacta), VTA (ventral tegmental area).*

2.6 The roles of amygdala in cognitive processes

Thanks to its capacity to trigger basic emotional responses on the basis of conditioning processes, amygdala also evolved the capacity to act as a link from emotional processes to cognitive ones, thus allowing the development of important new functionalities. In this section we discuss three fundamental new cognitive functions allowed (or improved) by Amg: affective labeling (2.6.1), goal-directed behavior (2.6.2), and planning and decision making (2.6.3).

2.6.1 Affective labeling

One of the most important memory functions of the brain is its capacity to quickly store specific events characterised by unique and arbitrary configurations of objects and events in space. This capability plays a very important role for organisms' survival as it allows them to store important information on the basis of a few experiences or, in extreme cases, even one single experience.

This functionality relies heavily on Hip and its peculiar anatomical and physiological properties. These properties have been specified at a theoretical level in McClelland, McNaughton, and O'Reilly (1995), have been modelled in Alvarez and Squire (1994), and can be summarised as follows (cf. Rolls and Kesner, 2006): (a) Hip has important reciprocal connections with many associative cortical areas (e.g. PFC, IT, PPC) and sub-cortical nuclei (e.g. NAcc and Amg); (b) Hip neurons have massive lateral connectivity; (c) Hip is one of the brain loci where rapid associative learning leading to Long Term Potentiation is strongly present; (d) Hip has been shown to reactivate during sleeping (Eschenko, Ramadan, Mölle, Born, and Sara, 2008; McClelland, McNaughton, and O'Reilly, 1995).

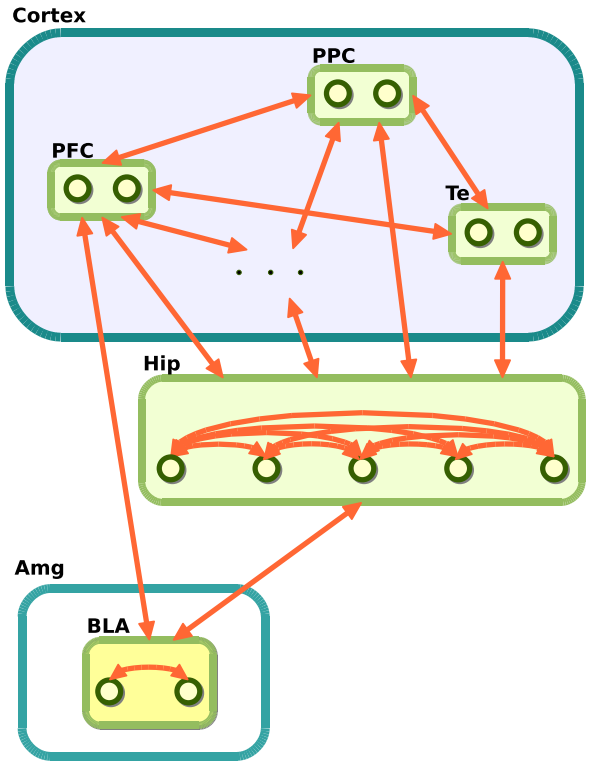


Figure 2.13: *Affective labeling: schematization of how amygdala 'tags' memories stored in Hip and cortex through emotional evaluations of stimuli and episodes. Plastic connections and innate connections are respectively indicated with gray and black arrows. Acronyms: Amg (amygdala), BLA (basolateral amygdaloid complex), Hip (hippocpampus), PFC (prefrontal cortex), PPC (posterior parietal cortex), Te (temporal cortex).*

On this basis, McClelland, McNaughton, and O'Reilly (1995) suggested that Hip plays an important role in *episodic memory* acquisition and consolidation. In particular, Hip can rapidly form neural associations between sub-clusters of its neurons and several different multimodal activation patterns that take place in different brain areas at the same time. Consequently, Hip can form representations

of any arbitrary polimodal pattern existing at a certain time. According to the authors, the later spontaneous reactivation (e.g. during sleep) of Hip clusters cause the reactivation of the corresponding patterns located in the various areas of the brain and so allows the formation of *direct connections* between the neurons corresponding to them (consolidation). With consolidation, probably the patterns initially stored in Hip fade away (within days/months), but they might even continue to be stored, at least in part, within Hip (Rolls and Kesner, 2006). The slow speed and intermixed order with which consolidation of different experiences takes place allows the formation of semantic long-term memories having a high degree of generalisation since the areas innervated by Hip can capture the *common structure* existing behind different experienced episodes.

BLA plays at least two important roles in the formation of episodic memories within Hip. First of all, it is important that only the experiences with high relevance for survival and reproduction are stored. As pivot of emotions, Amg contains the information needed to decide which events, either with a positive or negative valence, might have a high biological relevance, and so deserve to be stored in Hip. This allows the Amg to drive the Hip to store or not the various experiences. This first function is likely played by the Amg on the basis of its influence on neuromodulators (cf. section 2.5.2), which play a very important role in Hip learning.

A second, more direct, function played by Amg in episodic and semantic memories is based on the massive reciprocal connections it forms with Hip. These connections allow Amg to furnish Hip with the current emotional context, which is to be integrated with the other cognitive components that form the episodes to be stored. With the consolidation process driven by Hip, the information stored within Amg gets directly associated with other cortical and sub-cortical areas with which it is directly or indirectly (especially via PFC) connected. In this way, such information comes to play the role of a sort of emotional tag associated with the stored episodes. This association allows two fundamental processes to take place. First, it allows emotional reactions taking place withing Amg to contribute to the recall of memories stored within the Hip or within the areas with which the current affective context has been associated during consolidation (LaBar and Cabeza, 2006; Phelps, 2004). Second, when Hip, or the areas linked between them during consolidation, recall particular episodes, their association with Amg allow them to reactivate the emotional valence of such episodes within Amg itself so as to: (a) trigger the suitable brain and body regulations suitable for such episode (this might be important if the current situation is similar to the recalled episode), and (b) to get a feedback from Amg (via reciprocal connections) on the biological saliency of the recalled episode (this might be important when Hip processes excert a direct or indirect influence on action).

So, for example, if in the past a rat has experienced an attack from a predator after having perceived a particular noise in a certain location of the environment, a later sight of such place might trigger the recall of the noise (and hence trigger a useful priming effect which would facilitate its detection) and this might activate the related negative effects of the attack within Amg (thanks to a CS-US association). In turn, this reactivation might not only trigger a suitable regulation of body (e.g., making the body ready for flight or fight) and brain (e.g., enhancing the general arousal of cognitive processes such as attention), but also contribute to recall further useful memories within Hip (or within the areas connected during consolidation), for example the paths followed to reach a safe place after the attack.

2.6.2 Goal-directed behaviors

As mentioned in section 2.4, instrumental (or 'operant') learning represents, aside Pavlovian learning, one of the two fundamental processes underling individual learning in complex organisms (Domjan, 2006; Skinner, 1938; Thorndike, 1911). As we have seen, instrumental learning allows organisms to form stable S-R associations between stimuli and responses, initially produced by chance, if the latter allow obtaining rewards or avoiding punishments. The acquisition of S-R associations is well captured by reinforcement learning models (Barto, 1995; Sutton and Barto, 1998). Such S-R associations are acquired only with prolonged training and form efficient but rather rigid 'habits' that are performed independently of the current value of the pursued outcome (e.g. food, see below).

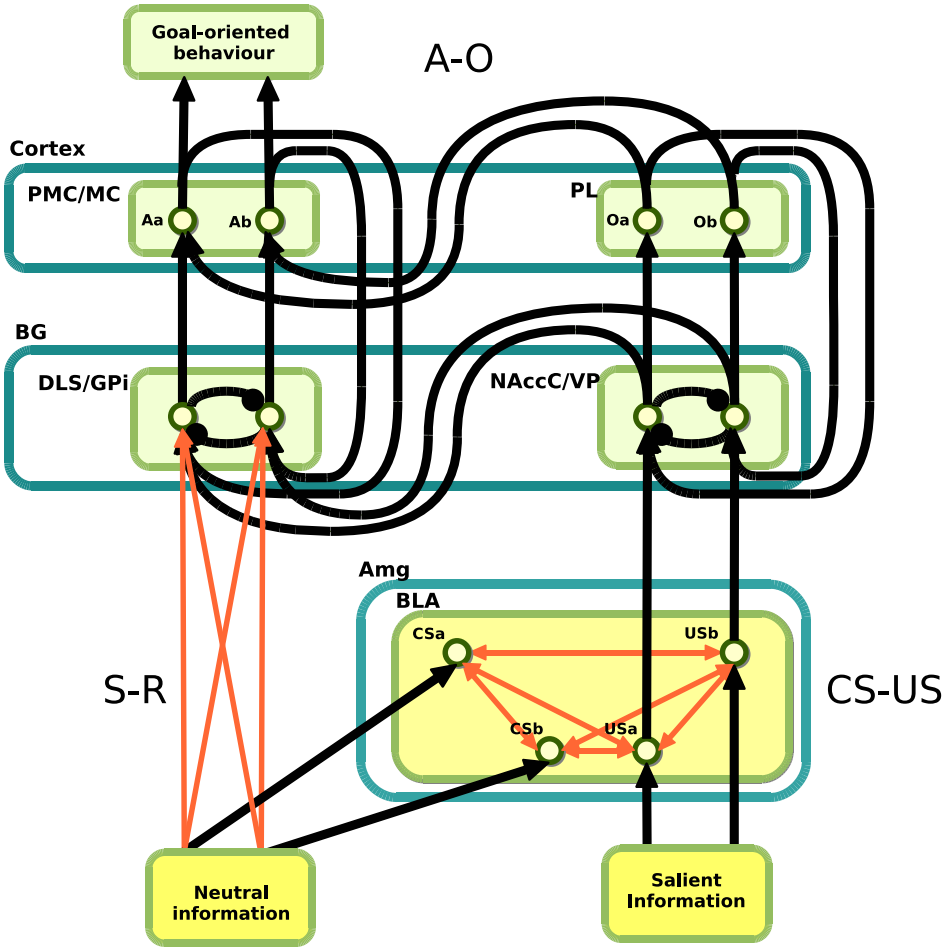


Figure 2.14: *Goal-directed behaviour: schematization of how amygdala contributes to bias the selection of instrumentally-acquired stimuli. Plastic connections and innate connections are respectively indicated with thin and thick arrows. Circle edges denote inhibitory connections whilst arrow edges denote excitatory connections. Acronyms: Amg (amygdala), BG (basal ganglia), BLA (basolateral amygdaloid complex), DLS (dorsolateral striatum), GPi (globus pallidus, internal segment), MC (motor cortex), NAccC (nucleus accumbens core), PL (prelimbic cortex), PMC (premotor cortex), VP (ventral pallidum).*

Basal ganglia are considered to be the main locus where operant conditioning associations take

place. In particular, the macro-loop formed by DLS with cortex (in particular PMC/MC) via the Th, is known to play a fundamental role in both the acquisition and the expression of S-R associations (Yin and Knowlton, 2006).

Other portions of BG, in particular the two macro striato-thalamo-cortical loops DMS-PFC/PC and NAcc/vmPFC, play a rather different role. In particular, the DMS-PFC/PC loop plays an important role in the initial phases of learning, when the S-R habits are not yet formed (Yin and Knowlton, 2006). The last NAcc/vmPFC loop, which has a higher interest for this review due to the strong projections it gets from the BLA, is very important for the guidance of behaviour (the 'action' – A) on the basis of the current value of its ultimate goal (the 'outcome' – O), for example the current potential value of the pursued food.

The behaviors modulated on the basis of A-O associations have a typical *goal-directed* nature in that they lead to select actions on the basis of a relation which goes *from the outcomes of actions to the actions themselves* and so *inverts* the temporal and causal relationship existing between them (actions cause the achievement of outcomes). In this respect, the goal-directed modulation of the selection of instrumental behaviours considered here represents a first fundamental departure from the S-R scheme which reaches its maximum degree of development with planning and complex decision making processes described in section 2.6.3.

The functionality accomplished by the A-O mechanisms has a fundamental adaptive role. Indeed, it allows internal body states and needs, via the MEA-BLA pathway, to bias the selection of different habits which might be triggered in a given situation. For example, as elegantly captured by the instrumental devaluation experiments mentioned below, if a rat can perform two or more different instrumentally acquired actions (i.e. different habits) to achieve two or more different outcomes (e.g. different resources satisfying different needs), the current configuration of its internal states and needs will allow it to decide on the fly which habit to select, without the need of re-learning. These mechanisms add a great flexibility to the rigid habits and are fundamental to allow animals to select between different courses of actions at each time on the basis of the needs related to the homeostatic regulation of body states (cf. section 2.5.1).

Balleine and Dickinson (1998) boosted a whole new research agenda directed to study A-O behaviours and to contrast them to S-R behaviours traditionally studied within the behaviourist approach. These authors give an operational definition of goal-directed behaviours based on two classes of experiments:

1. Goal-directed behaviours are sensitive to the degradation of the *A-O contingency*, that is the strength of the causal relationship existing between the performance of an action and the achievement of the related outcome (the contingency strength is measured on the basis of the relation existing between the probabilities of obtaining the outcome with and without the action). If this contingency is degraded, for example by delivering the outcomes non-contingently to the action, the considered action is performed less intensely or frequently in cases of goal-directed behaviours but not in case of habits (Balleine and Dickinson, 1998).

2. Goal-directed behaviours are *immediately* (i.e. without the need of re-training) sensitive to manipulations of the value that the organism assigns to the outcome (Balleine and Dickinson, 1998). For example, in a typical instrumental devaluation experiment (Balleine et al. (2003))

one of two foods ('Food1' and 'Food2') previously used to form two instrumental associations, 'PressLever1-Food1' and 'PressLever2-Food2' , is 'devalued' by letting the rat to freely access it (e.g. Food1). In a successive test, when exposed to both Lever1 and Lever2 the rat has a strong bias to select Lever2 associated to the currently non-devalued food (Food2).

.

Figure 2.14 presents a diagram which allows illustrating the most important mechanisms involved in goal-directed behaviour, for example the instrumental devaluation experiment illustrated above (the model has first been published in Mannella et al. (2007, 2009a) and is now being further refined). The components reported on the left of the figure mimic the reinforcement-learning based acquisition of S-R behaviours on the basis of prolonged training. This allows the rat to acquire the two habits 'PressLever1-Food1' and 'PressLever2-Food2' in the first phase of the experiment when the two levers are presented separately. When in the last test phase, which takes place after one of the two foods has been devalued (say Food1), the two levers are presented together, the rat exhibits a strong tendency to select on of the two levers (Lever2) thanks to the biasing effects that Amg exerts on the habits.

This important effect is played by Amg on the basis of three fundamental mechanisms:

1. While in the first phase of the experiment the rat instrumentally acquires the S-R habits, the creation of the contingency between the observation of each lever and the following reception of the corresponding food allows Amg to form the two CS-US contingencies Lever1-Food1 and Lever2-Food2.

2. In the devaluation phase, when the rat can freely accessible one of the two foods, implies that the rat gets satiated for such food (e.g. Food1): in the model, the resulting internal state inhibits the corresponding representation of food (US) within BLA.

3. As a consequence, when in the last phase the rat is exposed to the two levers, only one of the two representations the levers (CSs) within Amg can activate the corresponding US representation and so exert an influence on the corresponding S-R habits via NAccC (Corbit, Muir, and Balleine, 2001). Importantly, the actual biasing effects of Amg on habits, which instantiate the A-O associations within the model, is performed both via the striato-nigro-striatal connections ('dopamine spirals', Haber, 2003) and via PFC (in particular PL, Corbit and Balleine, 2003).

These mechanisms capture the essence on how Amg can increase the adaptation of animals by adding an important flexibility to the selection of the instrumentally acquired habits thanks to its capabilities of forming CS-US associations (within BLA) and of modulating their activation on the basis of internal states (detected by MEA).

2.6.3 Planning and decision making

Planning and decision making involving complex decisions can be he considered the hallmark of complex cognition in mammals. Planning consists in the mental generation of trajectories of future possible behaviours and states that can be achieved with them (Dagher, Owen, Boecker, and Brooks, 2001), whereas complex decision making involves the selection of a given alternative versus other alternatives on the basis of a complex calculation of the associated consequences, their values and

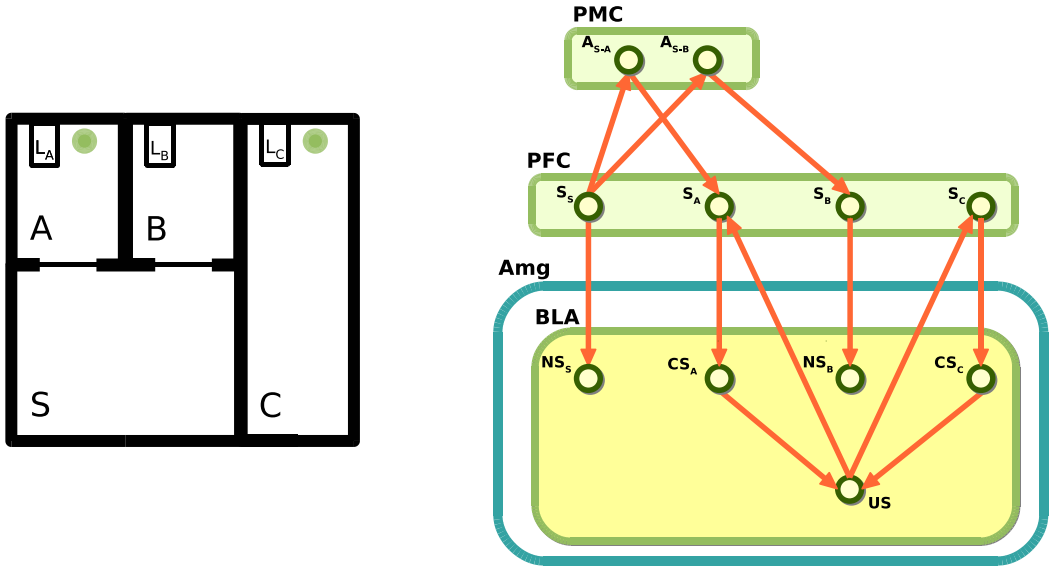


Figure 2.15: *Planning: schematization of how amygdala contributes to planning by furnishing values to anticipated states. Left: an hypothetical task involving planning. Right: a possible model to tackle the task. All connections in the model are learned in the various phases of the process (see text). A, B, C, S states corresponding to being in different chambers; L: lever; NS: neutral stimulus; CS: conditioned stimulus; S: state; A: action leading from one state to another. Acronyms: Amg (amygdala), BLA (basolateral amygdaloid complex), PFC (prefrontal cortex), PMC (premotor cortex).*

their probability of actually happening (Bechara, Tranel, Damasio, and Damasio, 1996). The core functionality underlying planning and complex decision making is the capacity of producing internal images of future states decoupled from current percepts but instead related to percepts which might be experienced in the future as a consequence of own actions (Miller and Cohen, 2001).

The development of these skills has reached the maximum level of sophistication in humans due to the evolution of an exceptionally extended and complex PFC cortex. In this respect, the PFC represents the brain area governing behaviour at the maximum level of abstraction and involving the longest future time scope (Miller and Cohen, 2001).

Amygdala plays a fundamental role in planning and complex decision making: it furnishes the *values to the imagined possible future states* so as to allow animals to select the suitable course of action which has the highest probability of achieving important biological advantages while reducing physical damages and costs to a minimum (Kringelbach and Rolls, 2004). In this respect, imagine a rat which has previously experienced food in a certain place in the environment but, on the way to it, it smells the presence of a predator, for example a cat. In this case the rat has to decide if continuing to move towards the food place, or, say, to detour and reach the food by following a much longer way, on the basis of the chances of encountering the predator along the straight fast path (e.g., as signalled by the intensity of the predator odour), the anticipated energy spent in the detour, the knowledge of the path to be followed in the case of the detour, the information from the body related to the current level of hunger and the residual amount of energy available, etc.

A possible experiment capturing this type of situation, which is inspired by the response-preconditioning experiments , is one where a rat is set in a chamber S from which it can access either one of two chambers A and B by entering their gated entrance (the gates prevent the rat from seeing the

inside of A and B from S). Now assume that each of the two chambers A and B contains a different distinguishable lever, respectively LeverA and LeverB, and that the rat is left free to explore this environment for a prolonged time. Also assume that the rat can experience a further level, LeverC, in a chamber C which does not communicate with none of S, A and B. Also assume that in a second training phase the same rat experiences a LeverA-food association in A, a LeverC-food association in C, and a Lever-no food condition in B. Now, if in a third phase the rat is set in S, one might expect that the rat would exhibit the tendency to enter A more than B as in A it would expect to see LeverA associated with food.

Figure 2.15 shows a sketch of a model which might be implemented to reproduce the role of Amg and PFC (in particular OFC, IL and PL) in the described experiment. The figure shows that the experience of the chambers A and B in the first phase might allow the rat to form associations between the representations of such chambers and the representation of chamber S in PFC, linked by suitable representations of the actions leading from S to A (A_{S-A}) and from S to B (A_{S-B}), and between the representations of S, A, B and C and the corresponding representations in Amg (CS_S, CS_A,CS_B, and CS_C). When in the third phase the rat is set in S, the PFC representation of S should cause an anticipatory reactivation of the internal representations of A and B (but not C). This activation would cause the corresponding representations within Amg to be activated (CSA and CSB): as only CS_A in turn reactivates the representation of food US, this might produce a feedback signal to the PFC representation A and this would strengthen the activation of A_{S-A} in comparison of A_{S-B}. This might cause trigger the performance of A_{S-A} with the aid of downstream areas, for example involving the striato-cortical loops involving in the specification of the detailed movements of action (in particular, DMS and DLS, and PMC and MC) with which PFC is directly and indirectly connected.

Amg and PFC, in particular BLA and OFC, play also a key role in complex decision making. This is for example demonstrated by the experiments of Winstanley, Theobald, Cardinal, and Robbins (2004) (cf. also Mobini, Body, Ho, Bradshaw, Szabadi, Deakin, and Anderson, 2002) who trained rats with two levers, one producing a small immediate amount of food and the second one producing a larger but delayed amount of food. Interestingly, rats which received a post-training lesion of BLA exhibited a higher tendency to select the immediate-food lever in comparison to shams whereas rats which received a post-training lesion of OFC exhibited a higher tendency to select the delayed-food lever. Although an agreed explanation of these experiments is not yet available (cf. Mobini, Body, Ho, Bradshaw, Szabadi, Deakin, and Anderson, 2002; Schoenbaum, Saddoris, and Stalnaker, 2007; Winstanley, Theobald, Cardinal, and Robbins, 2004), they show that OFC and BLA play a pivoting role in complex decision making.

It is interesting to relate these data on rats with those on complex decision making in humans. For example, humans with a damaged Amg/OFC/vmPFC perform poorly in tasks requiring the integration of information about imagined gains and costs in the financial domain. Bechara et al. (1994) developed a task, the Iowa Gambling Task (IGT), directed to study this kind of dysfunction. In the test subjects are allowed to choose an item from two decks of cards, one which produces low monetary gains with a high probability and one which produces high gains but also very high cost with a low probability so that the net gain this deck is lower than the net gain of the first deck. Whereas control subjects learn to choose cards from the first deck and also exhibit an increased skin conductance before selecting a card from the second deck, patients with damage to either the Amg or the vmPFC tend to prefer the

high-risk deck and also fail to show an increased skin conductance.

Bechara et al. (1996) have proposed that Amg and vmPFC play a central role in guiding choices in the IGT. The idea is that PFC generates possible future events (e.g., financial gains or damages) and these are evaluated by the PFC-Amg re-entrant loops (cf. the aforementioned abstract model) thanks to the capacity of Amg of activating the body reactions that would follow from the actual experience of such events. In this respect, these emotional body reactions plays the role of 'somatic markers' of such events that, once propagated back to PFC, support the selection or rejection of the alternative available courses of action.

2.7 Conclusion

Amygdala is a brain system which palys a key role in the affective regulation of body, brain and behaviour. This principled review has presented the general principles which might underly the inner functioning of amygdala, and has illustrated how they allow amygdala to play a key role within various sub-systems of brain. The review has first shown how Amygdala is capable of integrating information from internal states, innately relevant stimuli and innately neutral stimuli on the basis of three core functioning mechanisms: (a) amygdala associates the triggering of important basic behaviours (e.g. approaching and salivation), innately triggered by biologically relevant stimuli (e.g. food), to neutral stimuli (e.g., the sight of a landmark signaling the presence of food in the environment); (b) amygdala associates representations of neutral stimuli (e.g., of the landmark) to representations of biologically relevant stimuli (e.g., the food) so transferring all the properties of the latter ones (e.g., the capacity of triggering basic behaviours) to the former ones; (c) amygdala modulates such associations on the basis of internal states (e.g., satiation can stop the triggering of salivation caused by the sight of a landmark predicting food, or it can inhibit the re-activation of the internal representation of the food itself).

These mechanisms allow amygdala to play an important role in the regulation of three emotional processes fundamental for adaptation: (a) the regulation of body states; (b) the regulation of brain states via the principal neuromodulators; (c) the triggering of a number of basic behaviours relevant for organism's survival and reproduction. Moreover, they allow amygdala to contribute to exert an important emotional influence of three important high-level cognitive processes: (a) 'labeling' memories with emotional valence of stimuli and episodes; (b) biasing the selection of instrumentally-learned habits on the basis of the current valence of their ultimate goals; (c) furnishing the current value of stimuli and events to the processes of planning and complex decision making.

Both the overall picture and the specific claims proposed in this study have been developed by trying to fulfil two main constraints: on one hand, they have been rooted in the currently available empirical knowledge, and on the other hand, they have been developed on the basis of the functional/adaptivistic stance, and the goal of isolating general principles, typical of artificial-life and adaptive-behavior research. As a result, some of the ideas presented can be considered as acquired knowledge in the field of affective neuroscience, others constitute original hypotheses, well supported by available empirical knowledge, and finally some others (hopefully a minority) represent less supported hypotheses that may turn out to be just wrong from an empirical point of view.

In this respect, the authors are aware of the tentative nature of some of the ideas illustrated, but

nevertheless they decided to present them as the goal of the article was to contribute to build a coherent and biologically-constrained picture of the functioning of several brain sub-systems where amygdala plays a central role for the organization of adaptive behavior.

Hopefully, in this way, on one hand this review contributes to foster more theoretically oriented research within affective neuroscience, and on the other hand it contributes to produce more structured and informed research based on the simulation of the motivational and emotional aspects of adaptive behavior.

Published as: Mannella, Francesco, & Baldassarre, Gianluca. 2009. Navigation via pavlovian conditioning: A robotic bio-constrained model of autoshaping in rats. In: (ed), Ninth international conference on epigenetic robotics: Modeling cognitive development in robotic systems.

Chapter 3

Navigation via Pavlovian Conditioning: A Robotic Bio-Constrained Model of Autoshaping in Rats

Abstract

Within the autonomous robotics literature, bio-inspired models of navigation in organisms (e.g. rats) usually rely on instrumental conditioning processes based on the learning of associations between places in the environment and navigation actions leading to rewarded goal places. This paper presents a neural-network model capable of solving navigation tasks on the basis of Pavlovian conditioning processes which allow transferring innate approaching behaviours from biologically salient stimuli (e.g., food) to neutral stimuli (e.g., a landmark seen from far away and close to the food) ('autoshaping'). The overall architecture and functioning of the model is biologically constrained on the basis of relevant neuroscientific anatomical and physiological knowledge on amygdala, nucleus accumbens, and ventral tegmental area. The model is tested with a simulated robotic rat engaged in autoshaping and devaluation experiments. The results show that, although the model allows solving only simple navigation tasks, it produces fast learning and a flexible sensitivity of behaviour to internal states typical of Pavlovian processes. The model is also important for the investigation of adaptive behaviour in general as it clarifies the nature of some core mechanisms which play a key role in several forms of learning.

3.1 Introduction

Navigation is a fundamental adaptive behaviour which allows organisms to displace in space so to get in contact with resources scattered in the environment and use them to increase their survival and reproduction chances. For this reason, the brain machinery emerged during evolution to subserve navigation behaviours is rather sophisticated and based on multiple systems. Most models of animal navigation proposed within autonomous robotic literature are based on instrumental processes (for some reviews, see Filliat and Meyer, 2003a,b; Trullier et al., 1997). Instrumental processes allow organisms to form associations between stimuli and actions on the basis of the resulting reinforcing outcomes (Domjan, 2006). Some of the most influential models use reinforcement-learning algorithms (e.g., based on the Temporal Difference rule, Sutton and Barto, 1998) to form, via a *prologued* training, associations between places and the actions directed to achieve rewarded places. Those of these models which are more strongly biologically constrained assume that places are represented in 'place cells' of hippocampus (HIP) (O'Keefe et al., 1998) and that actions are selected and triggered in a reactive fashion by nucleus accumbens core (NAccC) (Arleo and Gerstner, 2000), or, alternatively, that actions are triggered in a proactive fashion via prefrontal cortex (PFC) (Martinet et al., 2009).

The important processes involving complex spatial elaborations performed by HIP, NAccC and PFC has led to overlook some processes underlying navigation behaviours which are simpler but also faster and more flexible than instrumental ones. In this respect, a main tenet of the paper is that an important class of these simpler processes are based on Pavlovian conditioning mechanisms. Pavlovian

conditioning (Lieberman, 1993) is an experimental paradigm in which a stereotyped 'unconditioned response' (UR), innately associated with, and triggered by, a biologically salient 'unconditioned stimulus' (US), might become associated with, and so triggered by (so becoming a 'conditioned response', CR), an innately neutral 'conditioned stimulus' (CS), if the CS regularly precedes the US. For example, the UR of salivation, innately triggered by the US of the taste or smell of food, might become associated and triggered by a CS consisting in the sight of food if the CS is repeatedly followed by the US.

Approaching food or conditioned stimuli (e.g., a light) is a typical UR/CR studied in Pavlovian experiments (in this case called 'autoshaping'). Autoshaping mechanisms allow organisms to approach (CR) a neutral stimulus (CS) if this has been regularly paired with an appetitive stimulus (US).

Pavlovian mechanisms related to approaching have a great evolutionary advantage. The approaching behaviour is formed by a set of motor routines which involve a complex rhythmic pattern of muscle activations which reduce the spatial distance with the target. In this respect, the advantage rendered by autoshaping mechanisms is that the formation of a *fast-learnable and simple association* between an US (e.g., food) and a CS (e.g., a big landmark close in space to the food and visible from far away) can allow organisms to *rapidly transfer the whole complex target-approaching behaviour* (UR) to the CS.

Pavlovian navigation has also a second important advantage of flexibility: the sensitivity to body states. In fact, internal representations of USs (via the activation of which approaching responses are triggered) can be directly modulated by internal states. For example, the satiation for a particular food (US) can prevent its internal representation from being activated by the activation of a CS associated to it, so stopping costly inuseful URs associated to it (e.g., salivation and approaching).

The main contribution of the paper is the proposal of a model which is a first important step towards a full detailed understanding of Pavlovian-based navigation processes in organisms. This not only has great relevance for neuroscience and psychology, but also is very important for autonomous robotics interested in learning processes for two reasons: (a) it suggests specific mechanisms for implementing quickly-learnable and flexible navigation behaviours; (b) the Pavlovian mechanisms investigated here play a key role in many learning processes, so they have an importance which spans well beyond navigation behaviours (see Mannella et al., 2009b).

The rest of the paper is organised as follows. Section 3.2 illustrates the biological constraints of the model, Section 3.3 the setup of the simulated experiments, and Section 3.4 the model in detail. Section 3.5 presents the results of the autoshaping and devaluation tests, whereas Section 3.6 draws the conclusions.

3.2 Biological Evidence on Pavlovian Navigation Mechanisms

This section presents biological evidence which on one side supports the claim that organisms acquire some kinds of navigation skills based on Pavlovian mechanisms, and on the other side furnishes the anatomical and physiological constraints used to design the architecture and functioning of the model.

A first piece of evidence is that lesions of HIP does not prevent the acquisition and expression of autoshaping behaviours (Parkinson et al., 2000). This is fundamental as rules out that the spatial computations performed by HIP underlie such behaviours.

Another important piece of evidence is related to the basolateral complex of AMG (BLA). BLA is the main locus where CS-US Pavlovian association processes take place (Cardinal et al., 2002; Knapska et al., 2007; McDonald, 1998; Pitkänen et al., 2000). Surprisingly, BLA is not necessary for learning and expression of autoshaping (Parkinson et al., 2000).

BLA, however, is necessary for the flexible modulation of Pavlovian mechanisms based on internal states. An example of this, relevant to this work, is that it is necessary to allow satiation for one food to inhibit not only approaching to such food but also approaching to a CSs associated with it (Blundell et al., 2003). This without the need of relearning.

BLA is also necessary for the functioning of *second order conditioning*, that is conditioning of a neutral stimulus on the basis of the presentation of another neutral stimulus previously associated with it (this can be done 'in extinction', i.e. without presenting the US after the first CS; Cardinal et al., 2002). This might be relevant to extend the model in the future and let it learn to approach a landmark (CS2) if this is followed by another landmark (CS1) previously associated with reward (US).

BLA is also capable of triggering phasic dopamine bursts via its connections with lateral hypothalamus (LH; Pitkänen et al., 2000). These types of dopamine signals are very important for learning.

Another important fact to consider is that the central complex of AMG (CEA) is needed for learning conditioned approach behaviours but not for expressing them (Cardinal et al., 2002). This property seems related to the capacity of CEA of causing a population diffused activation of the ventral tegmental area (VTA) and a consequent production of *tonic dopamine*: this acts as a necessary precondition for phasic dopamine to trigger learning. Tonic dopamine is also at the basis of *vigor* of actions, that is of the mechanisms for which the intensity and frequency of execution of actions can increase due to expectation of appetitive stimuli (cf. Niv et al., 2006).

A further important piece of evidence is that the ventral part of the striato-cortical system (Kandel et al., 2000) is needed to learn and express conditioned approach behaviours. In particular, lesions of the basal-ganglia and cortical components of such loops, namely respectively the nucleus accumbens core (NAccC; Cardinal et al., 2002) and anterior anterior cingulate cortex (ACC; Cardinal et al., 2002, 2003) prevent both learning and expression of conditioned approach.

3.3 The Simulated Rat, the Maze, and the Tasks

The robot used to test the model is a robotic rat ('ICEAsim') developed within the EU funded project ICEA on the basis of the physics 3D simulator Webots^{TM}. The model was written in Matlab^{TM} (Webots has an interface for Matlab code). The numerical integration of the equations of the model is performed with the Euler method and an integration time step of 0.05 (also used for the 3D simulator). The robotic setup used to test the model is shown in Figure 3.1 and it is now briefly described.

The training and test environment is composed by a grey-walled Y maze (only the two upper arms of it were used: the lower arm will be used in future work). Each upper arm contains a landmark, respectively red and blue for the two arms, which the rat can see from far away, and a rectangular food dispenser, which the rat can see only from the middle of the arm onward. The two food dispensers contain food A and food B respectively. When the rat touches a food dispenser it receives a rewarding signal corresponding to the ingestion of the food.

The simulated rat is a two-wheel robot equipped with various sensors. Among these, the tests

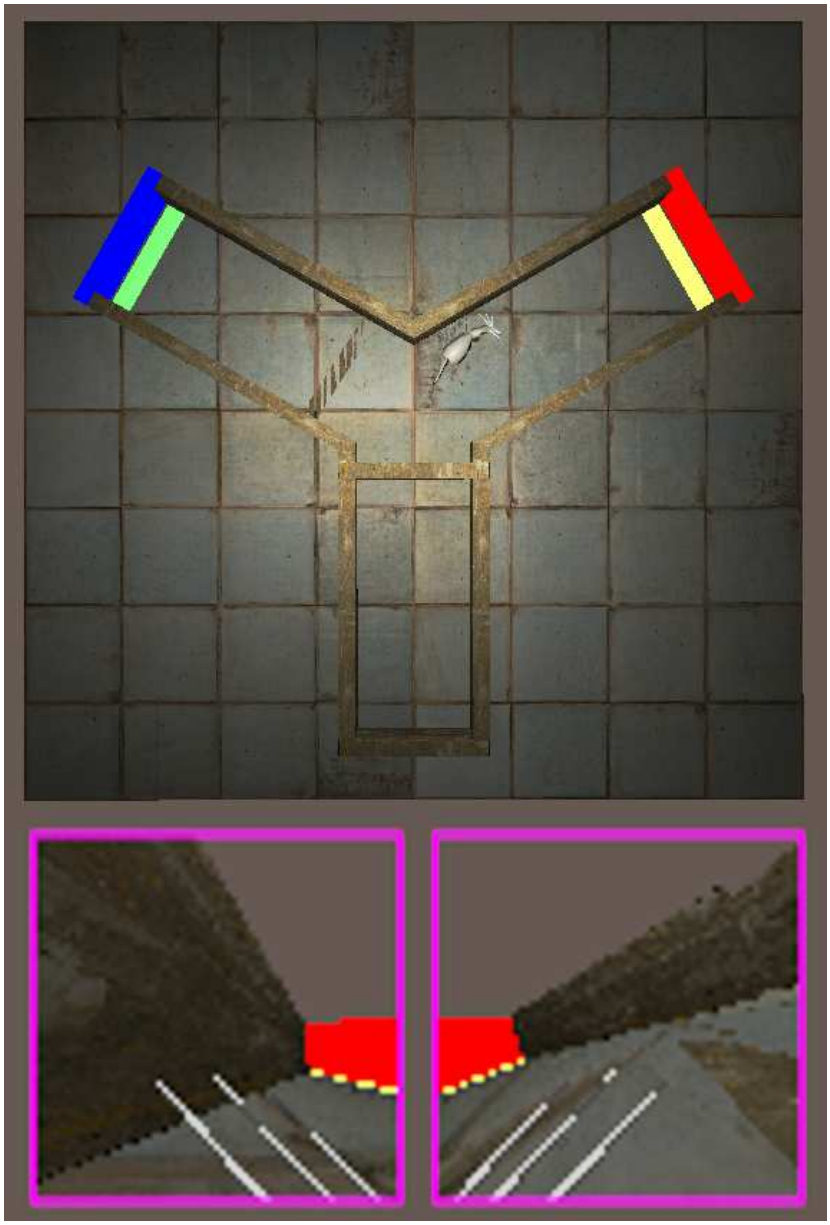


Figure 3.1: *Top: The simulated Y maze and robot. Bottom: The left and right retina images perceived by the rat while positioned as indicated in the top graph.*

reported here use two cameras (furnishing a panoramic 300 degrees view) and the whisker sensors. The rat uses the cameras to detect the red and blue landmarks and the two green and yellow food dispensers. Suitably tuned pre-processing colour filters allow the system to perceive stimuli as binary signals. Landmarks are seen from far away, for example from the crossing of the Y maze, but only when positioned in the frontal zone of the two retinas (approximately within 90^0). Also the food dispensers are visible only if within the frontal zone, but their visibility is limited to positions within a half-arm distance. The rat is also endowed with two binary sensors which detect the ingestion of respectively food A or B, and with two binary *internal* sensors respectively encoding satiety for either

food A or B.

The rat also uses the whiskers, activated with one if bent beyond a certain threshold and zero otherwise, to detect contacts with obstacles. The whiskers are used to control a low-level hardwired 'obstacle avoidance routine' which 'overwrites' all other actions and leads the rat away from obstacles.

The actuators of the rat are two motors which can independently control the speed of the two wheels. The system controls such speed by selecting one of three hardwired routines: 'turn left' and 'turn right', which lead the robot to respectively turn anticlockwise or clockwise on the spot, and 'go ahead' which leads the robot to move forward. If none of these routines is selected and active, the speed of wheels is set to zero. A further 'consummatory routine', mimicking eating, is triggered when the rat is on a dispenser and perceives the related US.

The rat undergoes three training/testing phases:

1. *Pre-training phase.* In this phase, the rat is first trained for 2 mins, divided in trials, in the food-B maze arm without the landmark and blocked with a wall at the other end; then it is trained in a similar condition in the food-A arm. Trials terminate either after 20 secs or when the rat ingests the food. In this phase the rat learns to associate the seen foods (CSs) with the ingested foods (USs).
2. *Training phase.* This phase lasts 2 mins, divided in trials as in the first phase, and involves the two upper arms. In this phase the rat learns to associate the landmarks (CSs) with the seen foods (CSs) and the ingested foods (USs).
3. *Devaluation phase.* This phase is composed of three sub-phases of 4 mins each: one with both fully-valued foods, one with the devalued food A, and one with the devalued food B. Each sub-phase is divided in trials as in the other two phases. In this phase the learning coefficients were set to zero to collect more controlled data. This phase allows testing if the rat has a tendency to explore more extensively the maze arm where the non-devalued food is located.

3.4 The model

This section uses the following conventions: bold capital letters ($\mathbf{X}$) represent matrices, bold small letters ($\mathbf{x}$) represent vectors and small letters (x) represent scalars. The notation $[x]^+$ means that the positive part of x is considered, while the notation $[x]^-$ means that the negative part of x is considered. The function $\phi(x, \theta)$ returns 1 if $x > \theta$, 0 otherwise. Note that each unit activation is here assumed to represent the firing rate of a population of neurons reached by a similar input pattern.

Figure 3.2 shows the architecture of the model based on three main components: (a) the AMG: this is responsible for implementing the stimuli associations of Pavlovian conditioning; (b) the striatocortical system formed by the ventral basal ganglia (VBG: these are a set of nuclei formed by the NAccC, the subtalamic nucleus, STN, and the susbstantia nigra pars reticulata, SNpr) the dorsomedial thalamus (DM) and the ACC: this is responsible for selecting the actions to execute; (c) the dopaminergic system formed by LH and VTA: this modulates both the learning processes and the speed of selection and duration of execution of actions (this is the model correspondent of vigor of actions, see Section 3.2).

With the exception the units of AMG (see Section 3.4.1), all the of the model are leaky integrators

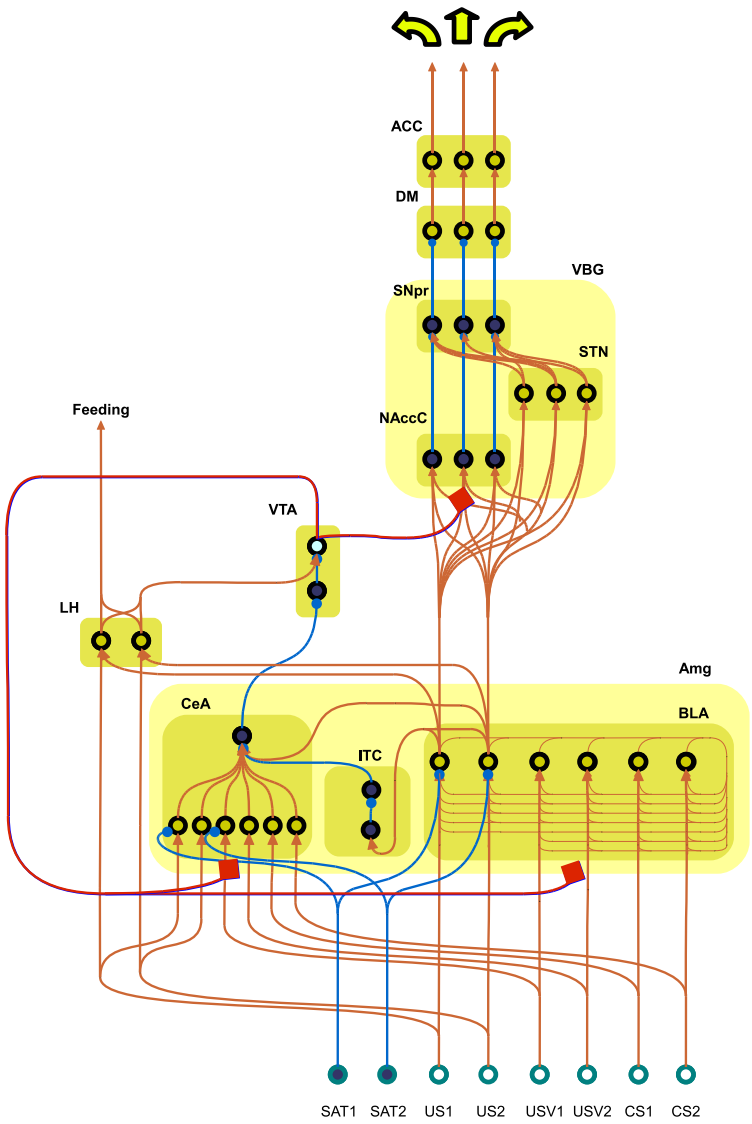


Figure 3.2: *The architecture of the model.*

as described in Amari (1977):

$$\begin{aligned} \tau \dot{u}_i &= -u_i + \kappa_u I + \sum_j w_{ij} \cdot v_j \\ v_i &= [\tanh[u_i]]^+ \end{aligned} \tag{3.1}$$

where u_i and v_i are respectively the potential and the activation of unit i, I is the input signal from either the external environment or the body, κ_u is a multiplying coefficient, and w_{ij} is the weight of an afferent connection from another unit j.

3.4.1 The Amygdala, an CS-CR and CS-US Associator

This section first describes the general functioning and learning of AMG units and then describes the specific functions of BLA and CEA.

BLA and CEA are each formed by six input units which receive one-to-one input signals from the six external input units of the model: two encoding visual conditioned stimuli, two encoding the two seen foods, and two encoding the taste of ingested food. Two additional internal input units of the model, respectively encoding the satiation for the two foods, send strong one-to-one inhibitory signals to the two units of BLA and CEA encoding the two food tastes. Another group of units (intercalated nuclei, ITC) serve as a disinhibitory interface between BLA and CEA (see Paré et al., 2004)

The units of BLA and CEA (denoted with $\mathbf{bla}$ and $\mathbf{cea}$) are different from the other units, in particular each one activates in correspondence to stimuli onset and then fades away (many single neurons in brain have this property). For each AMG unit, this onset-detection function is achieved on the basis of two leaky integrators, o_{in} and o_{out}:

$$\begin{aligned} \tau_1 \dot{o}_{in} &= -\, o_{in} + I \\ \tau_2 \dot{o}_{out} &= -\, o_{out} + [I - o_{in}]^+ \end{aligned} \tag{3.2}$$

This kind of activation is needed to allow the internal connections of BLA and CEA to be updated on the basis of a 'differential Hebb rule' (Mannella et al., 2007; Porr and Wörgötter, 2003). This rule captures the temporal correlation (or 'apparent causality') existing in incoming input patterns. In particular, if one has two units with two reciprocal connections, the rule tends to increase the weight of the connection which goes from the first unit to the second unit, and at the same time tends to decrease the weight that goes from the second unit to the first unit, if the first unit tends to be activated within a certain time window before the second unit. In detail, the learning rule works as follows. First the leaky traces of the derivatives of the activation of the onset units are computed:

$$\tau_{tr} \dot{tr} = -\, tr + \kappa_{tr} \cdot \dot{o}_{out} \tag{3.3}$$

where κ_{tr} is a multiplying factor.

Then a difference in the sign of the traces of the presynaptic and postsynaptic unit determines the amount of the increment of the weights:

$$\begin{aligned}\Delta w_{ij} = \\ &\eta \cdot \phi\left(da, \theta_{da}\right) \cdot \left(da - \theta_{da}\right) \cdot \\ &\left(\left[\dot{tr}_i\right]^- \cdot \left[\dot{tr}_j\right]^+ - \left[\dot{tr}_i\right]^+ \cdot \left[\dot{tr}_j\right]^-\right) \cdot \\ &\left(\theta_{w_{ij}} - w_{ij}\right)\end{aligned} \tag{3.4}$$

where $\theta_{w_{ij}}$ is a threshold near which the weights saturate, da is the dopaminergic level and θ_{da} is the threshold of dopamine level above which learning take place.

BLA. When connections between the AMG visual stimuli units and food-taste units are strengthened via the differential Hebb rule illustrated above (stimulus-stimulus associations), the former ones acquire the ability to activated the output unit in the same way as done by USs.

One of the output responses of BLA consists in triggering, via LH, the activation of VTA output units: this leads to a phasic dopaminergic signal underlying learning (see Section 3.4.3).

A second output reaches NAccC: this has the function of biasing the selection of actions taking place within VBG.

A last output reaches CEA, and allows BLA processes to excert control on the output of CEA.

As mentioned above, BLA US units are also reached by internal signals about satiety. Through these connections the activity of these units can be modulated by the rat internal states, for example here they can be suppressed by satiation. In this way, the US can dynamically change its motivational value. This property is also transferred to CSs if they have been associated to USs within AMG.

CEA. CEA has six input units and one output unit controlling VTA dopaminergic processes. All internal connections are trained with the differential Hebb rule mentioned above, with the exception of those carrying the information about the USs which are fixed ('innate'). This learning process allows the formation of CS-CR associations (stimulus-response associations).

CEA component triggers the dopaminergic system through a disinhibition of the internal population of VTA. Thus this mechanism is able to maintain tonic dopaminergic efflux upon baseline through time. This dopamine is not sufficient to trigger learning within NAccC but at the same time it is necessary to allow the BLA signal to VTA (via LH) to cause dopamine-based learning (see 3.4.3). Moreover, tonic dopamine acts as a multiplier of signals from BLA to NAccC, so implementing a 'vigor' function (see Section 3.2 and 3.4.2).

The internal signals related to satiety modulate the US input units of CEA similarly to what happens in the BLA, and so allow the modulation of its output.

The CEA receives input not only from external stimuli, but also from the BLA. This allows BLA to have access to the output of CEA (dopamine in this case).

3.4.2 The Striatocortical System

The VBG component is a simplified implementation of the basal ganglia 'GPR' model proposed by Gurney et al. (2001a,b). We implemented a three channel version of the model consisting of the basal ganglia 'direct pathway' (from NAccC to SNpr) and 'indirect pathway' (STN to SNpr; cf. Kandel et al., 2000). The three channels activate respectively the 'turn-left', 'go ahead', and 'turn right'

routines (see Section 3.3). As in the GPR model, the input to NAccC is amplified by dopamine:

$$\begin{aligned}\tau_{naccc}\dot{naccc}_i = &-naccc_i + \\ &\sum_j \left[w_{bla_j \to naccc_i} \cdot bla_j\right] \cdot \\ &(bl_{naccc} + w_{da \to naccc} \cdot da)\end{aligned} \tag{3.5}$$

where bla_j is the jth output unit of BLA and $_{bla_j \to naccc_i}$ is its connection weight to $naccc_j$, bl_{naccc} and $w_{da \to naccc}$ are respectively a baseline and a multiplication coefficient of the amplification effects of dopamine on input. The amplification effects of dopamine are very important as they are the means through which CEA can cause approaching in the absence of BLA.

Another important aspect of VBG is that the input signal it receives from BLA is affected by noise. This noise is generated in the form of a random number, uniformly drawn in [0, 1] with a probability of 0.05 at each step of the simulation, added to each VBG input signal received by BLA. The connections from BLA to NAccC are trained on the basis of the following Hebb rule modulated by dopamine:

$$\begin{aligned}&\Delta w_{bla_i \to naccc_j} = \\ &\eta_{bla \to naccc} \cdot \\ &(\phi\left[da, \theta_{da}\right] \cdot (da - \theta_{da})) \cdot \\ &(\phi\left[naccc_i, \theta_{naccc}\right] \cdot naccc_j) \cdot bla_j \cdot \\ &\left(\theta_{bla \to naccc} - w_{bla_i \to naccc_j}\right)\end{aligned} \tag{3.6}$$

where $\eta_{bla \to naccc}$ is a learning rate, θ_{naccc} is a learning threshold for the activation of NAccC units, and $\theta_{bla \to naccc}$ is a threshold for saturating the weights. Note that in this learning rule the information related to $naccc_j$ should be brought to the NAccC units by ACC-NAccC backward connections not explicitly simulated in the model.

3.4.3 The Dopamine System

The dopaminergic activity in the model depends on the LH-VTA system. VTA is formed by one input and one output unit. The input unit is activated by CEA and inhibits the output unit. The output unit receives also an excitatory input from LH and produces as output the dopaminergic signals. Figure 3.3 shows an example of the overall functioning of VTA. The first graph of the figure shows the negative input received by the input unit from CEA. The second graph shows the excitatory input received by the output unit from LH. The last two rows show respectively the activation of the input and output units. It can be seen that the inhibition of the input unit (caused by CEA) can augment dopaminergic activity but never lead it over a certain threshold, e.g. necessary to trigger learning of the dopamine target areas. Similarly, an excitatory signal (from LH) to the output unit is not sufficient to lead dopamine activity over the threshold when presented alone. This implies that both disinhibition and excitation are needed for the dopamine signal to trigger learning.

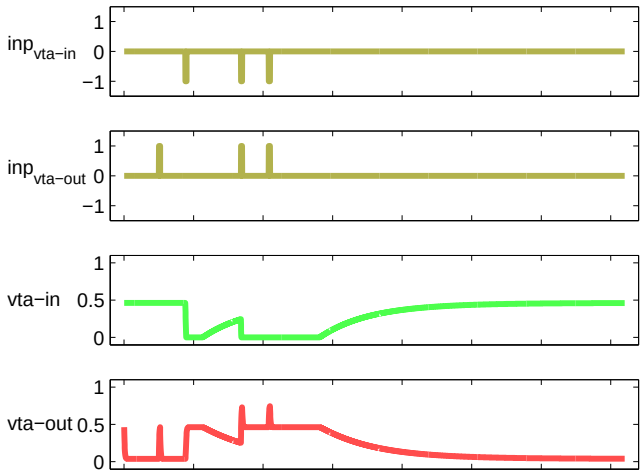


Figure 3.3: *An 'in-vitro' test on the VTA responses.*

3.5 Results

This section reports the outcome of the tests of the rat in the three learning/training phases described in Section 3.3.

During the pre-training phase, the rat initially randomly explores the maze arm where it is by triggering sporadic actions under the effect of noise affecting NAccC. Motion is rather slow due to the low levels of dopamine. Eventually this behaviour leads the rat to step on the food dispenser and eat the food (US). The resulting dopaminergic signal leads CEA to form associations between the CS seen-food unit and the with the output unit triggering the tonic dopamine in VTA, and BLA to form associations between the seen-food unit and the taste-food units. Learning of BLA and CEA lead the system to increase the frequency of selection of actions and the duration of their execution: overall the vigor of the rat seems increased when the rat sees the food.

Figure 3.4 shows the activation of BLA caused by these learning processes. Notice how the activation of the CS units cause an activation of the corresponding US units.

During the training phase, the rat initially explores the environment and speeds up its action when the food becomes in sight. This leads it to rapidly approach the food dispenser while the coloured landmark of the arm is visible. Within CEA, this causes the formation of the associations between the units encoding the seen landmarks and the output unit. In parallel, BLA forms associations between units encoding the seen landmarks and units encoding the sight and the taste of foods.

Figure 3.5 shows the connection weights formed during the pre-training and training phases. Notice how the system has formed positive connection weights from CS units to US units and negative weighs in the opposite direction thanks to the differential Hebb leaning rule.

Figure 3.4: *Example of activation of BLA during the pre-training phase. Notice the activation of the units in correspondence to CSs and USs onsets.*

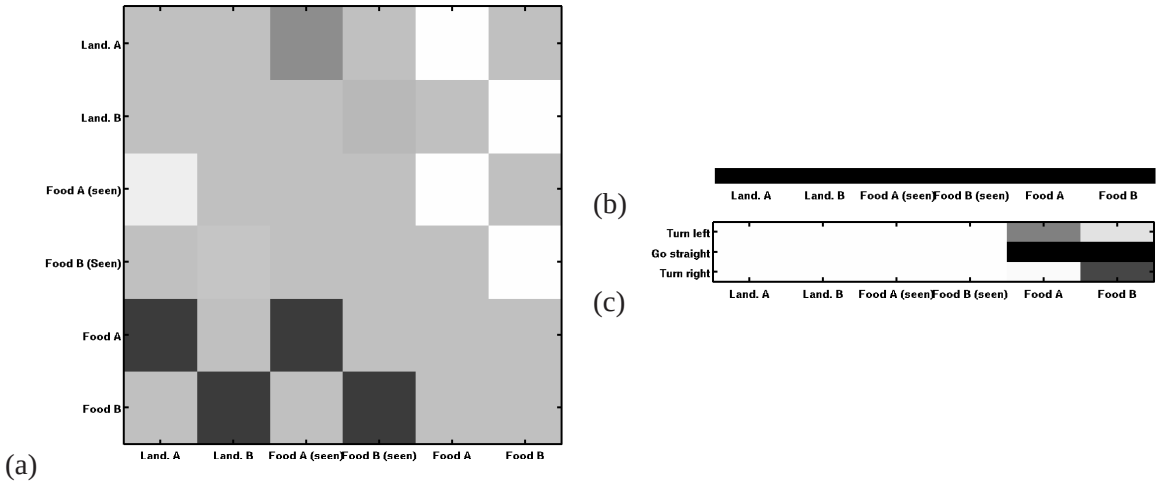


Figure 3.5: *Learned connection weights of the system after the pre-training and training phase. (a) BLA lateral-connection weights. (b) CEA internal connection weights. (c) BLA-NAccc connection weights.*

During the devaluation test the rat exhibits a tendency to move with a higher frequency and vigor towards the non-devalued food and the corresponding landmark (Figure 3.6).

Figure 3.7 shows the activations of the striatocortical system during the devaluation tests. The figure shows how the sight of the landmarks of the non-devalued food causes a higher chance of selection of the go-ahead action, and hence a higher chance of approaching and eating such food.

Interestingly, the intercalated neurons revealed important in this phase as they prevented the CEA from performing its non-selective effects on vigor (the CSs have access to the CEA output unit without being affected by satiety). Indeed, setting low values of the inhibition exerted by these neurons on CEA produced much less pronounced devaluation effects (data non reported).

3.6 Conclusions

This paper presented a bio-constrained model with the goal of giving a coherent overall picture of Pavlovian mechanisms underlying navigation behaviours. The model architecture and functioning was designed with a number of biological constrains in mind, in particular in relation to the specific brain areas which putatively correspond to its constituent parts: (a) one corresponding to amygdala which learns Pavlovian associations between innately-salient and neutral stimuli; (b) another one cor-

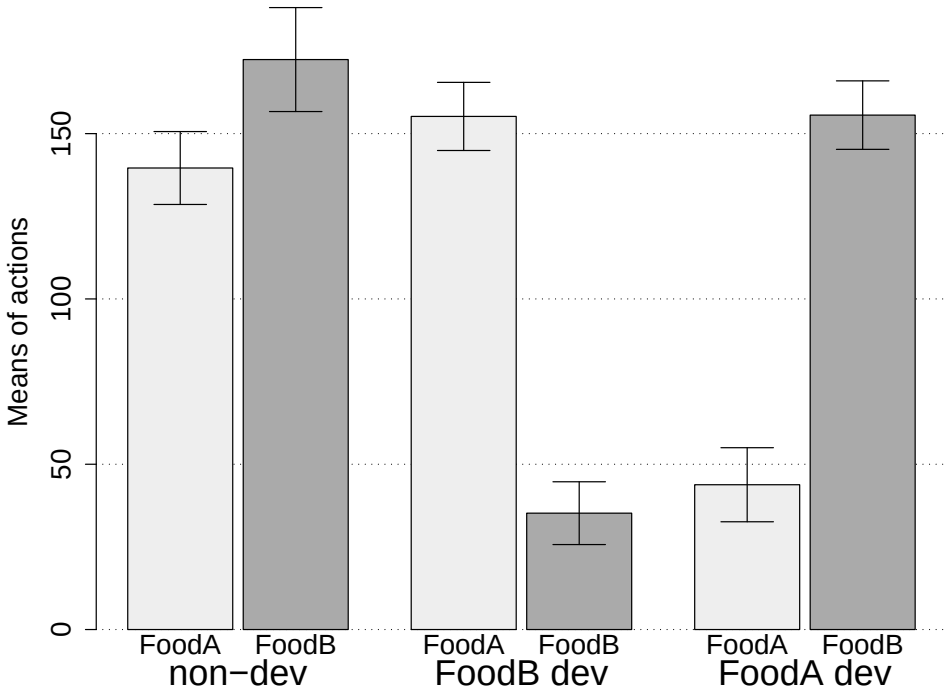


Figure 3.6: *Number of contacts with the (empty) dispenser during the devaluation test in three conditions: with no devaluation, with devaluation of food A, and with devaluation of food B.*

Figure 3.7: *Example of activation in time of the components of the striatocortical system. Note how activations are always biased toward the "go straight" action within NaccC and STN layers as far as no food is satiated. When food B is satiated only the vision of landmark B produces the biasing.*

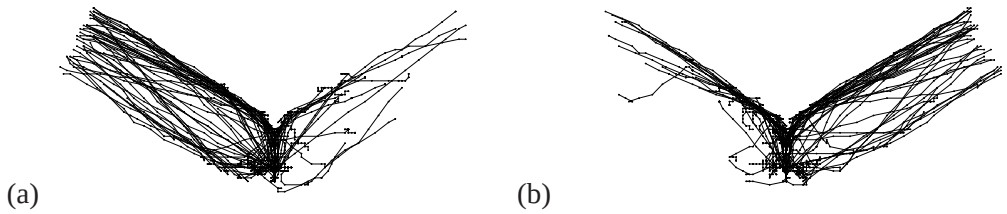


Figure 3.8: *A graphic representation of the movements of the agent during the test phases. (a) movements when food A is devalued. (b) movements when food B is devalued*

responding to nucleus accumbens which selects navigation actions; (c) and a last one corresponding to lateral hypothalamus and ventral tegmental area responsible for generating dopamine learning signals and vigor of action.

Preliminary results of the test of the model with autoshaping and devaluation experiments, run with a simulated rat, show that the behaviour exhibited by the system is qualitatively similar to the behaviour exhibited by real rats in corresponding experiments. Together with the biological constraints imposed to the architecture and functioning of the model, these results furnish a first proof of soundness of the hypotheses incorporated by the model.

The importance of the model for autonomous robotics resides in that the investigated Pavlovian mechanisms, although allow tackling only simple forms of navigation, might be relevant it for at least two reasons. The first is that they allow very fast learning typical of Pavlovian processes. The reason of this is that they rely upon the formation of simple associations between biologically salient stimuli and any relevant neutral stimuli associated with them. These associations allow transferring a complex but readily-available behaviour (basically: 'approach what you are looking') from the former ones to the latter ones.

The second reason is that they add flexibility to behaviours. The reason is that Pavlovian mechanisms allow internal body states to modulate the internal representations of the stimuli through which the navigation behaviours are triggered. For example, such mechanisms allow a rat to navigate towards the region where it expects to find a particular resource (say shelter), but not towards another region where it expects to find a second resource (say a certain kind of food), depending on the current needs for the two resources.

We are aware that much further work needs to be carried out to refine the model so that it can account for all the biological constraints and behavioural evidence reported in Section 3.2, especially in a more quantitative and detailed way with respect to what is done here. However, we believe that the model proposed here is a fundamental starting step towards this purpose.

Published as:Mannella, Francesco, Zappacosta, Stefano, & Baldassarre, Gianluca. 2008. “A computational model of the amygdala nucleis role in second order conditioning”. Pages 321330 of: Tani, Minoru Asada Jun, Hallam, John, & Meyer, Jean-Arcady (eds), *Proceedings of the tenth international conference on simulation of adaptive behavior: From amimals to animals 10*, vol. 5040. Osaka, Japan: Springer.

Chapter 4

The Role of the Amygdala in Second Order Conditioning: a Computational Model

Abstract

The mechanisms underlying learning in classical conditioning experiments play a key role in many learning processes of real organisms. This paper presents a novel computational model that incorporates a biologically plausible hypothesis on the functions that the main nuclei of the amygdala might play in first and second order classical conditioning tasks. The model proposes that in these experiments the first and second order conditioned stimuli (CS) are associated both (a) with the unconditioned stimuli (US) within the basolateral amygdala (BLA), and (b) directly with the unconditioned responses (UR) through the connections linking the lateral amygdala (LA) to the central nucleus of amygdala (CeA). The model, embodied in a simulated robotic rat, is validated by reproducing the results of first and second order conditioning experiments of both sham-lesioned and BLA-lesioned real rats.

4.1 Introduction

Individual learning plays a fundamental role in adaptive behavior of organisms, especially in most sophisticated ones like mammals. Some of the most important mechanisms underlying learning are those studied in classical (Pavlovian) conditioning experiments. In these experiments an animal experiences a systematic association between a neutral stimulus, for example a light (the “conditioned stimulus” or “CS”), and a biologically salient stimulus, for example food (the “unconditioned stimulus” or “US”), to which it tends to react with an innate set of responses appropriate for the US, for example orienting and approaching (the “unconditioned responses” or “UR”). After repeated exposure to couples of CS-US the animal produces the UR even if CS are presented alone.

Since Pavlov’s pioneering works Pavlov (1927), a lot of research has addressed classical conditioning phenomena producing a huge amount of behavioral and neural data Lieberman (1993). However, we still lack a comprehensive theory able to explain the full range of these empirical data. Trying to build detailed *biologically plausible* computational models is a necessary step to overcome this knowledge gap. The current most influential models on classical conditioning, those based on “temporal-difference reward prediction error” Schultz et al. (1997); Sutton and Barto (1998) , suffer of several limitations. The main reason is that they have been developed within the machine learning framework with the aim of building artificial machines capable of autonomously learning to perform actions useful for the user. For this reason they are suitable to investigate *instrumental conditioning* phenomena – a type of associative learning based on stimulus-actions associations – but less adequate to explain Pavlovian phenomena mainly based on stimulus-stimulus associations Dayan and Balleine (2002); O’Reilly et al. (2007).

A crucial question on classical conditioning regards the nature of the acquired association between the CS and the UR: is this association direct (CS-UR), as Pavlov himself seemed to claim Pavlov

(1927), or does it pass through the unconditioned stimuli (CS-US-UR), as Hull Hull (1943) suggested? The long-lasting debate on this topic Lieberman (1993) seems now settled in favor of both hypotheses: in fact, there is now strong empirical evidence supporting the co-existence of both CS-UR and CS-US associations Cardinal et al. (2002); Dayan and Balleine (2002). However, a clear understanding of the neural substrates which might be responsible for these two kinds of associations has yet to be gained. In particular, none of the computational models of classical conditioning based on the temporal-difference mechanisms, nor the models which have been proposed as alternatives to them Balkenius and Morén (1999, 2000); Dayan and Balleine (2002); O'Reilly et al. (2007), make any significant claim on this point.

Within the empirical literature, Cardinal et al. Cardinal et al. (2002) formulated an interesting hypothesis on the neural basis of stimulus-stimulus and stimulus-response Pavlovian associations. According to this hypothesis, the basolateral amygdala (BLA) stores the CS-US associations, whereas the central nucleus of amygdala (CeA) receives or stores the CS-UR associations (CS-UR associations encoded in the cerebellum Thompson et al. (2000) are not considered here).

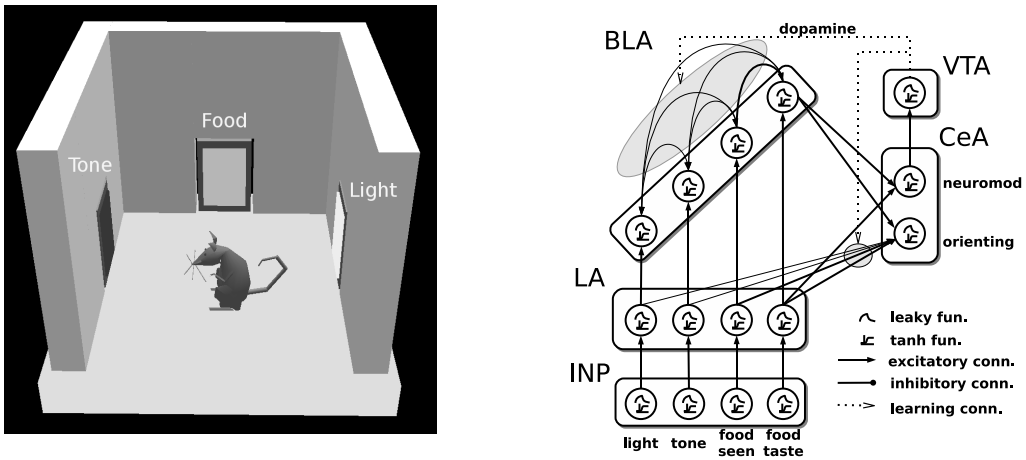
This paper presents an original computational model implementing that general hypothesis. In particular, it represents the first working model specifying the different functions played by the main sub-nuclei of amygdala in classical conditioning. The model, embodied in a simulated robotic rat, is validated by reproducing the results obtained with some first and second order conditioning experiments conducted with sham and BLA-lesioned real rats Hatfield et al. (1996).

Sect. 4.2 presents the target experiment and the simulated experimental setup. Sect. 4.3 describes the model's general functioning and the biological constraints taken into account. The mathematical details of the model are presented in the Appendix. Sect. 4.4 shows the results of the tests of the model and compares them with those obtained with real rats. Finally, Sect. 4.5 concludes the paper.

4.2 The target experiment and the simulated environment

The model is validated by reproducing second-order conditioning experiments on real rats (reported as experiment 1a in Hatfield et al. (1996)). The real experiment was conducted with 19 BLA-lesioned rats and 27 sham-lesioned rats, measuring the behaviours of walking, orienting and "food-cup" (insertion of head in the food dispenser). Namely, in the first phase both groups were trained for 8 sessions lasting $64\ \mathrm{min}$ each to acquire a first order conditioned behaviour. Each session was formed by a sequence of trials. In each trial a $10\ \mathrm{sec}$ light stimulus was presented, followed by the delivery of Noyes pellets (food) in the food dispenser. Recordings showed that both sham and lesioned rats were able to acquire first order conditioned behaviours. In the second phase the same rats were trained for 3 sessions of $64\ \mathrm{min}$ each to acquire a second order conditioned behaviour. A tone stimulus was presented for $10\ \mathrm{sec}$ followed by the light stimulus; every 3 trials a "reminder" of the light-food association was presented. The key result was that only sham rats acquired the second order CS-UR association. In accordance with other empirical evidences (see Cardinal et al. (2002) for a review), these experiments suggest that BLA plays a fundamental role in the formation of the association between the CS and the incentive value of the US, and that this association plays a key role in the acquisition of the CS-UR association in second order conditioning.

The real experiment was simulated through a robotic rat ("ICEAsim") developed within the EU



(a) (b)

Figure 4.1: *(a) A snapshot of the simulator, showing the simulated rat at the centre of the experimental chamber, the food dispenser (at the rat's right hand side), the light panel (behind the rat) and the tone panel (in front of the rat). (b) The architecture of the model: bold and plain arrows indicate innate and trained connections, respectively.*

project ICEA on the basis of the physics 3D simulator Webots^{TM}. The model was written in Matlab^{TM} and was interfaced with ICEAsim through a TCP/IP connection. The robotic setup is shown in Fig. 4.1. The environment is formed by a gray-walled chamber, and the stimuli are expressed by 3 panels (vision is used, as no sound is supported by Webots): food delivering in the dispenser occurs when the green panel turns on, light when the yellow one is on, and tone when the red one is on. When one of those stimuli elicits an orienting response within the controller (see Sect. 4.3), the rat turns toward the panel and then approaches it (these behaviors are hardwired). This behavioural sequence terminates when the rat touches the food-dispenser (that is assumed to correspond to a food-cup behaviour). Although the "degree of embodiment and situatedness" of the setup is rather limited, nevertheless a robotic test was used because in the future we plan to scale the model to more realistic scenarios (for example, the random-lasting time intervals elapsing between rats' orienting and food deliver already started to challenge the robustness of the associative learning algorithms used).

4.3 The model

This section presents a general description of the functioning of the model and the biological constraints that it satisfies, while a detailed mathematical description of it (included all the equations) is reported in the Appendix. A key feature of the model (Fig. 4.1) is the explicit representation of the three major anatomical components of the amygdala Pitkänen et al. (2000): the lateral amygdala (LA), the basolateral amygdala (BLA), and the central nucleus of amygdala (CeA). The model assumes that these components form two functional sub-systems: (1) the LA-CeA sub-system, which forms S-R associations, and (2) the BLA sub-system, which forms S-S associations. Note that in the following "neurons" have to be intended as units whose functioning abstracts the collective functioning of whole assemblies of real neurons.

4.3.1 The Stimulus-Response Associator (LA-CeA).

The LA is the main input of the amygdala system. It receives afferent connections from various sensory and associative areas of cortex, from thalamus, and from deeper regions within the brainstem, and it sends efferent connections both to BLA and to CeA. The model has an input layer (INP) of four leaky neurons (**inp**) activated by four binary sensors (**s**) which encode the presence/absence of four stimuli: light (s_{li}), tone (s_{to}), food sight (s_{fs}) and food taste (s_{ft}) (Eq. (4.1)). LA (**la**) is formed by four leaky neurons receiving one-to-one afferent connections from INP (Eq. (4.2)).

The CeA is one of the main output gates of amygdala. Its efferent connections innervate regions of the brainstem controlling mainly: (1) body and behavioral reactions through the hypothalamus and periaqueductal gray Phelps and LeDoux (2005); (2) the release of basic neuromodulators through the ventral tegmental area (dopamine), the locus coeruleus (norepinephrine), and the raphe nuclei (serotonin) Fudge and Emiliano (2003); Pitkänen et al. (2000); Rosen (2004). These neuromodulators play a fundamental role in learning processes but for simplicity this model considers only dopamine LaLumiere et al. (2005) (in particular it ignores the role that norepinephrine plays in AMG learning Berridge and Waterhouse (2003)). In the model CeA (**cea**) is formed by two leaky neurons, one (cea_{or}) encoding the rat's orienting behavior, and one (cea_{da}) connected to the ventral tegmental area (VTA) to produce the dopamine signal (da) (Eqs. (4.4) and (4.5)).

In the model, all LA neurons are connected to the orienting neuron of CeA (cea_{or}), whereas only the food taste neuron (la_{ft}) is connected to the neuromodulator neuron of CeA (cea_{da}). These connectivity allows stimuli representations of LA to be associated with the orienting behaviour in CeA but not with the dopamine neuromodulation. This is a key assumption to explain why LA-CeA associations can learn first order CS-US associations but not second order ones: conditioned stimuli cannot access the incentive value of rewarding stimuli.

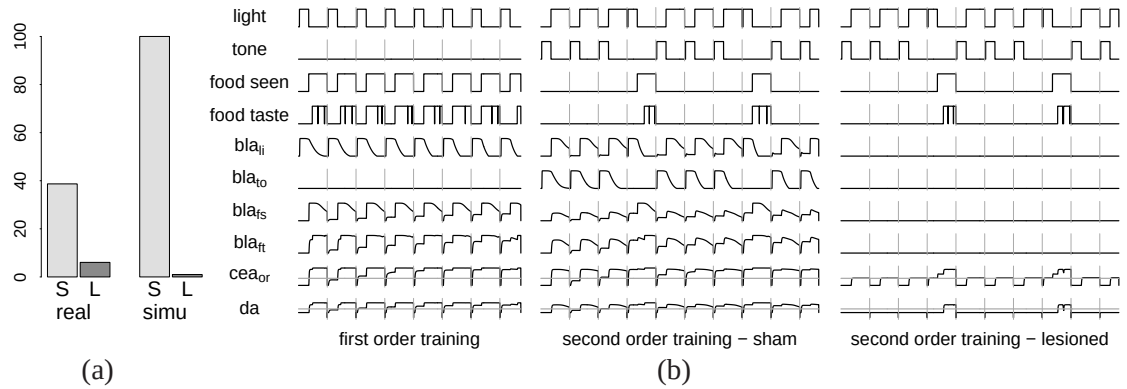
The connections from LA to CeA are trained on the basis of a Hebb rule. In particular, the strengthening of connections takes place in the presence of three conditions (Eq. (4.6)): (1) a high value of the trace of the LA activation onset (**la_tr**): the use of the *onset* makes learning happen only when LA neurons' activation precedes CeA neurons' activation, while the use of the *trace* allows overcoming the time gap between CS and UR; (2) a high activation of CeA neurons (cea_{or} and da); (3) a dopamine level (da) over its threshold (th_{da}).

4.3.2 The Stimulus-Stimulus Associator (BL).

The BLA has afferent connections from LA and efferent connections to CeA Rolls (2000); Saddoris et al. (2005). BLA is also interconnected with the orbitofrontal cortex and hippocampus, and sends efferent connections to the nucleus accumbens: all these connections are ignored here (see Mannella et al. (2007) for a model where BLA-nucleus accumbens connections play a key role).

In the model, BLA (**bla**) is formed by four leaky units which receive one-to-one connections from LA (**la**) and have all-to-all lateral connections (Eq. (4.7)). Only the neuron encoding food taste (bla_{ft}) is connected to CeA neurons. This implies that all neurons of BLA representing stimuli different from the US (bla_{ft}) can exert effects on the CeA output neurons only via lateral stimulus-stimulus connections with the BLA's US neuron.

Learning of BLA lateral connections is based on a time-dependent Hebb algorithm. The key



(a) (b)

Figure 4.2: *(a) Percentage of orienting behaviours of sham (S) and lesioned (L) rats in response to the tone after second order conditioning: data from real (first two bars) and simulated rats (last two bars). (b) Stimuli, activations of key neurons, and dopamine release in 3 conditions: first-order and second-order conditioning phases of a sham rat (first and second block, respectively), and second-order conditioning phase of a BLA-lesioned rat. Trials are separated by short vertical dotted lines; thresholds (for orienting behavior and dopamine learning) are represented as gray horizontal dotted lines.*

aspect of the algorithm is that it allows both the onset and the offset of BLA neurons preceding the onset of other BLA neurons to increase the connection from the former to the latter, provided that dopamine overcomes its threshold (Eqs. (4.8), (4.9), (4.10)). The sensitivity to the offset of stimuli was necessary due to the long duration of the CS stimuli, see Sect. 4.2 (cf. Mannella et al. (2007) for a simpler version of the algorithm using only the onset of presynaptic neurons).

4.4 Results

Figure 4.2a compares the percentage of times the tone elicits an orienting behaviour in real Hatfield et al. (1996) and simulated rats after the second order conditioning phase. The main result of the experiment has been qualitatively reproduced by the model: in both real and simulated rats a BLA lesion prevents second order conditioning to take place. The analysis of the detailed functioning of the model provides an explanation for this result. Figure 4.2b shows the activations of some key neurons of: (1) a simulated sham rat during the first order conditioning phase with the light-food contingency; (2) the same sham rat during the second order conditioning phase with the noise-light contingency; (3) a simulated BLA-lesioned rat during the second order conditioning phase.

Figure 4.2b, first block, shows the mechanisms underlying first order conditioning in a sham simulated rat. At the beginning of the first trial, the appearance of light activates the light-related BLA neuron (bla_{li}). After a while, the appearance of food activates the food-sight BLA neuron (bla_{fs}). The bla_{fs} pre-activates the BLA food-taste neuron (bla_{ft}) before the rat actually reaches the food thanks to a bla_{fs}-bla_{ft} excitatory connection which is assumed to be learned before the conditioning training (see the Appendix). In turn, the bla_{ft} triggers both the orienting behavior via the orienting CeA neuron (cea_{or}) and the release of dopamine (da) by the VTA via the CeA neuromodulation neuron. The release of above-threshold dopamine triggers the learning of both the connection between the light neuron in LA and the orienting neuron in CeA (implementing the CS-UR association) and the connections linking the light neuron with the food sight and food taste neurons in BLA (implementing the CS-US association). The result is that after a very few trials the bla_{fs} and bla_{ft} neurons start to be pre-activated as soon as the light is perceived. This results in an early activation of both CeA neurons

and, consequently, in an early dopamine release and an early orienting response to the light.

As in the target experiment, during the second order conditioning phase the rats are exposed to sequences of four trials composed by three tone-light presentations and one light-food "reminder". Thanks to the CS-US BLA association acquired during the first phase, in sham rats (Fig. 4.2b, second block) the presentation of light immediately triggers both orienting behavior and dopamine release. This ability of light to trigger dopamine release permits the acquisition of the second-order association between the tone and the URs (orienting response and dopamine release) in a manner which is completely analogous to what happens in the first-order conditioning with respect to light.

On the other hand, second order conditioning cannot take place in BLA-lesioned rats (Fig. 4.2b, third block). The reason is that in this case light can trigger only the orienting response via the connection linking the light representation in LA with the orienting neuron in CeA (the direct CS-UR association), but not the dopamine release, which requires the activation of the food-taste representation in either BLA (which is lesioned) or LA (which is activated only when food is effectively eaten). As a result, since synaptic modification depends on dopamine, no learning can takes place during second-order conditioning.

4.5 Conclusions

This paper presented an original computational model of the basic brain mechanisms underlying classical conditioning phenomena. The architecture and functioning of the model was constrained on the basis of neural empirical data on the amygdala. The fundamental assumption underlying the model is that the association between conditioned stimuli (CS) and unconditioned responses (UR) formed in classical conditioning experiments is due to two related but distinct mechanisms: (1) stimulus-stimulus associations (CS-US-UR) involving unconditioned stimuli (US) stored in the BLA; (2) direct stimulus-response associations (CS-UR) stored in the LA-CeA neural pathway.

The model was embedded in a simulated robotic rat and was validated by reproducing the behaviours exhibited by both sham and BLA-lesioned rats in first and second order conditioning experments. In particular, as in real rats, while after training the simulated sham rats react with UR (orienting) to both first and second order CS, BLA-lesioned simulated rats associate UR only to first order CS, but not to second order CS. The model is able to reproduce and explain these results thanks to the fundamental aforementioned assumption. During first order conditioning sham rats acquire both the direct CS-UR and the indirect CS-US-UR association. It is the first order CS-US association within BLA which permits the acquisition of the second order association as it allows the CS to reactivate the appetitive value of the US even when the US is absent. In contrast, BLA-lesioned rats can acquire direct first order CS-UR associations stored in the LA-CeA neural pathway but they cannot acquire the second order association because the first order CS has not access to the appetitive value of the US. To the best of the authors knowledge, this is the first model to propose such a specific computational hypothesis regarding the double association CS-US and CS-UR in classical conditioning.

Notwithstanding its strengths, the model suffers at least two significant limitations. First, the whole behavioral sequence triggered by the activation of the orienting neuron in CeA (orienting, approaching, and food-cup) is fully hard-wired. For this reason, the model cannot reproduce the results

on CeA-lesioned rats which are reported in the same article of the experiment targeted here Hatfield et al. (1996). Second, in contrast to most existing models of classical conditioning Balkenius and Morén (1999); Dayan and Balleine (2002); O'Reilly et al. (2007), the current model does not implement any mechanism for reproducing the exact timing of dopamine release observed in real animals. For this reason the model cannot reproduce another fundamental aspect of classical conditioning, that is extinction (the ability to re-learn not to respond to the CS if it stops to be followed by the US). We are currently working on improved versions of the present model for tackling both these limits.

Appendix: Mathematical details of the model

Throughout the Appendix, τ_x denotes the decay rate of a leaky quantity x, the sub-index $\cdot_p$ denotes the activation potential of the corresponding neuron, symbols $\mathbf{X}$, $\mathbf{x}$, and x are used respectively to denote matrices, vectors and scalars, the function φ is defined as $\varphi[x] = \max[0, x]$ and the function χ as $\chi[x] = 1\ if\ x \geq 0\ else\ \chi = 0$. The values of parameters are listed at the end of the Appendix.

LA-CeA: Functioning and Learning. INP ($\mathbf{inp}$) processes the input signal from sensors $\mathbf{s} = (s_{li},\ s_{to},\ s_{fs},\ s_{ft})'$ with a leak function:

$$\tau_{\mathbf{inp}} \cdot \dot{\mathbf{inp}} = -\mathbf{inp} + \mathbf{s}\ . \tag{4.1}$$

LA is formed by four leaky neurons ($\mathbf{la}$) activated as follows:

$$\tau_{\mathbf{la}} \cdot \dot{\mathbf{la}}_p = -\mathbf{la}_p + w_{\mathbf{inp,la}} \cdot \mathbf{inp}\ , \qquad \mathbf{la} = \varphi[\tanh[\mathbf{la}_p]] \tag{4.2}$$

where $w_{\mathbf{inp,la}}$ is the fixed weight of the connections from IMP to LA. The "double leak" processing of signals implemented by IMP and LA is used to smooth the derivative of LA (see Eq. (4.3)).

The trace of LA neurons ($\mathbf{la\_tr}$) is a leak function of the positive value of the derivative of their activation ($\dot{\mathbf{la}}$):

$$\tau_{\mathbf{la\_tr}} \cdot \dot{\mathbf{la\_tr}}_p = -\mathbf{la\_tr}_p + b_{\mathbf{la\_tr}} \cdot \varphi[\dot{\mathbf{la}}]\ , \qquad \mathbf{la\_tr} = \varphi[\tanh[\mathbf{la\_tr}_p]] \tag{4.3}$$

where $b_{\mathbf{la\_tr}}$ is an amplification coefficient.

CeA is formed by two leaky neurons ($\mathbf{cea}$) activated as follows:

$$\begin{aligned} \tau_{\mathbf{cea}} \cdot \dot{\mathbf{cea}}_p &= -\mathbf{cea}_p + \mathbf{W}_{\mathbf{la,cea}} \cdot \mathbf{la} + \mathbf{W}_{\mathbf{bla,cea}} \cdot \mathbf{bla} \\ \mathbf{cea} &= \varphi[\tanh[\mathbf{cea}_p]] \end{aligned} \tag{4.4}$$

VTA is formed by a dopamine leaky neuron (da) which activates as follows:

$$\tau_{da} \cdot \dot{da}_p = -da_p + bl_{da} + \mathbf{w}_{\mathbf{cea},da} \cdot \mathbf{cea}\ , \qquad da = \varphi[\tanh[da_p]] \tag{4.5}$$

where bl_{da} is the dopamine baseline.

The weights of the LA-CeA connections ($\mathbf{W}_{\mathbf{la,cea}}$) are updated with a three-element Hebb rule

involving CeA, LA's trace and dopamine:

$$\Delta \mathbf{W_{la,cea}} = \eta_{\mathbf{la,cea}} \cdot (\chi[da - th_{da}] \cdot da) \cdot \mathbf{cea} \cdot \mathbf{la\_tr}' \cdot (1 - |\mathbf{W_{la,cea}}|) \tag{4.6}$$

where $\eta_{\mathbf{la,cea}}$ is a learning rate, the term $(\chi[da - th_{da}] \cdot da)$ implies that learning takes place only when $da \geq th_{da}$, and the term $(1 - |\mathbf{W_{bla}}|)$ keeps the weights in the range $[-1, 1]$.

BLA: Functioning and Learning. BLA is formed by four leaky neurons (**bla**) activated as follows:

$$\begin{aligned} \tau_{\mathbf{bla}} \cdot \dot{\mathbf{bla}}_p &= -\mathbf{bla}_p + \mathbf{W_{bla}} \cdot \mathbf{bla} + (w_{\mathbf{la,bla}} \cdot \mathbf{la} + c_{\mathbf{bla}} \cdot \mathbf{la\_tr}) \\ \mathbf{bla} &= \varphi[\tanh[\mathbf{bla}_p]] \end{aligned} \tag{4.7}$$

where $c_{\mathbf{la\_tr}}$ is an amplification coefficient. According to this equation, with a transient constant input signal the activation of a BLA neuron presents a high initial peak (due to **la_tr**) followed by a lower constant value (due to **la**) and then by a smooth descent to 0 (due to the leak after the signal end): this activation has a derivative suitable for BLA learning (see below).

In order to train lateral connections of BLA, a trace of the derivative of the activation of BLA neurons **bla_tr** is computed as follows:

$$\tau_{\mathbf{bla\_tr}} \cdot \dot{\mathbf{bla\_tr}}_p = -\mathbf{bla\_tr}_p + \cdot \dot{\mathbf{bla}} \,. \tag{4.8}$$

Small values of this trace are ignored in the learning algorithm by considering the "cut trace" **bla_tr_cut** defined as: $\mathbf{bla\_tr\_cut} = \mathbf{bla\_tr}\ if\ |\mathbf{bla\_tr}| < th_{\mathbf{bla\_tr}}\ else\ \mathbf{bla\_tr\_cut} = 0$. Given the activation dynamics of BLA (Eq. (4.7)), the corresponding derivative (and, with some delay, its trace) presents: (1) an initial peak at signal onset; (2) a negative peak at the end of the signal onset; (3) a negative peak at the signal offset. The key point of the learning algorithm of BLA is that a connection between two neurons is potentiated in coincidence of a negative peak of the presynaptic neuron and a positive peak of the postsynaptic neuron. These two events mark a pre-synaptic-onset/post-synaptic-onset sequence (or a pre-synaptic-offset/post-synaptic-onset one). The matrix **S**, reported below, captures these conditions for all couples of neurons:

$$\mathbf{S} = \chi[\mathbf{bla\_tr\_cut}] \cdot \chi[-\mathbf{bla\_tr\_cut}]' - \chi[-\mathbf{bla\_tr\_cut}] \cdot \chi[\mathbf{bla\_tr\_cut}]' \,. \tag{4.9}$$

Denoting with pre and $post$ the presynaptic and postsynaptic neurons, **S** has an entry equal to 1 when $bla\_tr\_co_{pre} < 0$ and $bla\_tr\_co_{post} > 0$, equal to -1 when $bla\_tr\_co_{pre} < 0$ and $bla\_tr\_co_{post} > 0$, and equal to 0 otherwise. The learning rule of lateral connections is then:

$$\Delta \mathbf{W_{bla}} = \eta_{\mathbf{bla}} \cdot \chi[da - th_{da}] da \cdot (ltp_{\mathbf{bla}} \cdot \varphi[\mathbf{S}] + ltd_{\mathbf{bla}} \cdot \varphi[-\mathbf{S}])(1 - |\mathbf{W_{bla}}|) \tag{4.10}$$

where $\eta_{\mathbf{bla}}$ is a learning rate, $ltp_{\mathbf{bla}}$ is a long time potentiation coefficient, and $ltd_{\mathbf{bla}}$ is a short term depression coefficient.

Model's Parameters. The model's parameters were set as follows: $\tau_{\mathbf{inp}} = \tau_{\mathbf{la}} = \tau_{\mathbf{bla}} = 500$ ms, $\tau_{\mathbf{la\_tr}} = \tau_{\mathbf{bla\_tr}} = 5000$ ms, $\tau_{\mathbf{cea}} = 100$ ms, $\tau_{da} = 50$ ms, $w_{\mathbf{inp,la}} = 10$, $b_{\mathbf{la\_tr}} = 1000$, $w_{\mathbf{la,bla}} = 0.5$, $c_{\mathbf{bla}} = 60$, $bl_{da} = 0.3$, $th_{da} = 0.6$, $th_{\mathbf{la\_tr}} = 0.00001$, $\eta_{\mathbf{bla}} = 0.0005$, $\eta_{\mathbf{la,cea}} = 0.15$,

$ltp_{\mathbf{bla}} = 1.0$, $ltd_{\mathbf{bla}} = 0.3$. Some connections, assumed to be innate or pre-learned, are clumped to 1 (l=learned): $w_{blafs,ft} = 1$, $\mathbf{w}_{\mathbf{cea},da} = (1, 0)$, $\mathbf{W}_{\mathbf{la,cea}} = \begin{pmatrix} l & l & l & 1 \\ l & l & 1 & 1 \end{pmatrix}$, $\mathbf{W}_{\mathbf{bla,cea}} = \begin{pmatrix} l & l & l & 1 \\ l & l & l & 1 \end{pmatrix}$. The model's equations were integrated with the Euler method with a 50 ms step.

Published as:Mannella, Francesco, Zappacosta, Stefano, & Baldassarre, Gianluca. 2008. “A computational model of the amygdala nucleis role in second order conditioning”. Pages 321330 of: Tani, Minoru Asada Jun, Hallam, John, & Meyer, Jean-Arcady (eds), *Proceedings of the tenth international conference on simulation of adaptive behavior: From amimals to animals 10*, vol. 5040. Osaka, Japan: Springer.

Chapter 5

Brain Mechanisms underlying Learning of Habits and Goal-Driven Behaviour: A Computational Model of Instrumental Devaluation

Abstract

This paper presents an embodied biologically-plausible model investigating the relationships existing between classical and instrumental conditioning. The architecture and functioning of the model are constrained by anatomical and physiological assumptions drawn from the relevant neuroscientific literature. The model is validated by successfully reproducing the primary outcomes of instrumental-conditioning devaluation experiments conducted with normal and amygdala-lesioned rats (amygdala is a nucleus of the brain’s limbic system playing a key function in classical conditioning). These experiments are particularly important as they show how the sensitivity to internal states (such as satiety) exhibited by classical conditioning mechanisms can transfer to behaviors acquired on the basis of instrumental conditioning mechanisms. The work presented here is relevant for the behavioral and brain sciences as it based on a model, constrained and validated at both the neural and behavioral level, that indicates how internal states might modulate learning and performance of rigid habits so to produce the flexibility which is typical of goal-directed behaviour. Moreover, the present work is also relevant for autonomous robotics as it starts to investigate how the use of sophisticated motivational systems might allow building robots that are capable of exhibiting some of the flexibility of real organisms.

5.1 Introduction

The flexibility and capacity of adaptation of organisms greatly depends on their learning capabilities. For this reason, animal psychology has devoted great efforts to the study of learning processes. In particular, in the last century a huge body of empirical data have been collected around the two main experimental paradigms of ‘classical conditioning’ (Lieberman, 1993; Pavlov, 1927) and ‘instrumental conditioning’ (Domjan, 2006; Skinner, 1938; Thorndike, 1911).

‘Classical conditioning’ refers to an experimental paradigm in which a certain basic behaviour such as salivation or approaching (the ‘unconditioned response’ – UR), which is linked to a biologically salient stimulus such as food ingestion (the ‘unconditioned stimulus’ – US), gets associated to a neutral stimulus like the sound of a bell (the ‘conditioned stimulus’ – CS), after the neutral stimulus is repeatedly presented before the appearance of the salient stimulus. Such acquired associations are briefly referred to as ‘CS-US’ or ‘CS-UR’ associations (Lieberman, 1993; Pavlov, 1927).

‘Instrumental conditioning’ refers to an experimental paradigm in which an animal, given a certain stimulus/ contexts such as a lever in a cage (the ‘stimulus’ – S), learns to produce a particular action such as pressing the lever (the ‘response’ – R), which produces a certain outcome such as the opening of the cage (the ‘action outcome’ – O), if this outcome is consistently accompanied by a reward such as the access to food. In this case, the acquired associations are briefly referred to as either ‘S-R’

associations, when the reactive nature of the acquired behaviour is stressed, or 'A-O' associations, when the goal-directed nature of behaviour is stressed (Domjan, 2006; Skinner, 1938; Thorndike, 1911, see below).

This empirical work has been paralleled by the development, within the machine learning literature, of 'reinforcement learning algorithms' (Sutton and Barto, 1998, 1981), that is algorithms directed to provide machines with the capacity of learning new behaviors on the basis of rewarding stimuli (i.e. signals from the external environment that inform the machine about the achievement of desired goals). Interestingly, reinforcement learning algorithms have gained increasing interest within the empirical literature on animal learning as they represent theoretical models that can potentially furnish coherent explanations of organisms' learning processes. Indeed, one of such models, the so-called temporal-difference (TD) learning algorithm, is currently considered as the best theoretical account of several key empirical findings (Dayan and Balleine, 2002; Schultz, 2002).

Notwithstanding their success, standard reinforcement learning models suffer of several limitations from a biological point of view. In particular, three of the main drawbacks are as follows. First, such models ignore the role of internal states (e.g. hungriness vs. satiety related to a certain type of food) in modulating the effects of 'external' rewards (e.g. the receival of such a food). Such kind of effects are demonstrated by organisms, for example, in 'devaluation' experiments in which animals tend to change their reinforced behaviors in case the value of a rewarding stimulus, such as a food, is suddenly decreased through satiation or its association with poison. By ignoring the role of internal states in learning and behavior, current reinforcement learning models can not account for such effects.

Second, standard models tend to conflate the notions of classical conditioning (also called 'Pavlovian conditioning') and instrumental conditioning (also called 'operant conditioning'). On the contrary, accumulating empirical evidence indicates that classical and instrumental conditioning are based on different processes that rely on distinct neural systems. Furthermore, such processes interplay in complex ways (Dayan and Balleine, 2002), as demonstrated, for example, by phenomena like 'Pavlovian-Instrumental Transfer' (where a conditioned stimulus that is predictive of reward can energize the execution of instrumentally acquired behaviours), and 'incentive learning' (where, under certain conditions, the valence of an action's outcome need to be re learned to exert its effects on behaviour).

Different brain mechanisms underlying operant and classical conditioning There is nowadays a wide evidence that the amygdala (Amg - an almond shape group of nuclei within the temporal lobe, part of the brain limbic system) is a main actor in classical (pavlovian) conditioning processes linking CSs to both appetitive and negative USs (Baxter and Murray, 2002; Cardinal et al., 2003). In particular, the basolateral complex of amygdala (BLA) is necessary for a CS to acquire the same rewarding and motivational value of the US. When BLA is lesioned associations between CSs and URs can still be made, but the behavioural responses cannot be further transferred from a CS to another neutral stimulus, second order associations between USs and CSs being thus impeded (Hatfield et al., 1996). Furthermore lesions of BLA disrupt an animal's ability to keep a CS linked to the current motivational value of an US. When a rat is presented with a stimulus, let's say a light, associated with a food which it was previously satiated or nauseated of, in normal conditions the animal diminishes its appetitive

responses to the light as in the case it is presented with the devalued food (conditioned devaluation). Lesions to BLA disconnect the current value of the food US from that of the light CS, leading to appetitive behaviours when the animal is presented with the CS after CS-US conditioning and subsequent US devaluation (Hatfield et al., 1996). In permitting second order conditioning and conditioned devaluation behaviours, BLA processing is part of a complex associative system including orbitofrontal cortex (OFC). In fact OFC has been showed to be highly integrated with BLA, being necessary for second order associations in BLA to be correctly built (Saddoris et al., 2005; Schoenbaum et al., 2003), and for long term mantainance of those associations (Pickens et al., 2003).

When an association between a reinforcement or a punishment and actions made to produce it has to be learned, the role of the corticostriatal loops becomes evident (Yin and Knowlton, 2006). The striatum is the input portion of the basal ganglia, a set of forebrain subcortical nuclei playing an important function in voluntary movement; in rats, the striatum can be divided in (Yin and Knowlton, 2006): (1) dorsolateral striatum, mainly underlying motor- execution functions (Romanelli et al., 2005), (2) dorsomedial striatum, playing a role in motor- preparation, attention and cognitive functions (Lawrence et al., 2000), and (3) ventral striatum, considered an important interface (Mogenson et al., 1980) between the motivational processes taking place in the limbic system and the motor processes taking place within the rest of the basal ganglia and cortex. Each part of the striatum is reached by a different set of cortical afferents, from motor and premotor cortices (dorsolateral part) to associative cortices (dorsomedial part) and prefrontal cortical districts (ventral part). These cortical areas receive themselves inputs from the correspective striatal areas, through afferents from the basal ganglia outout nuclei to the thalamus (Haber, 2003; Haber et al., 2000; McFarland and Haber, 2002). A large set of data confirms that the dorsolateral striatum - cortical loops are necessary to form the S-R associations underlying instrumental behaviours at the level of motor reactions to stimuli (Yin and Knowlton, 2006; Yin et al., 2004). These associations are formed through long sessions of trials and errors, and, once learned, tend to be automatically recalled by specific stimuli in a fixed way, being insensitive to expectations about future. The behaviours emerging through this kind of learning are what we call ‘habits’. Ventral striatum, instead, seems to act as an interface between the processing of contingencies between actions and possible outcomes (A-O) within the ventromedial prefrontal cortex, and the elaboration of the current motivational value of outcomes within the limbic system, mainly Amg. This interaction between the A-O associative system and the CS-US associative system should be at the base of the devaluation behaviours, permitting to the animal to switch between outcomes to persecute, according with the current motivational value of each outcome. Within this theoretical framework, the A-O system is thought to guide learning of the S-R mechanisms during the initial phases of an instrumental training, then, once learned, S-R associations could directly trigger behaviours.

The role of ventral striatum as al limbic-motor interface includes a set of functional properties. In particular accumbal activation influences both general neural activation and goal selection in presence of a stimulus carrying a motivational value. The internal structure of ventral striatum can be divided in two areas: nucleus accumbens shell (NAccS) and nucleus accumbens core (NAccC) (Brauer et al., 2000).

A direct contribution of the ventral striatum over instrumental behaviours goes through NAccC. In fact, while lesions to NAccS doesn’t have any effect over devaluation tasks, NAccC disruption

abolishes devaluation behaviours in instrumental tasks similar to the task targeted in this paper (Corbit et al., 2001) (see section 5.2). Moreover, animals with NAccC lesions are also impaired in second order reinforcement tasks, where a CS (acquired instrumentally or through pavlovian conditioning) is able to lead a subsequent instrumental learning (Ito et al., 2004). Both instrumental devaluation and second order reinforcement share the necessity for the motivational value to be transferred, in a pavlovian way, from the US to a CS (being it the manipulandum in the devaluation task or the CS used as a reinforcer in second order reinforcement). While there is evidence that BLA processing is also necessary for instrumental devaluation to take place (Balleine et al., 2003), there is also evidence that direct interactions between BLA and NAccC underlie the ability to resolve second order reinforcement tasks (Di Ciano and Everitt, 2004). Activation by BLA influences the processing of the striato-cortical loop including NAccC and prelimbic cortex (PL) (Gorelova and Yang, 1997; Zahm, 2000). While doing so, it also lead to modulation of the activation of the dorsomedial striatal-cortical loop and finally of the dorsolateral striatal-cortical loop, both via striatonigrostriatal spiral projections (Haber et al., 2000), cortico-thalamo-cortical projections (McFarland and Haber, 2002) and cortico-cortical reentrant connections , (Haber, 2003), thus linking the elaboration of goal selection to the action selection mechanisms.

NAccS, instead, seems to be responsible of phenomena such as 'pavlovian - to - instrumental transfert', in which a CS, previously associated with an US, enhances activation when presented during an instrumental task where the reinforcer is different from that of the conditioning procedure (US') (Corbit et al., 2001). This property should emerge from a particular efferent pathway of this area, that includes a strong output to ventral tegmental area (VTA), both directly and through the ventral pallidum, the pedunculopontine tegmental nucleus (PPT) and the lateral hypothalamus (LH) (Usuda et al., 1998; Zahm, 2000). Together with a strong afferent projetion from VTA itself (Voorn et al., 1986), this pathway should exert a major influence on the overall modulation of dopaminergic activation and, through this, on dopamine-dependent sensomotory activation. Following this overall schema, BLA - NAccC pathway should be a direct channel through which pavlovian associative mechanisms exert their influence over reinforcement learning, permitting devaluation behaviours in instrumental tasks.

Therefore devaluation instrumental paradigms reveal to be the best framework to start overcoming the three limits of standard reinforcement learning algorithms discussed in the previous section, for through them 1) the necessity for an elaboration of the motivational value throught the internal states of the organism is revealed, 2) the interaction between pavlovian and instrumental learning processes can be analyzed and 3) an hypothesis can be drawn about the relation between habits and goal-directed behaviours.

This paper presents a novel computational model which is strongly rooted in the anatomy and physiology of the mammal brain and starts to address the drawbacks of current reinforcement learning models within the biological framework illustrated above (a preliminary version of the model was presented in Mannella et al., 2007). In particular, the model reproduces the results of an empirical experiment (Balleine et al., 2003) which demonstrates the phenomenon of *devaluation* in an instrumental conditioning task and proposes a coherent picture about the discussed possible neural mechanisms underlying it. The model is based on the following fundamental hypotheses discussed previously:

1. the *amygdala* constitutes the CS-US associator at the core of Pavlovian conditioning phenomena;

2. the *cortex-dorsolateral striatum* pathway, forming S-R associations, constitutes the main actor involved in instrumental conditioning;

3. the *cortex-ventral striatum* pathway, forming A-O associations, constitutes the main actor involved in goal selection processes and planning;

4. the *amygdala-ventral striatum* pathway 'bridges' classical conditioning processes happening in the amygdala and instrumental processes taking place in the basal ganglia.

By reproducing the basic results of both normal and lesioned rats the model provides significant evidence for these three fundamental hypotheses and, more importantly, it contributes to clarify the relationships existing between the neural structures and processes underlying them.

5.2 The target experiment and the simulated environment

The model presented here was tested within an embodied system because, as mentioned in the introduction, one of the long-term goal of this research is to build models that are based on sound anatomical and physiological neuroscientific evidence and that at the same time are capable of *scaling* to function in realistic robotic setups. Although we are aware that the role of the 'degree of embodiment and situatedness' of the model and simulations presented here is rather limited (e.g. the sensors and actuators used are rather simplified, low-level behaviors are hardwired, etc.), nevertheless the use of a robotic test forced us to design a model potentially capable to cope with the difficulties posed by more realistic setups. For example, the randomly variable duration of the trials, actions' execution, and rewarding effects posed interesting challenges to the robustness of the learning algorithms of the model.

The model was tested with a simulated robotic rat ('ICEAsim') developed within the EU project ICEA on the basis of the physics 3D simulator Webots^{TM}. The model was written in Matlab^{TM} and was interfaced with ICEAsim through a TCP/IP connection. The robotic setup used to test the model is shown in Figure 5.1 and it is now briefly described skipping irrelevant details. The training and test environment is composed by a grey-walled chamber containing a yellow lever, a red chain, and a food-dispenser that turns green or blue when respectively food A or food B is delivered in it. When 'pressed' or 'pulled', the lever and chain make respectively food A or B (the rewarding stimuli) available at the dispenser.

The simulated rat is a two-wheel robot equipped with various sensors. Among these, the experiments reported here use two cameras (furnishing a panoramic 300 degrees view) and the whisker sensors. The rat uses the cameras to detect the lever, the chain and the food dispenser, in particular their presence/absence (via their color) and their (egocentric) direction. The rat uses the whiskers, activated with one if bent beyond a certain threshold and zero otherwise, to detect contacts with obstacles. The rat is also endowed with *internal* sensors related to satiety for either food A or B (these sensors assume the value of one when the rat is satiated, and zero otherwise). The rat's actuators are two motors that can independently control the speed of the two wheels.

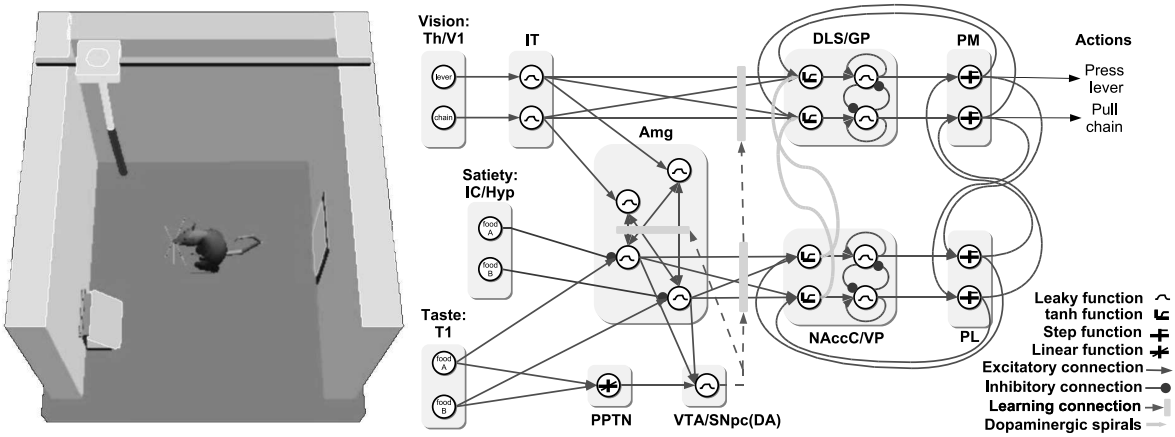


Figure 5.1: *Left: A snapshot of the simulator, showing the simulated rat at the center of the experimental chamber, the food dispenser (behind the rat), the lever (at the rat's left hand side) and the chain (at the rat's right hand side). Right: The architecture of the model.*

The information fed to the model is only related to the presence/absence of the lever and chain in the test chamber and food A and food B in mouth, whereas the other information is used to control a number low-level hardwired behavioral routines. These routines, triggered either by the model or directly by stimuli, are as follows: (1) 'obstacle avoidance routine': this routine, triggered by the whiskers, 'overwrites' all other actions to avoid obstacles; (2 and 3) 'lever press routine' and 'chain pull routine': these routines, activated by the model, cause the rat to approach the lever/chain on the basis of their visually detected direction; when the lever/chain are touched they activate the food delivery in the dispenser; (4) 'consummatory routine': when the dispenser turns green or blue (this signals the presence of food in it), the rat approaches and touches it ('consummation' of the food) so causing the perception of either food A or food B in mouth; the routine ends after the rat touches the dispenser ten times.

The devaluation experiment is divided in a training phase and two test phases. The training phase lasts 16 mins and the two test phases 4 mins each. Each phase is divided in trials that end either when the rat executes the correct action and consumes the food or after a 15 s timeout. In each trial the rat is set in the middle of the chamber with an orientation randomly set between the lever and the chain direction. In the trials of the training-phase either the lever and food A or the chain and food B are used in an alternate fashion and the rat is always 'hungry' (the two satiation sensors are set to 0). In the two test phases, the rat is respectively satiated either with food A (the satiation sensors for food A and B are respectively set to one and zero) or with food B. In all trials of the two test phases *both* manipulanda are present and the rat is evaluated in extinction (i.e. without delivery of food). The experiment (the three phases) was run 20 times with 'unlesioned' artificial rats and 20 times with 'lesioned' rats.

5.3 The Model

The model's input component is formed by three vectors of units activated by the sensors illustrated in Section 5.2. First, a vector $\mathbf{v1} = (v1_{lever}, v1_{chain})'$ of two units, corresponding to the primary visual cortex (V1), encodes the presence/absence of the lever and the chain in a 0/1 binary fashion.

Second, a vector $\mathbf{t1} = (t1_{foodA}, t1_{foodB})'$ of two units, corresponding to the primary taste cortex within the insular cortex (T1 Verhagen et al., 2004), encodes the presence/absence of food A and food B in the rat's mouth in a 0/1 binary fashion. Finally, a vector $\mathbf{ic} = (ic_{foodA}, ic_{foodB})'$ of two units, corresponding to both the processing by insular cortex (IC; Yeterian and Pandya, 1995) and the elaboration made by hypothalamus (Hyp; King, 2006), encodes the non-satiation/satiation for food A and food B again in a 0/1 binary fashion. Inputs from $\mathbf{v1}$ are further processed within a vector $\mathbf{it}$ of two leaky units, corresponding to the inferotemporal cortex (Yeterian and Pandya, 1995):

$$\tau_{\mathbf{it}} \cdot \dot{\mathbf{it}}_p = -\mathbf{it}_p + \mathbf{v1} \tag{5.1}$$

$$\mathbf{it} = \varphi[tanh[\mathbf{it}_p]]$$

where $\mathbf{it}_p$ encodes the activation potential of IT units, and $\varphi[x]$ is part of IT units transfer function (if $x \leq 0$ then $\varphi[x] = 0$, else $\varphi[x] = x$).

The model (Figure 5.1) is formed by three major sub-systems: (a) a S-S associator, corresponding to BLA; (b) a *static* S-R action selector, corresponding to the cortico-dorsolateral striatal pathway passing through IT, DLS, entopeduncular nucleus of basal ganglia (EP), and premotor cortex (PMC); (c) a *dynamic* S-S-R associator, corresponding to the cortico-ventral striatal pathway passing through Amg, ventral striatum (VS), in particular to the nucleus accumbens core (NAccC), ventral pallidum (VP), and prelimbic cortex (PL). Note that 'static' and 'dynamic' terms are used here to refer respectively to associators which do not or which do implement associations which can be modulated on-the-fly by internal states. Now the three sub-systems are presented in detail.

5.3.1 The Amygdala, an S-S Associator, and the Dopamine Learning Signal

The S-S associator implements Pavlovian conditioning through the association between CSs and USs ('stimulus substitution'). In real brains this role seems to be played by BLA (Baxter and Murray, 2002; Cardinal et al., 2003). There are massive reciprocal connections between BLA and several brain areas, including inferotemporal cortex (IT), prefrontal cortex (PFC), and hippocampus (Hip) (McDonald, 1998; Pitkänen et al., 2000; Price, 2003; Rolls, 2005). Furthermore, BLA receives inputs from insular cortex (IC), Hypothalamus (Hyp) and posterior intralaminar nuclei of thalamus (PIL) (McDonald, 1998; Pitkänen et al., 2000; Shi and Davis, 1999). These connections underlie an interplay between processes related to perceived or represented external context and stimuli (IT, PFC, Hip) and processes related to internal states (IC, Hyp, PIL). In general, BLA can be seen as playing the function of assigning a "subjective valence" (i.e. a mark of biological relevance) to external previously-neutral events on the basis of the animal's internal states (needs, motivations, etc.), and to use this to both regulate learning processes and directly influence behavior.

The S-S model's associator, considered to abstract the processes taking place in the BLA, performs 'asynchronous learning/synchronous functioning' associations. First, the associator associates between them asynchronous stimuli perceived at different times (namely, it associates CSs with USs): this associative learning takes place only if USs cause a dopamine (DA) release (see below). Then, once the association is established, CSs are able to synchronously re-activate the USs' representations

in BLA.

Mathematically, the associator is composed by a vector $\mathbf{bla} = (bla_{lever}, bla_{chain}, bla_{foodA}$,-$bla_{foodB})'$ of four laterally-connected leaky units that process the input signals as follows:

$$\begin{aligned}\tau_{\mathbf{bla}} \cdot \dot{\mathbf{bla}}_p = &-\mathbf{bla}_p + \\ &\mathbf{W}_{it-bla} \cdot \mathbf{it} + \mathbf{W}_{t1-bla} \cdot \mathbf{t1} + \\ &\mathbf{W}_{ic-bla} \cdot \mathbf{ic} + \mathbf{W}_{bla} \cdot \mathbf{bla}\end{aligned} \tag{5.2}$$

$$\mathbf{bla} = \varphi[tanh[\mathbf{bla}_p]]$$

where $\mathbf{bla}_p$ are the activation potentials of BLA units, $\mathbf{W}_{it-bla}$ is the matrix of connection weights between IT and BLA, $\mathbf{W}_{t1-bla}$ is the matrix of connection weights between T1 and BLA, $\mathbf{W}_{ic-bla}$ is the matrix of connection weights between IC and BLA, and $\mathbf{W}_{bla}$ is the matrix of all-to-all lateral connection weights within BLA. Note that while external stimuli play the function of input signals to the model, internal stimuli *modulate* the internal representations of external stimuli. In particular, ic_{foodB} and ic_{foodB} assume a value of either zero or one when the corresponding satiation is respectively low or high, and this activation inhibits the hedonic internal representation of such foods within BLA via inhibitory connections (see Section 5.3.3). This assumption is supported by evidence indicating that a similar computation is performed in the secondary taste areas of the prefrontal/insular cortex which are connected to BLA (Rolls, 2005). This part of the model is particularly important because, as we shall see, it mediates the influence of the shifts of primary motivations on both learning and behaviour.

The associator's learning is based on the *onset* of input signals, detected as follows. First, 'leaky traces' $\mathbf{tr}$ of the derivative of $\mathbf{bla}$, $\dot{\mathbf{bla}}$, trunked to positive values, are computed:

$$\tau_{\mathbf{tr}} \cdot \dot{\mathbf{tr}} = -\mathbf{tr} + C_{BLA} \cdot \varphi[\dot{\mathbf{bla}}] \tag{5.3}$$

where C_{BLA} is a coefficient used to amplify the small signals $\dot{\mathbf{bla}}$. Second, the derivatives of the elements of $\mathbf{tr}$ are computed. Notice that when positive these derivatives detect the onset of the original signals, whereas when negative they detect the fact that some time elapsed since such onset took place.

The weights between BLA's units are updated on the basis of the signs of $\dot{\mathbf{tr}}$ and the DA signal (see below). In particular, when (and only when) the derivative of the presynaptic unit's trace is negative and the derivative of the postsynaptic unit's trace is positive (i.e. when the presynaptic unit fires *before* the postsynaptic unit) the related connection is strengthened. Instead, when (and only when) the derivative of the presynaptic unit's trace is positive and the derivative of the postsynaptic unit's trace is negative (i.e. when the presynaptic unit fires *after* the postsynaptic unit) the related connection is weakened. This condition is encoded, for all couples of units, in the Boolean matrix $\mathbf{L}$: each element of this matrix is equal to + 1 for synapses to be strengthened and equal to -1 for synapses to be weakened. Formally, the rule used to update weights of BLA's lateral connections is as follows:

$$\Delta \mathbf{W}_{bla} = \eta_{bla} \cdot \varphi[da - th_{da}] \cdot \mathbf{L} \cdot (1 - |\mathbf{W}_{bla}|) \tag{5.4}$$

where η_{bla} is a learning rate coefficient, da is the dopamine signal, and th_{da} is a threshold over which dopamine elicits learning.

DA release (corresponding to activation of the ventral tegmental area, VTA, and in the substantia nigra pars compacta, SNpc) is triggered by BLA through its units bla_{foodA} and bla_{foodB} which issue signals to the dopamine unit of VTA/SNpc through connections which are an abstraction of the connections going from BLA to Hyp and hence to the VTA (Petrovich et al., 2002,?). Notice that, thanks to these unlearned connections, BLA units bla_{foodA} and bla_{foodB} come to represent the 'hedonic impact of food', that is internal representations of food directly causing the DA signals underlying learning (as simulated here) and motivation (not represented here). DA release is also triggered by the 'primary reward signals' (which is signals associated to stimuli which cause DA learning signals via unlearned connections) received from the pedunculopontine tegmental nucleus (PPT) denoted as ppt ($ppt = t1_{foodA} + t1_{foodB}$) (Kobayashi and Okada, 2007; Pan and Hyland, 2005). Specifically, the DA signal is computed as follows:

$$\tau_{da_p} \cdot \dot{da_p} = -da_p + da_{baseline} + w_{bla-da} \cdot (bla_{f_A} + bla_{f_B}) + w_{ppt-da} \cdot ppt \tag{5.5}$$

$$da = \varphi[tanh[da_p]]$$

DA not only drives learning taking place within the BLA (Equation 5.4) but it also drives learning involving the two action selectors pivoting respectively on the DLS (see Section 5.3.2, Equation 5.9) and NAccC (see Section 5.3.3,Equation 5.13).However, note that although here the choice of DA as a unique signal underlying learning in all the three brain systems was done for simplicity and because it was not in contrast with neuroscientific evidence, norepinephrine (NE) might alternatively or complementarily drive associative learning in BLA, as suggested by some empirical evidence (McIntyre et al., 2002). Indeed, at least in some cases DA might not be suited to drive learning in BLA. In fact, BLA is known to implement associations between neutral stimuli and both positive and *negative* stimuli whilst DA seems to be mainly involved in signalling stimuli with positive valence (citeShultz, getting formal) whereas stimuli with negative valence tend to cause its depression (citepaper on pinch; depression might be useful for active avoidance learning).

Note that in the experiments reported in Section 5.4 the lesions of rats' BLA were simulated by clumping the BLA units at a zero activation.

5.3.2 The Dorsolateral Corticostriatocortical Pathway: A Static S-R Action Selector

The static action selector learns 'habits', that is rigid S-R associations, through reinforcement learning processes. In real brains, this function might be implemented by the corticostriatal loops involving

in particular the DLS and premotor cortex (PMC) or motor cortex (MC) (Yin and Knowlton, 2006; here for simplicity only PMC will be considered). At the input level, this sub-system receives signals from IT, feedback input signals from PMC, and modulation DA signals from VS (see Section 5.3.3). At the output level, represented by the PMC, the sub-system selects the actions to be executed (e.g., lever-press or chain-pull).

In the model this component is formed by four layers of units encoded in four vectors: (a) a leaky-unit layer $\mathbf{it}$ corresponding to IT;(b) a non-leaky unit layer $\mathbf{dls} = (dls_{lever}, dls_{chain})'$ corresponding to DLS and encoding the total signals in favour of the execution of the two actions ('votes');(c) a leaky-unit layer $\mathbf{ep} = (ep_{lever}, ep_{chain})'$ corresponding to the EP, formed by reciprocally inhibiting units which implement a competition in order to select one of the two actions (this is an abstraction of the selective function which in real brains might be implemented by the re-entrant thalamo-cortical connections, Dayan and Balleine, 2002);(d) a layer $\mathbf{pmc} = (pmc_{lever}, pmc_{chain})'$ corresponding to PMC, representing the selected action with a 0/1 binary code.Note that some of these notations are mainly aimed at understanding the model architecture. In particular, the implementation of dorsolateral basal ganglia presented here is focused on the overall functionality of selection via competition more than on the its detailed micro-architecture (for a more accurate computational model of basal ganglia, see Gurney et al., 2001a, Gurney et al., 2001b, and Humphries et al., 2006).

IT is connected through all-to-all connections to DLS. DLS (non-leaky) units receive the signals from IT, together with the feedbacks from PMC, which can be though of as 'votes' in favour of the selection of either one of the two actions. Importantly, these votes are modulated by NAccC activation $\mathbf{naccc}$ (see Equation 5.10) which should be considered an abstraction of the striatonigrostriatal connections (seeSection 5.3.3 and 5.1; the way of representing the modulatory effect of DA used here is as in Humphries et al. (2006)):

$$\mathbf{dls}_p = (\mathbf{W}_{(it-dls)} \cdot \mathbf{it} + \mathbf{pmc} + \mathbf{naccc}) \cdot (1 + C_{naccc} \cdot \mathbf{naccc}) \tag{5.6}$$

$$\mathbf{dls} = \varphi[tanh[\mathbf{dls}_p + dls_{baseline}]]$$

The selection of actions is performed on the basis of these votes through a competition taking place between the leaky units of EP:

$$\tau_{ep} \cdot \dot{\mathbf{ep}}_p = -\mathbf{ep}_p + C_{ep} \cdot \mathbf{dls} + \mathbf{ep}_{baseline} + \mathbf{n}_{ep} + \mathbf{W}_{ep} \cdot \mathbf{ep} \tag{5.7}$$

$$\mathbf{ep} = \varphi[tanh[\mathbf{ep}_p]]$$

where C_{ep} is a coefficient scaling the DLS votes, $\mathbf{ep}_{baseline}$ is a baseline activation, $\mathbf{n}_{ep}$ is a noise vector with components uniformly drawn in $[-n_{ep}, +n_{ep}]$, and $\mathbf{W}_{ep}$ are the EP lateral connection weights.

When one of the EP units overcomes an activation threshold th_{pmc}, the corresponding unit of

PMC is set to one (otherwise PMC units are kept at zero) and the corresponding action is executed. PMC 's activation is also influenced by the activation of PL units $\mathbf{pl}$ (see Equation 5.12):

$$\mathbf{pmc} = \psi[(\mathbf{ep} + \mathbf{pl}) - th_{pmc}] \tag{5.8}$$

where $\psi[x]$ is the step function (if $x \leq 0$ then $\psi[x] = 0$, else $\psi[x] = 1$). Once the execution of the routine corresponding to the selected action terminates, the connection weights between IT and DLS, $\mathbf{W}_{it-dls}$, are modified according to the DA signal (this is null in the case the wrong action is selected):

$$\Delta\mathbf{W}_{it-dls} = \eta_{it-dls} \cdot \varphi[da - th_{da}] \cdot \mathbf{dls} \cdot \mathbf{it}' \tag{5.9}$$

where η_{it-dls} is a learning coefficient. Note that here PMC's feedback to DLS is essential to allow the Hebbian product $\mathbf{dls} \cdot \mathbf{it}'$ to strengthen the connection weights between correlating stimuli and actions in the presence of DA.

5.3.3 The ventral Corticostriatocortical Pathway: A Dynamic S-R Action Selector

The dynamic action selector learns (S-)S-R associations through a reinforcement learning process that exploits the information encoded in the BLA's S-S associations (e.g., the 'lever-hedonic value of food A' association). In real brains this function might be implemented by the neural pathway connecting the BLA nuclei of BLA to the ventral striatum, in particular to the portion of it called nucleus accumbens core (NAccC; see Corbit et al., 2001 and Baxter and Murray, 2002). This pathway sends signals ('votes') to VP which then selects *desired states* via the prelimbic cortex (PL). These 'desired states' are the potential outcomes of actions, in the model 'food A with hunger-for-food-A expected after execution of a lever-press action' and 'food B with hunger-for-food-B expected after execution of a chain-pull action'. As we shall see below, as these outcomes can participate to trigger the execution of specific actions in the DLS (habit) pathway, the BLA- NAccC-DLS can be said to implement inverted O-A relations which allow desired outcomes (i.e. USs) elicited by CSs in BLA, to contribute to select actions at the level of the habit pathway either biasing the DLS competition via the BLA-NAccC-DLS pathway or by 'overwriting' the action selection in PMC via the BLA-NAccC-VP-PL-PMC pathway.

In the model, the BLA-NAccC neural pathway is implemented by an all-to-all connection matrix $\mathbf{W}_{bla-naccc}$ linking the BLA's hedonic representations of food (here denoted as $\mathbf{bla}_{us} = (bla_{foodA}, bla_{foodB})'$), to the NAccC (non-leaky) units. Similarly to DLS and PMC, NAccC also receives re-entrant input signals from PL (these signals play an important role in for learning, see below):

$$\mathbf{naccc}_p = \mathbf{W}_{bla-naccc} \cdot \mathbf{bla}_{us} + \mathbf{pl} \tag{5.10}$$

$$\mathbf{naccc} = \varphi[tanh[\mathbf{naccc}_p + naccc_{baseline}]]$$

NAccC units play a function similar to DLS units for EP in that they represent 'votes' that bias the competition taking place between the VP leaky-units units and directed to select one 'desired

outcome':

$$\tau_{vp} \cdot \dot{\mathbf{vp}}_p = -\mathbf{vp}_p + C_{vp} \cdot \mathbf{naccc} + \mathbf{vp}_{baseline} + \mathbf{n_{vp}} + \mathbf{W}_{vp} \cdot \mathbf{vp} \tag{5.11}$$

$$\mathbf{vp} = \varphi[tanh[\mathbf{vp}_p]]$$

where C_{vp} is a coefficient scaling the votes, $\mathbf{vp}_{baseline}$ is a baseline activation of VP, $\mathbf{n}_{vp}$ is a noise vector with components uniformly drawn in $[-n_{vp}, +n_{vp}]$ $\mathbf{W}_{vp}$ are the EP lateral connection weights.

When one of the VP units reaches the activation threshold th_{pl}, the corresponding desired outcome is activated in PL (as PMC, PL has binary activations):

$$\mathbf{pl} = \psi[(\mathbf{vp} + \mathbf{pmc}) - th_{pl}] \tag{5.12}$$

Note that PL activation is also influenced by PMC activation $\mathbf{pmc}$: this signal has a very important function for updating BLA-NAccC connection weights as it can carry the information related to the executed action, represented in PMC, to the expected outcomes, represented in PL, and then backward to the NAccC which can then form suitable associations with the representations of BLA. In particular, similarly to IT-DLS connections, BLA-NAccC connections $\mathbf{W}_{bla-naccc}$ are modified after action execution on the basis of the DA-dependent Hebbian rule involving the activations of BLA and NAccC (on its turn influenced by the re-entrant signals from PL):

$$\Delta\mathbf{W}_{bla-naccc} = \eta_{bla-naccc} \cdot \varphi[da - th_{da}] \cdot (\mathbf{bla}_{us} \cdot \mathbf{naccc}) \tag{5.13}$$

where $\eta_{(bla-naccc)}$ is a learning rate coefficient.

The importance of the BLA-NAccC dynamic action selector resides in the fact that its 'votes' for the various actions can be modulated *on-the-fly* by the organism's motivational states, in particular by satiety for either one of the two foods. In general, this mechanisms opens' up the possibility for the motivational-sensitive Pavlovian system (mainly the BLA in the model) to exert a direct effect on actions without the need to pass through re-learning processes, as it will be exemplified by the devaluation experiments illustrated in the next section.

5.4 Results

This section describes the basic functioning of the model on the basis of Figures 5.2 and 5.4. The figures show the activations of various units related to the lever (data related to the chain are omitted as qualitatively similar) during both the training (5.2) and testing phases (5.4) of an experiment run with a non-lesioned simulated rat. (5.4) also shows the activations of the same units in the two test phases for rats with three kind of lesions.

At the beginning of the training phase, the baseline activations of DLS and NAccC (dls_{lever} $naccc_{lev}$), together with noise, are sufficient to occasionally trigger the execution of an action (m_{lever})

by the competition taking place in PMC (pmc_{lever}). When the behavioral routine corresponding to the selected action is appropriate for the environment configuration ('lever press' in the presence of lever), the dispenser becomes yellow, the rat approaches it and consumes the corresponding food (s_{foodA}). The food consumption activates the internal hedonic representation of food in BLA (bla_{foodA}) and hence the units in VTA/SNpc with the consequent release of DA in DLS. This drives the learning of the dorsal corticostriatal instrumental pathway.

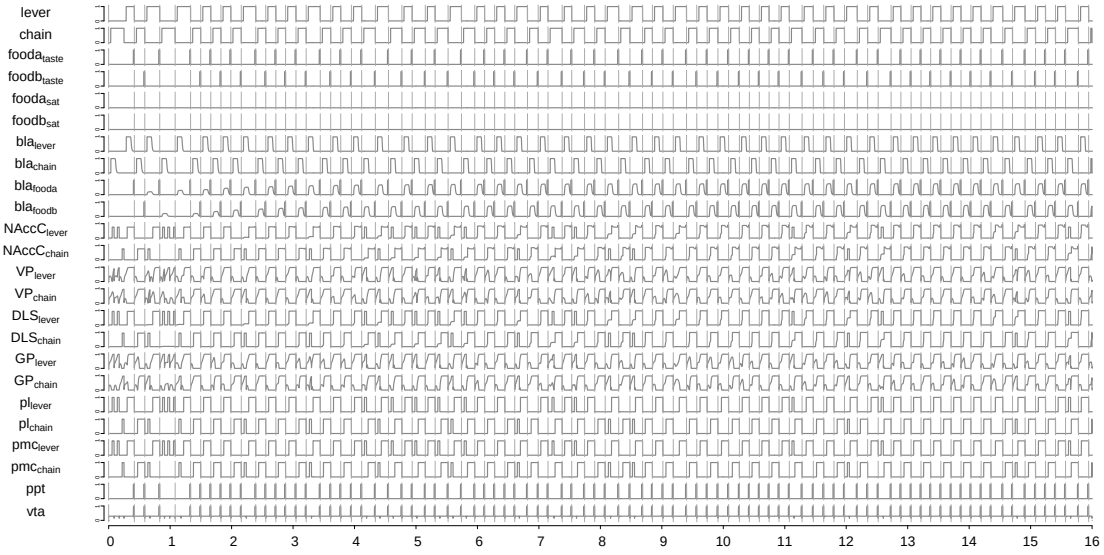


Figure 5.2: *Activations of some key units of a non- lesioned rat during the training phase; Trials are separated by short vertical lines.*

The effect of these events is that after a few learning trials the model learns to reliably and fast perform the action which is appropriate to the current context. The progress of learning can be seen in terms of: (a) the increase of DLS's votes for the press lever action (dls_{lever}) in the trials in which the lever is present; (b) the increase of the regularity of the peaks of the food A amygdala units (bla_{foodA}); (c) the DA release in VTA-SNpc ($vta - SNpc$).

When instrumental S-R associations begin form due to instrumental learning, the vision of the neutral stimuli of the lever (s_{lever}, bla_{lev}) starts to be reliably followed, within a relatively small time interval, by the food perception (s_{foodA}) and the consequent DA release (da). This contingency and the DA signal allow the Pavlovian learning taking place within BLA to 'take off' and form S-S associations between the lever and BLA's food A representation. This is evident from the fact that after a few successful trials the bla_{foodA} unit's activation not only show a peak when the food A is delivered but are also pre- activated by the presence of the lever: this reveals that a Pavlovian association is being acquired between the conditioned stimulus (lever) and the unconditioned stimulus (food). The pre-activation of the bla_{foodA} unit due to the perception of the conditioned stimulus is responsible for the early DA release da which anticipates the future delivery of reward: this mimics an important well-known phenomenon observed in real animals (Schultz, 2002).

The last important learning process takes place in the BLA-NAccC pathway. The rat's consumption of food A activates both the BLA's hedonic representation of it (bla_{foodA}) and, via the VTA/SNpc,

which results in a strong DA signal. This creates a strong association between the hedonic representation of food and the last executed action. The key point here is that once the S-S associations are formed in the BLA, conditioned stimuli such as the lever can trigger the activation of the BLA's hedonic representation of the related food and, via these, influence PUT's action selection via NAccC. This is shown by the fact that, after some training, NAccC starts to be activated and to vote for the correct actions ($naccc_{lever}$). The importance of the formation of this Stimuli-BLA-NAcc-PM pathway resides in the fact that it constitutes the fundamental bridge between the the Pavlovian processes happening in the amygdala and the instrumental processes happening in the basal ganglia pathway (cortex–dorsal striatum– putamen–thalamus–frontal cortex). We argue that this pathway plays a central role in the flexibility demonstrated by real organisms. In particular, it is through this pathway that instant motivational manipulations that characterize Pavlovian conditioning are able to affect instrumentally learned behaviors, as in the devaluation experiments now illustrated.

During the two test phases, the satiety of respectively food A or B are kept at one, i.e. at their maximum level (the other satiety level is kept at zero). The satiety for a food causes a strong inhibition to the BLA's hedonic representation of such food. As a consequence both the direct consumption of that food and the perception of the conditioned stimulus previously associated with it cannot elicit the related BLA's hedonic reaction. This is shown by the lack of bla_{foodA}'s activation during the second test phase when the rat is satiated with food A. The perception of both the lever and the chain leads

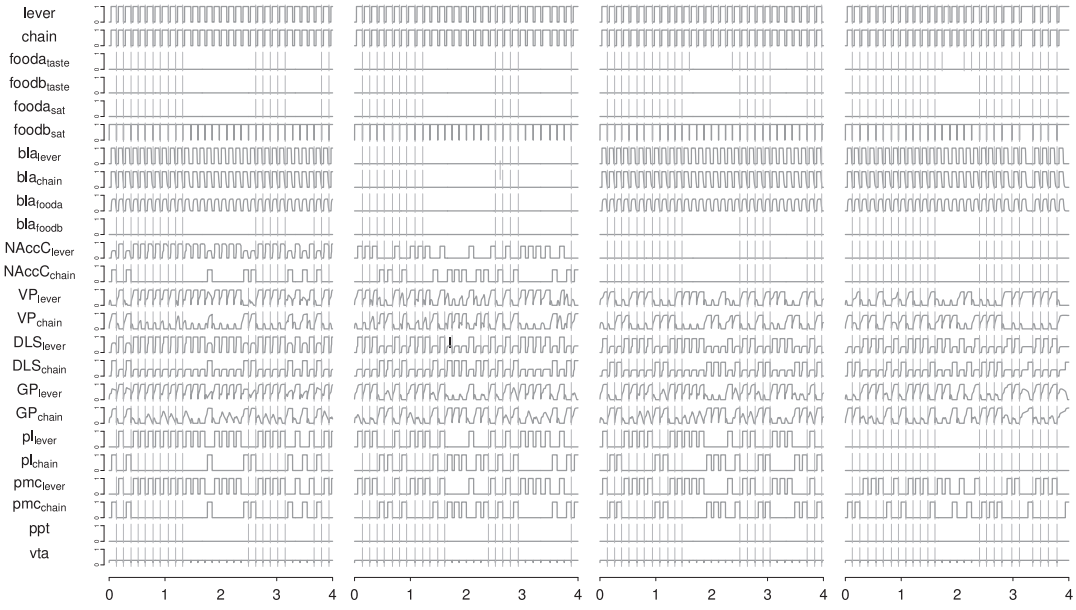


Figure 5.3: *Activations of some key units in four test phases. Each block shows a test phase where satiation for food B was induced. The first block shows activations in a rat with no lesions (SHAM). The second block shows activations in a rat where a lesion to the BLA component was produced before the training session. The third block shows activations in a rat where connections between NAccC and DLS (SPIRALS) were destroyed before trainning. The fourth block shows activations in a rat where connections from PL to PMC (PL-PMC) were lesioned before training. Trials are separated by short vertical lines.*

PUT to 'vote' for both the lever press and chain pull actions at the same time. This rules out the

influences of the S-R instrumental pathway on action selection. Note that this experimental condition was precisely designed by Balleine et al. (2003) to stop the effects of habits that would otherwise 'mask' the motivation-sensitive Pavlovian influence on action selection. On the other hand, satiation stops only one of the two influences of the BLA-NAccC pathway on action selection in that it inhibits only the amygdala representation of the conditioned stimulus which has been satiated (compare the $nacccc_{lever}$ activation in the two test phases). The fact that the BLA-NAccC pathway 'votes' only for the action associated with the non-satiated food breaks the symmetry and makes the related action reliably win the competition in PM (compare the pmc_{lever} and m_{lev} activations in the two test phases).

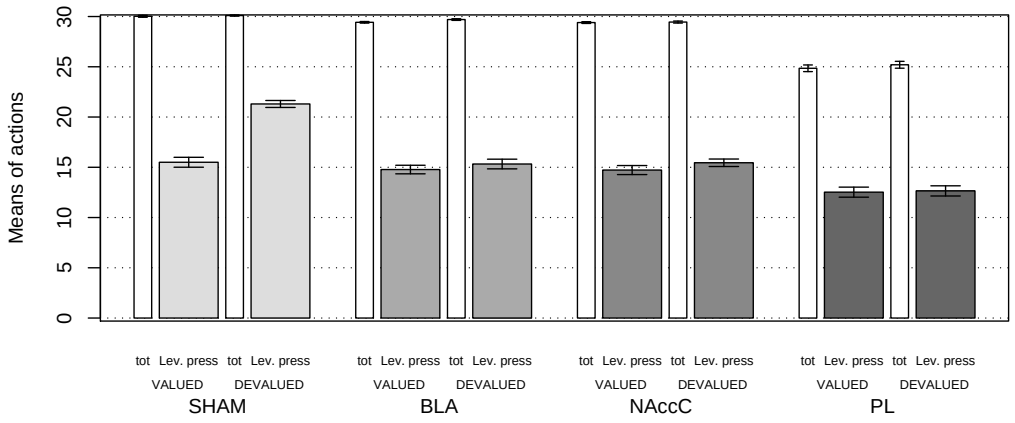


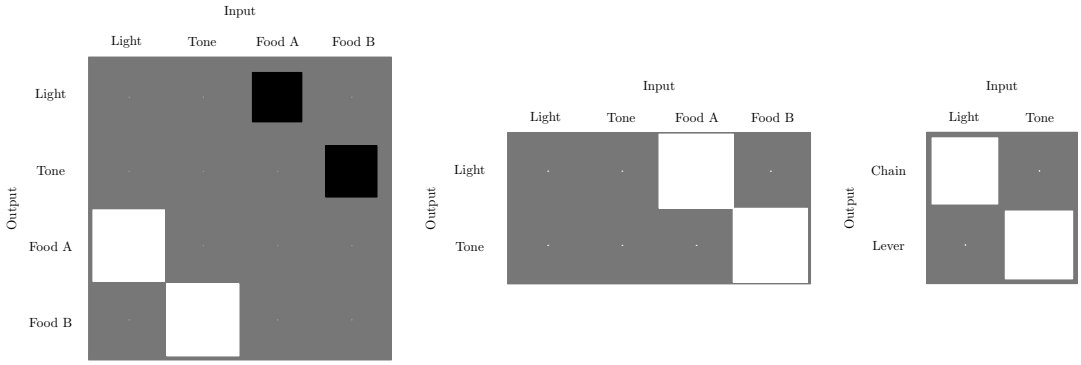
Figure 5.4: *Means of responses to lever during tests where rats were devalued or not for food B (left bars), compared with means of responses of BLA- lesioned rats in the same tests (right bars). Rats with BLA lesions (BLA), lesions of NAccC (NAccC), lesions of PL (PL), or no lesions (SHAM) are compared. Standard errors and total responses are also shown*

The comparison between the lesioned and non-lesioned conditions (see Figure 5.4) reproduces the basic finding of the target experiment of Balleine et al. (2003) and confirms the aforementioned interpretation of the devaluation tests: as it happens in real rats, a lesion to the BLA pathway linking the amygdala to the NAccC prevents the devaluation of food from having any effect on the action selection process. More in particular, during the four minutes of test non-lesioned (SHAM) rats perform the action associated to the non-devalued (ND) food 21.3 times on average whereas they perform the action associated to the devalued (D) food 8.8 times on average ($t = -10.2203, df = 39, p-value < 0.001$). On the contrary, BLA-lesioned (BLA) rats select actions randomly: the averages of performed actions associated with the non-devalued and the devalued foods are respectively 15.325 and 14.375 ($t = -0.772, df = 39, p-value > 0.05$). Furthermore the findings of Corbit et al. (2001) and Corbit and Balleine (2003) about lesions of NAccC an PL on instrumental devaluation are also confirmed (see Fig. 5.4).

Lesions to NAccC or PL prevents in simulated as in real rats prevents the devaluation of food from having any effect on the action selection process.During the four minutes of test, NAccC-lesioned (NAccC) rats select actions randomly: the averages of performed actions associated with the non-

devalued and the devalued foods are respectively 15.45 and 14.00 ($t = -1.3746, df = 39, p - value > 0.05$). Also PL-lesioned (PL) rats select actions randomly: the averages of performed actions associated with the non-devalued and the devalued foods are respectively 12.65 and 12.55 ($t = -0.1701, df = 39, p - value > 0.05$).

These results show the plausibility of the hypothesis for which the BLA-NAccC pathway bridges the Pavlovian processes happening in the amygdala with the instrumental processes happening in the cortex-basal ganglia pathway, so allowing the current state of animals' motivational systems to modulate *on the fly* their action selection mechanisms.



(a) (b) (c)

Figure 5.5: *In (a), (b) and (c) the maps of weights of learned connections within the model are showed. White and black squares respectively correspond to positive and negative weights. The areas of squares are proportional to the absolute values of the corresponding weights. (a) shows weights connecting all units within BLA. (b) shows weights of the connection between BLA and NAccC. (c) shows weights of the connection between IT and DLS.*

5.5 Conclusions and Future Work

This paper presented an embodied model of some important relations existing between Pavlovian and instrumental conditioning. The model's architecture and functioning was constrained with relevant neuroscientific knowledge on the brain anatomy and physiology. The model was validated by successfully reproducing the primary outcomes of some instrumental conditioning devaluation tests conducted with normal and amygdala-lesioned rats. These tests are particularly important for studying the Pavlovian- instrumental interplay as they show how the sensitivity to motivational states exhibited by the Pavlovian system can transfer to instrumentally acquired behaviors.

To the best of the authors' knowledge, the model represents the first attempt to propose a comprehensive interpretation of the aforementioned phenomena, tested in an embodied model. The works most closely related to this one are those of Armony et al. (1997), Dayan and Balleine (2002), Morén and Balkenius (2000), and O'Reilly et al. (2007). The model presented here differs from these works in that it proposes an embodied model (absent in all mentioned researches), presents a fully developed model (Dayan and Balleine, 2002, presented only a 'sketched' model), and tackles the issue of the relations existing between Pavlovian and instrumental conditioning (Armony et al., 1997, Morén and Balkenius, 2000, and O'Reilly et al., 2007, focussed only on Pavlovian conditioning).

Notwithstanding the proposed model has these several strengths, it will be improved along many directions in future work.The first limit of the work is that the model was tested with an embodied system where input signals were heavily pre-processed before being fed into the model in the form of 'localistic representations' (one neuron-one object), and where actions could be specified at a rather abstract level by relying on hardwired low-level behavioral routines. In the future the whole model, or some of its parts (e. g. the amygdala component), will be tested with more challenging embodied systems where the model will be fed with realistic distributed input patterns (e.g., the activations of retina's pixels) and will be required to issue low-level motor commands (e.g., the desired displacement and turning speed).Second, the model has several limitations with respect to available biological evidence. For example, it does not learn to inhibit the dopamine signal at the onset of the USs if these are preceded by CSs, as it happens in real organisms (Schultz, 2002). This prevents the model from performing 'extinction' (i.e., to un-learn a classical conditioning association or an instrumental response if these are not followed anymore by a reward) and from stopping the weights' update. In future work, the model will be added this capability by drawing ideas from other works, for example O'Reilly et al. (2007). Moreover, the model cannot reproduce classical- conditioning based modulation of the *vigor* with which instrumental actions are performed (Niv et al., 2006), nor it is capable of triggering innate actions on the basis of classical-conditioning (e.g. approaching an US, or approaching a CS after this has been associated to an US; Dayan and Balleine, 2002). Finally, the model assumes that the selection ofactions takes place within premotor cortex. However, there is strong evidence (Redgrave et al., 1999) that in real brains action selection takes place at the level of the DLS itself, and so PM activations might only reflect such selection without causing it (cf. Cisek, 2007). This possibility, however, opens up the problem of how the NAccC might influence such action selection, as requested for the Pavlovian processes to exert an influence on instrumental processes. In this respect, an interesting neural pathway through which this influence might be implemented are the striato-nigro- striatal connections (or 'dopaminergic spirals'; Haber et al., 2000). These topics will be addressed in future work.

Notwithstanding these limitations, the proposed model represents an important step in the construction of an integrated picture on how animals' motivational systems can both drive instrumental learning and directly regulate behavior. Constructing such a picture is of paramount importance from the scientific point of view as psychology and neuroscience have now amassed a large body of evidence and knowledge on the phenomena investigated here which would greatly benefit of theoretical systematization. As mentioned in Sect. 5.1, although this papers has mainly a scientific relevance, the research agenda of the work presented here has also a potential interest for overcoming the limited autonomy of current robots. In fact, a way to tackle these limits is to attempt to understand the mechanisms underlying organisms' behavioural flexibility so as to use them in designing robot's controllers. In this respect, notwithstanding the motivational and emotional regulation of behavior is very important for behavioural flexibility, it has been almost completely overlooked by autonomous robotics. For this reason Parisi (2004) has advocated the need of an 'Internal Robotics' research agenda dedicated to the study of these processes.In line with this, recently machine learning and robotics communities have been devoting increasing efforts to the study of autonomous learning by trying to improve the standard reinforcement learning algorithms mentioned in Sect. 5.1 on the basis of ideas coming from the study of real organisms (Weng et al., 2001; Zlatev and Balkenius, 2001). In this respect, the

investigations on emotional regulation of learning and behaviour in animals, such as those reported here, are expected to produce important insights on possible new principles and techniques to be used to design more powerful learning algorithms exhibiting a degree of autonomy similar to that of real organisms (see Barto et al., 2004, and Schembri et al., 2007, for two examples of this).

Chapter 6

Conclusions

The approach of all work described in this thesis consisted in employing tools from computational neuroscience in order to explain data acquired by psychobiological research on emotions and motivations and to furnish a strong operational theoretical framework to interpret them.

The achievements of the research presented in the thesis can be grouped into two areas. First, an operational hypothesis was given on the functional processes taking place within amygdala and their interactions with the other functional brain systems by collecting the neuroscientific data and analyzing it through "computational" lens (see chapter 2). Second, this theoretic framework was used to build computational models of some of the systems centered on amygdala processing. The use of computational models allowed furnishing specific computational hypotheses about: (*1*) how different associative learning mechanisms are implemented within the amygdaloid system, (*2*) how such mechanisms elicit the activation of unlearned responses to the environment, (*3*) how such mechanisms bias cognitive processes of choice and decision making . Specific models were developed in chapters 3), 4), and 5 to investigate all these processes.

Specifically, the model presented in chapter 3 explains how Pavlovian mechanism add flexibility to unlearned behaviours, allowing internal body states to modulate the internal representations of the stimuli through which the cue-guided navigation behaviour (one of the most important unlearned behaviours) is triggered. Such mechanisms allow a rat to navigate towards a region of space where it expects to find a particular resource, but not towards another region where it expects to find a second resource, depending on the current needs for the two resources. Furthermore, the model shows the role these Pavlovian mechanisms in focusing attention. Biasing the navigation behaviours, the Pavlovian mechanisms within amygdala contribute to bring under the focus of attention specific portions of the world depending on internal needs. The importance of the contribute of this chapter resides in that the specific mechanisms underlying unlearned navigation behaviours are not yet fully understood.

The model presented in chapter 4 about internal associative Pavlovian mechanisms implements the hypothesis that the association between conditioned stimuli (CS) and unconditioned responses (UR) formed in classical conditioning experiments is due to two related but distinct mechanisms: (1) stimulus-stimulus associations (CS-US-UR) involving unconditioned stimuli (US) stored in the BLA; (2) direct stimulus-response associations (CS-UR) stored in the LA-CeA neural pathway. The importance of this investigation resides in the fact that the relations between these two associative processes and their location within amygdala is not yet fully understood.

The model described in chapter 5 about the interaction between Pavlovian mechanisms implemented within the amygdala and cognitive processes implemented within striatocortical loops furnishes a computational hypothesis on how Pavlovian mechanisms can bias instrumental actions in order to produce goal-directed behaviour. The novelty of this investigation resides in the fact that, although much data has been furnished on goal-directed behaviours, habit behaviours, and Pavlovian

processes, an overall picture of their relations is emerging only now and computational models can greatly help this synthesis effort.

Future work related to this research will include an investigation in three directions. First, an exploration on the very nature of the reward signal: is it genetically determined? Is it built, at least in part, during the first stages of life? Second an analysis of the mechanisms that, starting from the processing of incentive salience, produce both plasticity of neural populations leading to learning and amplification of the general activity of the agent. Third, two neural subsystems should be investigated at a lower computational level: on one side, the system including the basolateral complex of amygdala, the orbitofrontal cortex, their strict interconnections and the relationship with the associative learning features of amygdala; on the other side, the mechanism defined by the heavily reentrant connections existing between the medial complex of amygdala and the ventromedial hypothalamus, that could underly the very mechanisms of modulation of the incentive value.

The embodiment of simulations should also be improved so as to achieve a more realistic reproduction of the environment, the body of the subjects of the target experiments and the interactions between them. Furthermore, the sensory processing of the simulated organisms should be improved in order to face situations in which learning occurs in the presence of stimuli with partially overlapping features.

Appendix A

Neural networks

A.1 Natural and artificial neural networks

Artificial neural networks can be considered as simplyfied models which capture the essence of the functioning of the brain, and, more generally, of the nervous system. The basic, fundamental units of the nervous system are *neurons*, special types of cells capable of trasmitting elettrical signals. The number of neurons in the human brain is about 10^{11}-10^{12}, and each neuron is connected to about 10^3-10^4 other neurons. There are a number of different kinds of neurons, but there is a general structure that underly all of them (except for a few rare subtypes of neuron, such as analog neurons in the mammalian retina). This structure can be divided in four parts, namely the dendrites, the soma, the and the axon. The electrical signal emmitted by neurons are called action potentials or spikes and are constituted by rapid, binary, electical impulses propaging mostly through axons. When the action potential reaches the end of the axon, it triggers the emission of some chemicals, called neurotransmitters, which are released in the space between the axon and the dendrite of another neuron, the synaptic cleft (the synapses are regions where neurons are connected). The neurotrasmitters bind to the receptors of the post-synaptic neuron and cause, through a chain of events, either the depolarization or the hyperpolarization of the membrane of the receiving neuron. A depolarization corresponds to an excitation in that it favors the emission of a spike in the post-synaptic neuron, while a hyperpolarization corresponds to an inhibition in that it oppose spike emission. Changes in the polarization of the neurons propagate passively from the dendrites to the cell body, where their effects are integrated. If at the origin of the axon the depolarization reaches a certain threshold, an action potential is generated. After the spike, there is a brief refractory period in which the neuron is slightly hyperpolarized and cannot generate another action potential (for a detailed account, see Kandel, Schwartz, and Jessel, 2000). Generally speaking, an artificial neural network is a collection of artificial neurons, units or nodes, linked to each other by connection weights. There are several classes of neural models, that simulate neurons and neural networks at any scale and with any level of sophistication: from the detailed models of single neurons that simulate the effects of particular chemicals on ionic channels (which are the mechanisms through which the membran potential changes), to neural networks consisting of several thousand of abstract, idealized neurons (for an overview, see Floreano and Mattiussi, 1996).

Classical connectionist neural models (McClelland and Rumelhart, 1986) represent the state of a neuronat any given moment by its activation, which correspond to the average firing rate of a real neuron. Connection weights are represented by real numbers that corresponds to the number and strength of synapses between two neurons. Weights can be either positive, corresponding to excitatory synapses, or negative, corresponding to inhibitory synapses. A neuron's activity is a function of the sum of the excitatory and inhibitory inputs that comes from all other neurons connected to it. The

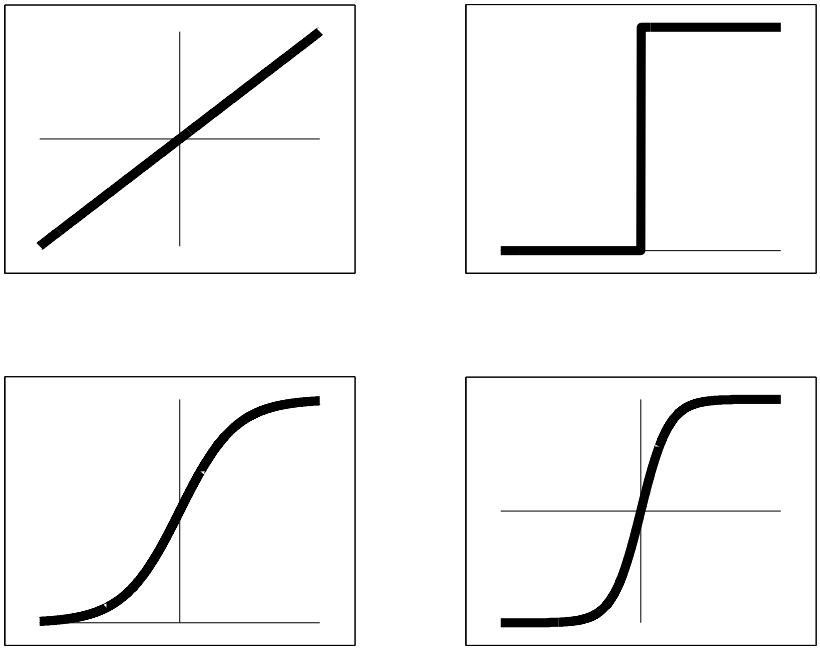


Figure A.1: *Some examples of neurons' activation functions. Linear:* $y = x$*. Step: Sigmoid (or logistic):* $y = \frac{1}{1+e^{-x}}$*. Hyperbolic tangent:* $y = \tanh(x)$*.*

value of each excitation or inhibition arriving to a (post-synaptic) neuron through a connection is in turn calculated by multiplying the activity of the pre-synaptic neuron by the weight of the connection that link the two neurons. The sum of these contributions is called activation potential. Formally, at any given moment, the activation a_i of a given neuron is given by the following formula:

$$a_i = f(\sum w_{ij} a_j)$$

where w_{ij} is the connection weight that connect the j_{th} neuron, with activation a_j, to neuron i, and $f(x)$ is the so called activation function, which determines how the neuron reacts to stimulation. There are a number of possible activation functions, some of which are depicted in figure A.1.

A class of more realistic models take into account the natural decay of the potential leading to the activation of neurons during time. Each i-th unit is defined by a potential p given by a linear ordinary differential equation called leaky integrator (Amari, 1977, see):

$$\tau \dot{p_i} = -p_i + I + \sum_j w_{ij} f\left(p_j\right)$$

where τ is the decaying rate of the unit, I is the external input, w_{ij} is the value of the connection weight from the j-th unit to the i-th unit and $f\left(w_{ij} p_j\right)$ is a function of the sum of the inputs from other units, the activation of each unit $f\left(p_i\right)$ being defined, for example, as:

$$a_i = \begin{cases} 0 & \text{if } p \leq 0 \\ \tanh p_i & \text{if } p_i > 0 \end{cases} .$$

The units of these models (commonly known as dynamic "firing-rate" or "population rate" models

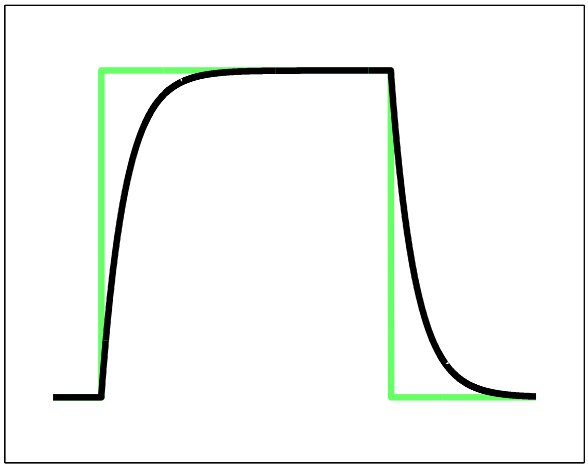


Figure A.2: *an example of the activation of a firing-rate unit (black line) in response to an input (green line).*

(Amari, 1977; Gerstner and Kistler, 2002b)) are intended to represent the mean activity of populations (or fields) of real neurons (see Fig. A.2).

Models of the activity of neurons can be even more sophisticated, allowing the reproduction of the spiking activation of single neurons. One way of doing it is the "integrate-and-fire" neuron model (Gerstner and Kistler, 2002b), built adding a threshold th_p to the amplitude of a leaky-integrator function. When the potential p gets over the threshold the activation of the neuron is set to its maximum and the potential p is reset to a value min_p under its baseline (see Fig. A.3):

$$a_i = \begin{cases} 0 & \text{if } p_i \leq 0 \\ \tanh p_i & \text{if } p_i > 0 \text{ and } p_i < th_p \\ th_p & \text{if } p_i >= th_p \end{cases}$$

$$p_i = \begin{cases} p_i & p < th_p \\ min_p & \text{if } p_i >= th_p \end{cases} .$$

Other more complex neuron models can reproduce the complex behaviors of different types of neurons through compact dynamic systems (e.g. Izhikevich, 2004) or taking into account the interactions between the internal ionic currents (the Hodgkin-Huxley model and its derivated, see Gerstner and Kistler (2002b)).

All simulations within this thesis are based on population firing-rate models. This level of abstraction has been chosen because the research presented here aims at the exploration of neural substrates of behaviour at a system level. This perspective implies two classes of constraints: (*1*) first, the technical difficulties to be faced in order to model each of the single components of the studied systems, would go beyond the scope of our study. For example, by following a more detailed approach the onset activation of units within the models of amygdala presented here (see sections 4.3, 3.4 and 5.3) should be reduced to the real interactions between single neurons that produce that behaviour. (*2*) Second, the computational power needed to implement the models presented here at the level of spiking neurons should not allow us to analyze the properties of the entire system within a reasonable research

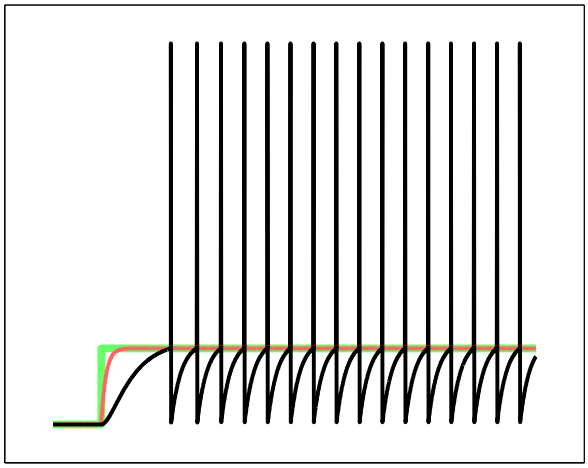


Figure A.3: *An example of the activation of a integrate-and-fire unit (black line) in response to an input (green line). The red line indicates a leaky-integrator function of the input, modelling the summation of all synaptic currents.*

time .Furthermore, some of the issues to be faced at the single-neuron level of modeling could reveal to be independent, that's our hope, from their abstraction at the level of populations. Again, the onset activation of the units within amygdala can be a good example as the solution that can be found by reproducing it with spiking neurons does not seem to depend on the synchronization between the spikes of different populations, but on the changes in the mean firing rates of neurons (data not shown).

A.2 Learning algorithms

The way a neural network responds to inputs depends on (a) its architecture and (b), the connection weights. If some of the weights of a given neural network are changed, then its behavior will change. A fundamental feature of neural networks is their capacity to learn, that is, their capacity to adjust connection weights in such a way that the overall behavior gets better (according to some criterion)[1]. A number of learning algorithms have been developed in the literature. Here we will focus on one family of them, the hebbian learning models.

In his most-famous book *The Organization of Behavior*, Donald Hebb proposed a possible rule for synaptic modification according to which : "*When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased*" (Hebb, 1949). In the connectionist research is termed the Hebb(ian) rule and it is often paraphrased as "Neurons that fire together wire together", In mathematical terms:

$$\Delta w_{ji} = \lambda x_i y_j \ , \tag{A.1}$$

[1]The change in connection weights of a neural network corresponds to the increase or decrease of the number of synaptic connection (and their efficacy) that happens between two connected neurons of the real brain due to brain activity. These changes are the mechanisms that underlie brain's plasticity, that is, the capacity of brain to continually adapt to new circumstances.

where Δw_{ij} represents the change of the weight that connects neuron j with neuron i, x_i and y_j are the activations of neurons i and j, respectively, and λ is the learning rate, typically a number in the range $[01]$ which determines the rate of change of the connection. It is a correlation learning rule: concurrent activation of neurons strengthen the connection between them. It has been proven (see Linsker, 1988) that this rule is a maximizer of the output variance, and plays an important role as an abstract input-output mapping in information- theory.

Such a setup has been largely used in the early formulations of neural networks to deal with problems where no desired target is a- priori known (unsupervised learning). Self- organization of activations of neurons have been studied mainly with a focus on the capacity of such networks to perform pattern associations. Indeed, using the basic Hebbian learning rule (A.1) it is sufficient that the set of inputs are mutually orthogonal to let a standard single-layered architecture with a linear activation function (perceptron) a good associator. By the way mutual orthogonality of input patterns may be biologically implausible, as well as mathematically unprovable. Extreme interference (and subsequent growth blow-up) may occur whenever noisy, overlapping or incomplete stimuli are presented since changes are accumulated all over the training phase. Therefore, the network may not be able to learn associations. This limitation of Hebbian learning can be overcome by modifying both the learning rule and the architecture of the network itself.

The learning rule (A.1) can be modified in different ways. Mathematically speaking, the learning rule (A.1) may be seen as a simplified form of a general law

$$\Delta w_{ji} = F(w_{ji}, x_i, y_j) \tag{A.2}$$

where F is a suitable function (see Gerstner and Kistler, 2002a,b). A straightforward way to control the dangerous effects of unbounded growth is to plug in the formula a "forgetting" or "memory decay" term:

$$\Delta w_{ji} = \lambda x_i y_j - \eta w_{ji} \tag{A.3}$$

where η is the decay factor. On the other hand (A.3) forces every connection to collapse to a null baseline when the input is absent.

Another way of control divergence is to assume a normalization. That may be physiologically motivated by the boundedness of some factors involved in the synaptic competition (such as the number of receptor molecules, the surface area of the postsynaptic membrane or the energy resources). In this case we have the following discrete formulation

$$\Delta w_{ji} = \frac{w_{ji} + \lambda x_i y_j}{\sqrt{\sum_k \left(w_{jk} + \lambda x_k y_j\right)^2}} - w_{ji} \tag{A.4}$$

in such a way that the Euclidean norm of the weights is set to 1 on each time step.

Another Hebbian rule named after Oja (see Oja, 1982) is given by:

$$\Delta w_{ji} = \lambda y_j \left(x_i - y_j w_{ji}\right) \tag{A.5}$$

In this case, if the input sequence is regarded as a stochastic process, then the output is able to extract

the statistically most significant factor. That is, it can be related with a statistical technique called principal component analysis (PCA) or "Karhunen-Loéve feature extraction".

The sign postulated by the initial conjecture may be switched to $-$ as well. That is the case of anti-Hebbian rule:

$$\Delta w_{ji} = -\lambda x_i y_j \tag{A.6}$$

introduced (see Földiák, 1990; Lisman, 1989) to describe the dynamics of excitatory and inhibitory (EXIN) networks. Indeed, adding lateral inhibitions makes the inhibitory component of the network beeing capable of minimizing cross-correlations and decorrelating associated output activations. That capacity improves the competition between nodes(see Marshall, 1995; Spratling and Johnson, 2002) [2]

A.2.1 Differential Hebbian Learning

Spike Timing Dependent Plasticity

Spike timing of neurons plays an important role in the synaptic plasticity, as recent studies and technologies have proven (see Abbott and Nelson, 2000; Caporale and Dan, 2008). Firstly, the sequential order of spikes plays an important role: presynaptic spikes preceding postsynaptic spikes or postsynaptic spikes preceding presynaptic spikes (known as $post$-pre or pre-$post$ spiking, respectively) may trigger quite different effects.

Pioneer studies on the hippocampus shown that pre-$post$ spiking causes long-term potentiation (LTP) of the synapse (see Bliss and Gardner-Medwin, 1973; Bliss and Lømo, 1973; McNaughton, 2003). Similar results have been found in neocortical areas (see Artola and Singer, 1987), in the amygdala (see Chapman, Kairiss, Keenan, and Brown, 1990; Clugnet and LeDoux, 1990), and in the midbrain reward circuit (see Liu, Pu, and Poo, 2005). In this type of experiments LTP was obtained by either the lonely high- frequency stimulation (HFS) of the presynaptic neuron or by low-frequency stimulation (LFS) with large (overshoot) postsynaptic depolarization.

In the inverse situation, in which inputs signals follow target spiking, a long-term depression (LTD) has been observed. Such a behavior was observed in hippocampus (see Debanne, Gähwiler, and Thompson, 1994; Dudek and Bear, 1992), in neuromuscular (see Dan and Poo, 1992). In this type of experiments LPD was obtained by either the lonely low-frequency stimulation (LFS) of the presynaptic neuron or by pairing it with small (undershoot) postsynaptic depolarization.

Currently many of the cellular mechanisms that are involved in LTP have been unveiled. The main responsible components are shown to be the postsynaptic N-methyl D-aspartate (NMDA) receptors, which are are highly sensitive to the membrane potential. Due to their high permeability for calcium,

[2] If a Taylor expansion of the general Hebbian rule (A.2) is considered (with respect to the learning rate λ around 0) we have the following relation

$$\begin{aligned}\Delta w_{ji} = {} & c_0(w_{ji}) + \\ & c_1^{post}(w_{ji}) y_j + c_1^{pre}(w_{ji}) x_i + \\ & c_2^{post}(w_{ji}) {y_j}^2 + c_2^{pre}(w_{ji}) {x_i}^2 + c_2^{corr}(w_{ji}) x_i y_j \ ,\end{aligned} \tag{A.7}$$

which contains all the modifications presented above, if the functions $c_h^x(w_{ji})$ are set in a suitable way ($h = 0, \ldots, 2$ is the degree of the term, pre stands for $presynaptic$, $post$ stands for $postsynaptic$, and $corr$ stands for the $correlation$).

they generate a local chemical signal that is largest when the back-propagating action potential (BAP) in the dendrite arrives shortly after the synapse was active (pre-$post$ spiking). Large postsynaptic calcium transients are known to trigger LTP. The mechanism for LTD is less well understood, but is thought to involve inactivation of ion channels.

The temporal difference between the offset of the presynaptic signal and the onset of the postsynaptic one, that is the length of interstimulus interval (ISI), has been recently shown to be crucial for associative plasticity (see Levy and Steward, 1983). Moreover, the timing sensitivities are on the order of milliseconds, that is only a fixed temporal window lets synaptic plasticity to occur. The studies of spike timing dependent plasticity (STDP) have initially focused on the changes in synaptic potentials and activations of neurons, rather than analyzing the electrochemical mechanisms underlying those behaviours. STDP was firstly detected in neurons in the neocortex (Markram, Lübke, Frotscher, and Sakmann, 1997). Dual patch clamping techniques were used to repetitively activate pre-synaptic neurons 10 milliseconds before the postsynaptic target neurons, and the strength of the synapse turned out to increase. When the activation order was reversed so that the presynaptic neuron was activated 10 milliseconds after its postsynaptic target neuron, the strength of the pre-to-$post$ synaptic connection decreased. The phenomenon was observed later in the cerebellum (Bell, Han, Sugawara, and Grant, 1997) and in various other preparations, with a lot of variations in the time-window and the shape of the curve of plasticity with respect to the spiking timing.

STDP may be seen a differential Hebbian learning because the plasticity processes underlying registrations in all these studies depend on temporal difference between changes of the presynaptic signal and changes of the postsynaptic one.

Models of differential hebbian learning

The Hebbian rule (A.1) and its modifications model the efficacy of synaptic transmission (known as synaptic plasticity) in such a way that if the presynaptic activation persistently concur to cause the postsynaptic target neuron to increase its activation their synaptic efficiency is modified (increased in the plain rule) in the long run. Following the very same rationale proposed by Hebb's conjecture, rather than concurrent levels of activations, concurrent changes of activations underlie the learning mechanism (see Klopf, 1986; Kosko, 1986). Modelling Hebbian learning in terms of variations fits more realistically with the contingency features requested for learning to take place in animals. That is straightforward in classical conditioning, where the learning of asynchronous signals (the conditioned and the unconditioned one) has to be achieved (see Sutton and Barto, 1981, 1990; Sutton, Barto, and Watkins, 1989).

Mathematically speaking, the differential Hebbian rule corresponding to the standard Hebbian rule (A.1) is given by:

$$\Delta w_{ji} = \sigma x'_i y'_j \,, \tag{A.8}$$

where the derivatives replaces the activations. It turns out that differential Hebbian rule is a covariance[3] learning rather than a correlation learning (Choi, 2003, 2006), and it is related to the problem

[3]Sejnowski and Tesauro (1989) have suggested a learning rule named "covariance" rule, of the form $\Delta w_{ij} = \eta \left(x_i - \overline{x}_i\right)\left(y_j - \overline{y}_j\right)$, with its pre-synaptic version: $\Delta w_{ij} = \eta x_i \left(y_j - \overline{y}_j\right)$, and its post-synaptic version: $\Delta w_{ij} = \eta \left(x_i - \overline{x}_i\right) y_j$ where $\overline{x}_i$ and $\overline{y}_j$ represent the mean activations of the units x_i and y_j, respectively. The use of those rules

of independent component analysis (ICA).

Another interesting approach is the isotropic sequence order (ISO) learning, whose differential learning rule follows:

$$\Delta w_{ji} = \eta x_i y'_j \ , \tag{A.9}$$

where y'_j is the derivative of the output signal (whose activation function is linear), and x_j is the traced input, that is decaying trace of the input signal (see Porr and Wörgötter, 2003; Wörgötter and Porr, 2005). The trace is important in order to define a time window in which an output signal is eligible for learning.

A general differential Hebbian learning rule can be considered:

$$\begin{aligned} \Delta w_{post,pre} = {} & \sigma_\alpha \left[u_{pre}{}'\right]^+ \left[u_{pos}{}'\right]^+ + \sigma_\beta \left[u_{pre}{}'\right]^+ \left[u_{pos}{}'\right]^- + \\ & \sigma_\gamma \left[u_{pre}{}'\right]^- \left[u_{post}{}'\right]^+ + \sigma_\delta \left[u_{pre}{}'\right]^- \left[u_{post}{}'\right]^- + \\ & \eta_\alpha u_{pre} \left[u_{pos}{}'\right]^+ + \eta_\beta u_{pre} \left[u_{pos}{}'\right]^- + \\ & \eta_\gamma \left[u_{pre}{}'\right]^+ u_{post} + \eta_\delta \left[u_{pre}{}'\right]^- u_{post} \ , \end{aligned} \tag{A.10}$$

where the coefficients σ's refer to the products between derivatives (the subindeces denotes the signs of the parts considered) and the coefficients η's refer to the mixed products between a signal and a derivative. Since for any function f the basic relation $f = [f]^+ - [f]^-$ holds, (A.10) is a generalization of the previous ones, since if $\sigma_\alpha = -\sigma_\beta = -\sigma_\gamma = \sigma_\delta = \sigma$, $\lambda = 0$ and η's are also null the Kosko learning rule is obtained, while if the σ's are null, $\lambda = 0$, $\eta_\gamma = \eta_\delta = 0$ and $\eta_\alpha = -\eta_\beta$'s the ISO learning rule is obtained;

This general rule can be parametrized to obtain different behaviours. In figure A.2.1 three examples are shown. On the left, column A shows a comparison between registrations of changes in the amplitude of EPSPs at several time differences between signals, made on a slice culture from rat visual cortex (top figure) (Froemke and Dan, 2002) and the corresponding curve of variation of weights in the model suitably parametrized. Column B, at the centre, shows another study done on slice cultures taken from rat hippocampus (Nishiyama, Hong, Mikoshiba, Poo, and Kato, 2000) and the correspondent curve from the model. Finally, column C, shows registrations from different cells in the rat hippocampus (Woodin, Ganguly, and Poo, 2003) and the correspondent data from the model.

This family of learning rules can be used within spiking neuron models given that u_{prev} and u_{post} values correspond to the currents at the level of the presynaptic cleft (u_{prev}) and at the level of the postsynaptic dendrite. Furthermore it can be used within population firing-rate models, given that units have an onset-dependent activation with habituation. This kind of activation can be achieved by computing the unit activation with the following system of leaky integrators (see fig. A.2.1):

guarantees that the weights do not increase indefinitely since they will decrease every time the activation of the sending or receiving neurons is lower than usual (pre- and post- synaptic rules, respectively), or the differences between the mean and the present activations of the two neurons are of different sign (co-variation rule). Though not as powerful as other learning rules, all those variations of the Hebb rule are interesting because they can be viewed in the framework of differential hebb, being within the framework of the Taylor expansion (A.7).

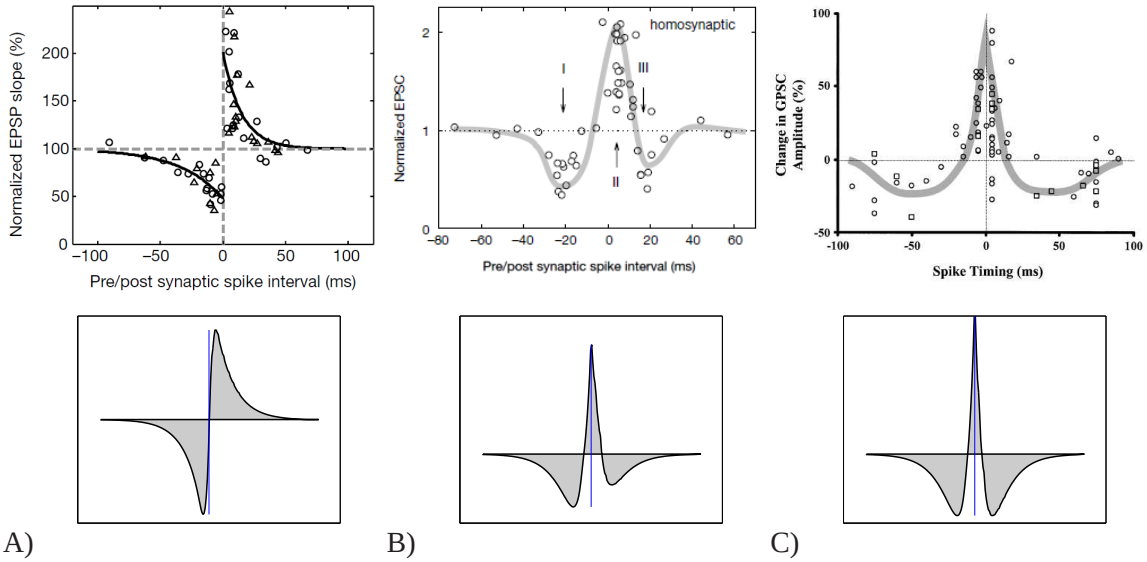


Figure A.4: *examples of modelling STDP data with the use of the learning rule A.10, choosing suitable values of the coefficients. On the top, data from three studies on different cells from different cerebral regions are shown. On the bottom the corresponding data taken from simulations (see text for references)*

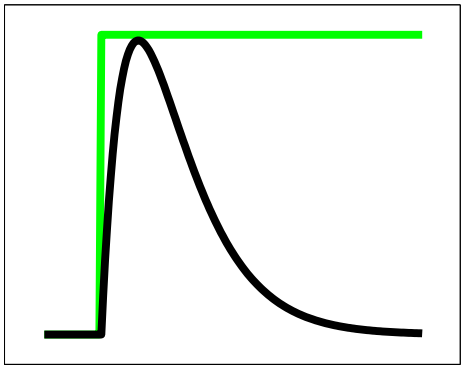


Figure A.5: *The onset-dependent activation described by A.11 is shown in black. The green line defines the input signal.*

$$\begin{aligned} \tau \dot{p}_{int} &= -p_{int} + I \\ \tau \dot{p}_{out} &= -p_{out} + [-p_{int} + I]^{+} \end{aligned} \tag{A.11}$$

This is the way in which differential Hebbian learning is modeled within the amygdala component in all models described in the thesis.

Appendix B

Acronyms

)

Table B.1 presents all the acronyms used in the paper.

Amg				Amygdala
	BLA			Basolateral amygdaloid complex
		LA		Lateral amygdaloid nucleus
			Ld	Lateral dorsal amygdaloid nucleus
				Lda Lateral dorsal amygdaloid nucleus, anterior part
				Ldp Lateral dorsal amygdaloid nucleus, posterior part
			Lvm	Lateral ventromedial amygdaloid nucleus
			Lvl	Lateral ventrolateral amygdaloid nucleus
		BL		Basolateral amygdaloid nucleus
			B	Basal amygdaloid nucleus
			AB	Accessory basal amygdaloid nucleus
	CEA			Central extended amygdala
		CeA		Central amygdaloid nucleus
			CLC	Central amygdaloid nucleus, laterl capsular suddivision
			CL	Central lateral amygdaloid nucleus
			CM	Central medial amygdaloid nucleus
		BNST		Bed Nucleus of the stria terminalis
	MEA			Medial extended amygdala
		MeA		Medial amygdaloid nucleus
			Mv	Medial amygdaloid nucleus, ventral part
			Md	Medial amygdaloid nucleus, dorsal part
	ITC			Intercalated nuclei
BG				Basal ganglia
	STR			Striatum
		DLS		Dorsolateral striatum
		DMS		Dorsomedial striatum
		NAcc		Nucleus accumbens
			NAccC	Nucleus accumbens core
			NAccS	Nucleus accumbens shell
	PAL			Pallidum

		DP	Dorsal pallidum
		GPi	Globus pallidus, internal segment
		VP	Ventral pallidum
CB			Cerebellum
DI			Disgranular insular cortex
	DIv		Disgranular insular cortex, visceral part
	DIg		Disgranular insular cortex, gustatory part
En			Endopiriform nucleus
Hyp			Hypothalamus
	LH		Lateral hypothalamus
	PO		Preoptic nucleus of hypothalamus
	VMH		Ventromedial hypothalamus
	PVN		Paraventricular nucleus of hypothalamus
Hip			Hippocpampus
	S		Subiculum
MB			Midbrain
	VTA		Ventral tegmental area
	PAG		Periaqueductal gray
	SNpc		Substantia nigra, pars compacta
	MEV		Midbrain trigeminal nucleus
	PPT		Pedunculopontine tegmental nucleus
	DR		Dorsal raphe
My			Medulla
	NST		Nucleus of the solitary tract
	AMB		Nucleus ambiguus
	DMX		Dorsal motor nucleus of the vagus nerve
MC			Motor cortex
OB			Olfactory bulb
P			Pons
	PB		Parabrachial nucleus
	LDT		Laterodorsal tegmental nucleus
	LC		Locus coeruleus
	NRPC		Nucleus reticularis pontis caudalis
PaRh			Parietal rhinal cortex
PC			Piriform cortex
PFC			Prefrontal cortex
	AC		Anterior cingulate cortex
	vmPFC		Ventromedial prefrontal cortex
		PL	Prelimbic cortex
		IL	Infralimbic cortex
	OFC		Orbitofrontal cortex
PMC			Premotor cortex

PRC		Perirhinal cortex
SI		Substantia innominata
Te		Temporal cortex
	Te2	Temporal cortex, Area 2
	Te3	Temporal cortex, Area 3
Th		Thalamus
	LG	Lateral geniculate nucleus
	MG	Medial geniculate nucleus
	ILN	Infralaminar nucleus
	VPMpc	Ventral posteromedial nucleus, parvicellular part

Table B.1: Acronyms used throughout the paper to refer to the anatomical brain areas of interest.

Bibliography

Abbott, L. F. and S. B. Nelson (2000, Nov). Synaptic plasticity: taming the beast. *Nat Neurosci 3 Suppl*, 1178–1183.

Ahn, S. and A. G. Phillips (2002, Dec). Modulation by central and basolateral amygdalar nuclei of dopaminergic correlates of feeding to satiety in the rat nucleus accumbens and medial prefrontal cortex. *J Neurosci 22*(24), 10958–10965.

Alvarez, P. and L. R. Squire (1994, Jul). Memory consolidation and the medial temporal lobe: a simple network model. *Proc Natl Acad Sci U S A 91*(15), 7041–7045.

Amari, S.-i. (1977, Aug). Dynamics of pattern formation in lateral-inhibition type neuralfields. *Biol Cybern 27*(2), 77–87.

Amat, J., M. V. Baratta, E. Paul, S. T. Bland, L. R. Watkins, and S. F. Maier (2005, Mar). Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nat Neurosci 8*(3), 365–371.

Amat, J., E. Paul, L. R. Watkins, and S. F. Maier (2008, Jul). Activation of the ventral medial prefrontal cortex during an uncontrollable stressor reproduces both the immediate and long-term protective effects of behavioral control. *Neuroscience 154*(4), 1178–1186.

Amat, J., E. Paul, C. Zarza, L. R. Watkins, and S. F. Maier (2006, Dec). Previous experience with behavioral control over stress blocks the behavioral and dorsal raphe nucleus activating effects of later uncontrollable stress: role of the ventral medial prefrontal cortex. *J Neurosci 26*(51), 13264–13272.

Arbib, M. A. and J.-M. Fellous (2004). Emotions: from brain to robot. *Trends Cogn Sci 8*(12), 554–561.

Arias, C. and M. G. Chotro (2007, Mar). Amniotic fluid can act as an appetitive unconditioned stimulus in preweanling rats. *Dev Psychobiol 49*(2), 139–149.

Arleo, A. and W. Gerstner (2000, Sep). Spatial cognition and neuro-mimetic navigation: a model of hippocampalplacecell activity. *Biol Cybern 83*(3), 287–299.

Armony, J. L., D. Servan-Schreiber, L. M. Romanski, J. D. Cohen, and J. E. LeDoux (1997, Mar). Stimulus generalization of fear responses: effects of auditory cortex lesions in a computational model and in rats. *Cereb Cortex 7*(2), 157–165.

Arnold, M. B. (1960). *Emotion and personality. Psychological aspects*, Volume 1. New York, Academic Press.

Artola, A. and W. Singer (1987). Long-term potentiation and nmda receptors in rat visual cortex. *Nature 330*(6149), 649–652.

Aston Jones, G. and J. D. Cohen (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci 28*, 403–450.

Avila-Garcia, O. and L. Cañamero (2004). Using hormonal feedback to modulate action selection in a competitive scenario. In S. Schaal, A. J. Ijspeert, A. Billard, S. Vijayakumar, J. Hallam, and J. A. Meyer (Eds.), *From animals to animats 8: Proceedings of the 8th international conference on the simulation of adaptive behavior*, Cambridge, MA, pp. 243–252. MIT Press.

Balkenius, C. (1993). The roots of motivations. In J. A. Mayer, H. L. Roitblat, and S. W. Wilson (Eds.), *From Animals to Animats II: Proceedings of the Second International Conference on Simulation of Adapt Behav*, Cambridge, MA. MIT Press.

Balkenius, C. and J. Morén (1999, July). Dynamics of a classical conditioning model. *Autonomous Robots 7*(1), 41–56.

Balkenius, C. and J. Morén (2000). A computational model of context processing. In J.-A. Meyer, A. Berthoz, D. Floreano, H. L. Roitblat, and S. W. Wilson (Eds.), *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behaviour*, Cambridge, Mass., pp. 256–265. The MIT Press.

Balleine, B. W. (2005, Dec). Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol Behav 86*(5), 717–730.

Balleine, B. W. and A. Dickinson (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology 37*(4-5), 407–419.

Balleine, B. W. and S. Killcross (2006, May). Parallel incentive processing: an integrated view of amygdala function. *Trends Neurosci 29*(5), 272–279.

Balleine, B. W., S. A. Killcross, and A. Dickinson (2003, Jan). The effect of lesions of the basolateral amygdala on instrumental conditioning. *J Neurosci 23*(2), 666–675.

Bandler, R., K. A. Keay, N. Floyd, and J. Price (2000, Sep). Central circuits mediating patterned autonomic activity during active vs. passive emotional coping. *Brain Res Bull 53*(1), 95–104.

Bar-Gad, I., G. Morris, and H. Bergman (2003, December). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog Neurobiol 71*(6), 439–473.

Bard, A. P. (1934). On emotional expression after decortication with some remarks on certain theoretical views, ii. *Psychol Rev 41*, 424–449.

Barnes, N. M. and T. Sharp (1999, Aug). A review of central 5-HT receptors and their function. *Neuropharmacology 38*(8), 1083–1152.

Barto, A. G. (1995). Models of information processing in the basal ganglia. In J. C. Houk, J. Davis, and D. Beiser (Eds.), *Models of information processing in the basal ganglia*, pp. 215–232. Cambridge, MA: MIT Press.

Barto, A. G., S. Singh, and N. Chentanez (2004). Intrinsically motivated learning of hierarchical collections of skills. In *International Conference on Developmental Learning (ICDL)*, LaJolla, CA.

Baxter, M. G. and E. A. Murray (2002, Jul). The amygdala and reward. *Nat Rev Neurosci 3*(7), 563–573.

Bechara, A., A. R. Damasio, H. Damasio, and S. W. Anderson (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition 50*(1-3), 7–15.

Bechara, A., H. Damasio, A. R. Damasio, and G. P. Lee (1999, Jul). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J Neurosci 19*(13), 5473–5481.

Bechara, A., D. Tranel, H. Damasio, and A. R. Damasio (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb Cortex 6*(2), 215–225.

Bell, C. C., V. Z. Han, Y. Sugawara, and K. Grant (1997, May). Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature 387*(6630), 278–281.

Berridge, C. W. and B. D. Waterhouse (2003, Apr). The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Res Brain Res Rev 42*(1), 33–84.

Berridge, K. C. (2007, Apr). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl) 191*(3), 391–431.

Berridge, K. C. and T. E. Robinson (1998, Dec). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev 28*(3), 309–369.

Berridge, K. C. and T. E. Robinson (2003, Sep). Parsing reward. *Trends Neurosci 26*(9), 507–513.

Blair, H. T., G. E. Schafe, E. P. Bauer, S. M. Rodrigues, and J. E. LeDoux (2001). Synaptic plasticity in the lateral amygdala: A cellular hypothesis of fear conditioning. *Learn Memoryory 8*(5), 229–242.

Blair, H. T., F. Sotres-Bayon, M. A. P. Moita, and J. E. LeDoux (2005). The lateral amygdala processes the value of conditioned and unconditioned aversive stimuli. *Neuroscience 133*(2), 561–569.

Bland, S. T., D. Hargrave, J. L. Pepin, J. Amat, L. R. Watkins, and S. F. Maier (2003, Sep). Stressor controllability modulates stress-induced dopamine and serotonin efflux and morphine-induced serotonin efflux in the medial prefrontal cortex. *Neuropsychopharmacol 28*(9), 1589–1596.

Bliss, T. V. P. and A. R. Gardner-Medwin (1973, July). Long-lasting potentiation of synaptic transmission in the dentate area of the unanaesthetized rabbit following stimulation of the perforant path. *J Physiol 232*(2), 357–374.

Bliss, T. V. P. and T. Lømo (1973, Jul). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol 232*(2), 331–356.

Blundell, P., G. Hall, and S. Killcross (2001, Nov). Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. *J Neurosci 21*(22), 9018–9026.

Blundell, P., G. Hall, and S. Killcross (2003, Aug). Preserved sensitivity to outcome value after lesions of the basolateral amygdala. *J Neurosci 23*(20), 7702–7709.

Brauer, K., M. Häusser, W. Härtig, and T. Arendt (2000, Mar). The core-shell dichotomy of nucleus accumbens in the rhesus monkey as revealed by double-immunofluorescence and morphology of cholinergic interneurons. *Brain Res 858*(1), 151–162.

Cabib, S. and S. Puglisi-Allegra (1994, May). Opposite responses of mesolimbic dopamine system to controllableanduncontrollable aversive experiences. *J Neurosci 14*(5 Pt 2), 3333–3340.

Cabib, S. and S. Puglisi-Allegra (1996, Dec). Stress, depression and the mesolimbic dopamine system. *Psychopharmacology (Berl) 128*(4), 331–342.

Cador, M., T. W. Robbins, and B. J. Everitt (1989). Involvement of the amygdala in stimulus-reward associations: interaction with the ventral striatum. *Neuroscience 30*(1), 77–86.

Calandreau, L., A. Desmedt, L. Decorte, and R. Jaffard (2005). A different recruitment of the lateral and basolateral amygdala promotes contextual or elemental conditioned association in pavlovian fear conditioning. *Learn Memory 12*(4), 383–388.

Cañamero, L. (2005). Emotion understanding from the perspective of autonomous robots research. *Neural Networks 18*(4), 445–455.

Cañamero, L. D. (1997). Modeling motivations and emotions as a basis for intelligent behavior. In W. Lewis Johnson (Ed.), *Proceedings of the First International Conference on Autonomous Agents*, New York, pp. 148–155. The ACM Press.

Cannon, W. B. (1920). *Bodily Changes In Pain Hunger Fear And Rage*. D.Appleton And Company.

Caporale, N. and Y. Dan (2008, Jan). Spike timing-dependent plasticity: a hebbian learning rule. *Annu Rev Neurosci 31*(1), 25–46.

Cardinal, R. N., J. A. Parkinson, J. Hall, and B. J. Everitt (2002, May). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev 26*(3), 321–352.

Cardinal, R. N., J. A. Parkinson, J. Halla, and B. J. Everitt (2003). The contribution of the amygdala, ventral striatum, and prefrontal cortex to emotion and motivated behaviour. In *Cognition and Emotion in the Brain (International Symposium on Limbic and Association Cortical Systems: Basic, Clinical and Computational Aspects).*, Toyama, Japan, pp. 347–370. Elsevier.

Cardinal, R. N., J. A. Parkinson, G. Lachenal, K. M. Halkerston, N. Rudarakanchana, J. Hall, C. H. Morrison, S. R. Howes, T. W. Robbins, and B. J. Everitt (2002, Aug). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behav Neurosci 116*(4), 553–567.

Cardinal, R. N., J. A. Parkinson, H. D. Marbini, A. J. Toner, T. J. Bussey, T. W. Robbins, and B. J. Everitt (2003, Jun). Role of the anterior cingulate cortex in the control over behavior by Pavlovian conditioned stimuli in rats. *Behav Neurosci 117*(3), 566–587.

Cecconi, F. and D. Parisi (1993). Neural networks with motivational units. In J. A. Meyer, H. L. Roitblat, and S. W. Wilson (Eds.), *From animals to animats 2: Proceedings of the 2nd International Conference on Simulation of Adapt Behav*, Cambridge, MA, pp. 346–355. MIT Press.

Chapman, P. F., E. W. Kairiss, C. L. Keenan, and T. H. Brown (1990). Long-term synaptic potentiation in the amygdala. *Synapse 6*(3), 271–278.

Chen, J., M. Nakamura, T. Kawamura, T. Takahashi, and D. Nakahara (2006, Mar). Roles of pedunculopontine tegmental cholinergic receptors in brain stimulation reward in the rat. *Psychopharmacology (Berl) 184*(3-4), 514–522.

Choi, S. (2003). Differential ica. In O. Kaynak, E. Alpaydin, E. Oja, and L. Xu (Eds.), *Proceedings of the Joint International Conference of Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP 2003)*, Volume 2714, pp. 68–75. Springer Verlag.

Choi, S. (2006). Differential learning algorithms for decorrelation and independent component analysis. *Neural Networksorks 19*(10), 1558–1567.

Cisek, P. (2007, Sep). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos Trans R Soc Lond B Biol Sci 362*(1485), 1585–1599.

Clugnet, M.-C. and J. E. LeDoux (1990, Aug). Synaptic plasticity in fear conditioning circuits: induction of ltp in the lateral nucleus of the amygdala by stimulation of the medial geniculate body. *J Neurosci 10*(8), 2818–2824.

Corbit, L. H. and B. W. Balleine (2003, Nov). The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res 146*(1-2), 145–157.

Corbit, L. H., J. L. Muir, and B. W. Balleine (2001, May). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *J Neurosci 21*(9), 3251–3260.

Dagher, A., A. M. Owen, H. Boecker, and D. J. Brooks (2001, May). The role of the striatum and hippocampus in planning: a PET activation study in Parkinson's disease. *Brain 124*(Pt 5), 1020–1032.

Damasio, A. R. (1994). *Descartes Error: Emotion, Reason and the Human Brain*. New York: G.P. Putnams Sons.

Dan, Y. and M.-M. Poo (1992, June). Hebbian depression of isolated neuromuscular synapses in vitro. *Science 256*, 1570–1573.

Davis, M. (1992a). The role of the amygdala in fear and anxiety. *Annu Rev Neurosci 15*, 353–375.

Davis, M. (1992b, Jan). The role of the amygdala in fear-potentiated startle: implications for animal models of anxiety. *Trends Pharmacol Sci 13*(1), 35–41.

Davis, M. and P. J. Whalen (2001, Jan). The amygdala: vigilance and emotion. *Mol Psychiatry 6*(1), 13–34.

Dayan, P. (2002). Motivated reinforcement learning. In T. G. Dietterich, S. Becker, and Z. Ghahramani (Eds.), *Advances in Neural Information Processing Systems 14*, Cambridge, MA, pp. 11–18. MIT Press.

Dayan, P. and B. Balleine (2002). Reward, motivation and reinforcement learning. *Neuron 36*, 285–298.

De Olmos, J. S., C. A. Beltramino, and G. F. Alheid (2004). Amygdala and extended amygdala. In G. Paxinos (Ed.), *The Rat Nervous System*. (third ed.). Elsevier accademic press.

Debanne, D., B. H. Gähwiler, and S. M. Thompson (1994, Feb). Asynchronous pre- and postsynaptic activity induces associative long-term depression in area ca1 of the rat hippocampus in vitro. *Proc Natl Acad Sci USA 91*(3), 1148–1152.

Di Ciano, P. and B. J. Everitt (2004, Aug). Direct interactions between the basolateral amygdala and nucleus accumbens core underlie cocaine-seeking behavior by rats. *J Neurosci 24*(32), 7167–7173.

Dickinson, A. (1985, February). Actions and habits: The development of behavioural autonomy. *Philos Trans R Soc Lond B Biol Sci 308*, 67–78.

Domjan, M. (2006). *The principles of learning and behavior* (Fifth ed.). Belmont, CA: Thomson Wadsworth Pub. Co.

Dudek, S. M. and M. F. Bear (1992, May). Homosynaptic long-term depression in area ca1 of hippocampus and effects of n-methyl-d-aspartate receptor blockade. *Proc Natl Acad Sci USA of the United States of America 89*(10), 4363–4367.

Duffy, E. (1957, Sep). The psychological significance of the concept of arousal or activation. *Psychol Rev 64*(5), 265–275.

Eschenko, O., W. Ramadan, M. Mölle, J. Born, and S. J. Sara (2008). Sustained increase in hippocampal sharp-wave ripple activity during slow-wave sleep after learning. *Learn Memory 15*(4), 222–228.

Filliat, D. and J.-A. Meyer (2003a). Map-based navigation in mobile robots - I. a review of localisationstrategies. *J Cog Sys Res 4*(4), 243–282.

Filliat, D. and J.-a. Meyer (2003b). Map-based navigation in mobile robots - II. a review of map-learningandpath-planing strategies. *J Cog Sys Res 4*(4), 283–317.

Floreano, D. and C. Mattiussi (1996). *Manuale sulle reti neurali*. Bologna: Il Mulino.

Floresco, S. B. (2007, Nov). Dopaminergic regulation of limbic-striatal interplay. *J Psychiatry Neurosci 32*(6), 400–411.

Földiák, P. (1990). Forming sparse representations by local anti-hebbian learning. *Biol Cybern 64*(2), 165–170.

Froemke, R. C. and Y. Dan (2002, March). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature 416*(6879), 433–438.

Fudge, J. L. and A. B. Emiliano (2003). The extended amygdala and the dopamine system: another piece of the dopamine puzzle. *J Neuropsychiatry Clin Neurosci 15*(3), 306–316.

Fudge, J. L. and S. N. Haber (2000). The central nucleus of the amygdala projection to dopamine subpopulations in primates. *Neuroscience 97*(3), 479–494.

Gallagher, M., P. W. Graham, and P. C. Holland (1990, Jun). The amygdala central nucleus and appetitive Pavlovian conditioning:lesionsimpair one class of conditioned behavior. *J Neurosci 10*(6), 1906–1911.

Gerstner, W. and W. M. Kistler (2002a, December). Mathematical formulations of hebbian learning. *Biol Cybern 87*(5-6), 404–415.

Gerstner, W. and W. M. Kistler (2002b). *Spiking neuron models: single neurons, populations, plasticity*. Cambridge: Cambridge University Press.

Gorelova, N. and C. R. Yang (1997, Feb). The course of neural projection from the prefrontal cortex to the nucleus accumbens in the rat. *Neuroscience 76*(3), 689–706.

Grahn, R. E., S. Maswood, M. B. McQueen, L. R. Watkins, and S. F. Maier (1999, Mar). Opioid-dependent effects of inescapable shock on escape behavior and conditioned fear responding are mediated by the dorsal raphe nucleus. *Behav Brain Res 99*(2), 153–167.

Gurney, K., T. J. Prescott, and P. Redgrave (2001a, Jun). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol Cybern 84*(6), 401–410.

Gurney, K., T. J. Prescott, and P. Redgrave (2001b, Jun). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biol Cybern 84*(6), 411–423.

Haber, S. N. (2003, Dec). The primate basal ganglia: parallel and integrative networks. *J Chem Neuroanat 26*(4), 317–330.

Haber, S. N., J. L. Fudge, and N. R. McFarland (2000, Mar). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci 20*(6), 2369–2382.

Han, J. S., R. W. McMahan, P. Holland, and M. Gallagher (1997, May). The role of an amygdalo-nigrostriatal pathway in associative learning. *J Neurosci 17*(10), 3913–3919.

Hatfield, T., J. S. Han, M. Conley, M. Gallagher, and P. Holland (1996, Aug). Neurotoxic lesions of basolateral, but not central, amygdala interfere with pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci 16*(16), 5256–5265.

Hebb, D. O. (1949). *The Organization of Behaviour*. New York: Wiley, John and Sons.

Hiroi, N. and N. M. White (1991, Jul). The lateral nucleus of the amygdala mediates expression of the amphetamine-produced conditioned place preference. *J Neurosci 11*(7), 2107–2116.

Holland, O. and D. McFarland (2001). *Artificial Ethology*. Oxford University Press: Oxford.

Hull, C. L. (1943). *Principles of behavior*. Appleton-century-crofts.

Humphries, M. D., R. D. Stewart, and K. N. Gurney (2006, Dec). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *J Neurosci 26*(50), 12921–12942.

Ito, R., T. W. Robbins, and B. J. Everitt (2004, Apr). Differential control over cocaine-seeking behavior by nucleus accumbens core and shell. *Nat Neurosci 7*(4), 389–397.

Iversen, S., L. Iversen, and C. B. Saper (2000). The autonomic nervous system and the hypothalamus. In E. R. Kandel, J. H. Schwartz, and T. M. Jessell (Eds.), *Principles of neural sciences* (fourth ed.). McGraw Hill.

Izhikevich, E. M. (2004, Sep). Which model to use for cortical spiking neurons? *IEEE Trans Neural Networks 15*(5), 1063–1070.

James, W. (1884). What is an emotion? *Mind 9*, 188–205.

Jolkkonen, E., R. Miettinen, M. Pikkarainen, and A. Pitkänen (2002). Projections from the amygdaloid complex to the magnocellular cholinergic basal forebrain in rat. *Neuroscience 111*(1), 133–149.

Jolkkonen, E. and A. Pitkänen (1998, May). Intrinsic connections of the rat amygdaloid complex: projections originating in the central nucleus. *J Comp Neurol 395*(1), 53–72.

Kandel, E. R., J. H. Schwartz, and T. M. Jessel (2000). *Principles of Neural Science* (4 ed.). New York: McGraw-Hill.

Kelley, A. E. and K. C. Berridge (2002, May). The neuroscience of natural rewards: relevance to addictive drugs. *J Neurosci 22*(9), 3306–3311.

King, B. M. (2006, Feb). The rise, fall, and resurrection of the ventromedial hypothalamus in the regulation of feeding behavior and body weight. *Physiol Behav 87*(2), 221–244.

Klopf, H. A. (1986). A drive-reinforcement model of single neuron function: An alternative to the hebbian neuronal model. *AIP Conference Proceedings 151*, 265–270.

Knapska, E., K. Radwanska, T. Werka, and L. Kaczmarek (2007, Oct). Functional internal complexity of amygdala: focus on gene activity mapping after behavioral training and drugs of abuse. *Physiol Rev 87*(4), 1113–1173.

Kobayashi, Y. and K.-I. Okada (2007, Mar). Reward prediction error computation in the pedunculopontine tegmentalnucleusneurons. *Ann N Y Acad Sci*.

Kosko, B. (1986). Differential hebbian learning. *AIP Conference Proceedings 151*(1), 277–?282.

Kringelbach, M. L. and E. T. Rolls (2004, Apr). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol 72*(5), 341–372.

LaBar, K. S. and R. Cabeza (2006, Jan). Cognitive neuroscience of emotional memory. *Nat Rev Neurosci 7*(1), 54–64.

LaLumiere, R. T., T.-V. Buen, and J. L. McGaugh (2003, Jul). Post-training intra-basolateral amygdala infusions of norepinephrine enhance consolidation of memory for contextual fear conditioning. *J Neurosci 23*(17), 6754–6758.

LaLumiere, R. T., E. M. Nawar, and J. L. McGaugh (2005). Modulation of memory consolidation by the basolateral amygdala or nucleus accumbens shell requires concurrent dopamine receptor activation in both brain regions. *Learn Memory 12*(3), 296–301.

Lange, C. G. (1912). The mechanism of the emotions. In B. Rand (Ed.), *The Classical Psychologists*. Boston: Houghton Mifflin.

Langton, C. (1987). *The First International Conference on the Simulation and Synthesis of Living Systems (ALifeI)*.

Lanuza, E., K. Nader, and J. E. LeDoux (2004). Unconditioned stimulus pathways to the amygdala: effects of posterior thalamic and cortical lesions on fear conditioning. *Neuroscience 125*(2), 305–315.

Lawrence, A. D., L. H. Watkins, B. J. Sahakian, J. R. Hodges, and T. W. Robbins (2000, Jul). Visual object and visuospatial cognition in Huntington's disease: implications for information processing in corticostriatal circuits. *Brain 123 (Pt 7)*, 1349–1364.

Lazarus, R. S., J. R. Averill, and E. M. Opton (1970). Towards a cognitive theory of emotion. In M. Arnold (Ed.), *Feelings and emotions: the Loyola Symposium*. New York, Academic Press.

LeDoux, J. (1996). *The emotional brain: the mysterious underpinnings of emotional life*. New York , NY: Simon and Schuster.

LeDoux, J. (2003). *Synaptic self: how our brains become who we are*. New York, NY: Penguin Books.

Lee, H. J., F. Groshek, G. D. Petrovich, J. P. Cantalini, M. Gallagher, and P. C. Holland (2005, Apr). Role of amygdalo-nigral circuitry in conditioning of a visual stimulus paired with food. *J Neurosci 25*(15), 3881–3888.

Levy, W. B. and O. Steward (1983). Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience 8*(4), 791–797.

Lewis, M. D. (2005, April). Bridging emotion theory and neurobiology through dynamic systems modeling. *Behav Brain Sci 28*(2), 169–194.

Lieberman, D. A. (1993). *Learning, behaviour and cognition*. Brooks/Cole.

Linsker, R. (1988, Mar). Self-organization in a perceptual network. In *Computer*, Volume 21, pp. 105–117. IEEE Computer Society Press.

Lisman, J. E. (1989, Dec). A mechanism for the hebb and the anti-hebb processes underlying learning and memory. *Proc Natl Acad Sci USA 86*(23), 9574–9578.

Liu, Q.-s., L. Pu, and M.-m. Poo (2005). Repeated cocaine exposure in vivo facilitates ltp induction in midbrain dopamine neurons. *Nature 437*(7061), 1027–1031.

Maier, S. F. (1984). Learned helplessness and animal models of depression. *Prog Neuropsychopharmacol Biol Psychiatry 8*(3), 435–446.

Maier, S. F. and L. R. Watkins (2005). Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neurosci Biobehav Rev 29*(4-5), 829–841.

Mannella, F., M. Mirolli, and G. Baldassarre (2007). The role of amygdala in devaluation: a model tested with a simulated robot. In L. Berthouze, C. G. Prince, M. Littman, H. Kozima, and C. Balkenius (Eds.), *Proceedings of the Seventh International Conference on Epigenetic Robotics*, pp. 77–84. Lund University Cognitive Studies.

Mannella, F., M. Mirolli, and G. Baldassarre (2009a). The interplay of pavlovian and instrumental processes in devaluation experiments: a computational embodied neuroscience model tested with a simulated rat. Cambridge: Cambridge University Press. In press.

Mannella, F., M. Mirolli, and G. Baldassarre (2009b). The roles of amygdala in the affective regulation of body, brain and behaviour. *Connection Science*, submitted.

Mannella, F., S. Zappacosta, and G. Baldassarre (2008, Jun). A computational model of the amygdala nuclei's role in second order conditioning. In M. A. J. Tani, J. Hallam, and J.-A. Meyer (Eds.), *Proceedings of the Tenth International Conference on Simulation of Adapt Behav: From Amimals to Animals 10*, Volume 5040, Osaka, Japan, pp. 321–330. Springer.

Maren, S. (2005, Feb). Building and burying fear memories in the brain. *Neuroscientist 11*(1), 89–99.

Markram, H., J. Lübke, M. Frotscher, and B. Sakmann (1997, Jan). Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science 275*(5297), 213–215.

Marshall, J. A. (1995). Adaptive perceptual pattern recognition by self-organizing neural networks: context, uncertainty, multiplicity, and scale. *Neural Networksorks 8*(3), 335–362.

Martinet, L. E., D. Sheynikhovich, J. A. Meyer, and A. Arleo (2009). A cortical column model for spatial navigation planning. *Journées Francophones de Planification, Décision et Apprentissage*.

McClelland, J. L., B. L. McNaughton, and R. C. O'Reilly (1995, Jul). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev 102*(3), 419–457.

McClelland, J. L. and D. E. Rumelhart (1986). *Parallel distributed processing: explorations in the microstructure of cognition*, Volume 2. Cambridge,MA: The MIT Press.

McDannald, M., E. Kerfoot, M. Gallagher, and P. C. Holland (2005, Feb). Amygdala central nucleus function is necessary for learning but not expression of conditioned visual orienting. *Behav Neurosci 119*(1), 202–212.

McDonald, A. J. (1998, Jun). Cortical pathways to the mammalian amygdala. *Prog Neurobiol 55*(3), 257–332.

McDonald, R. J. and N. M. White (1993, Feb). A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav Neurosci 107*(1), 3–22.

McFarland, N. R. and S. N. Haber (2002, Sep). Thalamic relay nuclei of the basal ganglia form both reciprocal and nonreciprocal cortical connections, linking multiple frontal cortical areas. *J Neurosci 22*(18), 8117–8132.

McIntyre, C. K., T. Hatfield, and J. L. McGaugh (2002, Oct). Amygdala norepinephrine levels after training predict inhibitory avoidance retention performance in rats. *Eur J Neurosci 16*(7), 1223–1226.

McNaughton, B. L. (2003, April). Long-term potentiation, cooperativity and hebb's cell assemblies: a personal history. *Philos Trans R Soc Lond B Biol Sci 358*(1432), 629–634.

Meyer, J.-A. and S. W. Wilson (Eds.) (1991). *From Animals to Animats 1: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, Cambridge, MA. MIT Press.

Miller, E. K. and J. D. Cohen (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci 24*, 167–202.

Mirolli, M. and D. Parisi (2003). Artificial organisms that sleep. In W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, and J. Ziegler (Eds.), *Advances in Artificial Life. Proceedings of the 7th European Conference on Artificial Life*, Berlin, pp. 377–386. Springer Verlag.

Mobini, S., S. Body, M. Y. Ho, C. M. Bradshaw, E. Szabadi, J. F. W. Deakin, and I. M. Anderson (2002, Mar). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology (Berl) 160*(3), 290–298.

Mogenson, G. J., D. L. Jones, and C. Y. Yim (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol 14*(2-3), 69–97.

Montebelli, A., C. Herrera, and T. Ziemke (2007). On cognition as dynamical coupling: An analysis of behavioral attractor dynamics. *Connect Sci*.

Morén, J. and C. Balkenius (2000). A computational model of emotional learning in the amygdala. In J.-A. Meyer, A. Berthoz, D. Floreano, H. L. Roitblat, and S. W. Wilson (Eds.), *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behaviour*, Cambridge, Mass. The MIT Press.

Murphy, R. (2002). Emotion-based control of cooperating heterogeneous robots. *IEEE T Robotic Autom 18*, 744–757.

Nader, K., P. Majidishad, P. Amorapanth, and J. E. LeDoux (2001). Damage to the lateral and central, but not other, amygdaloid nuclei prevents the acquisition of auditory fear conditioning. *Learn Memory 8*(3), 156–163.

Nelson, R. J. and B. C. Trainor (2007, Jul). Neural mechanisms of aggression. *Nat Rev Neurosci 8*(7), 536–546.

Nishiyama, M., K. Hong, K. Mikoshiba, M.-M. Poo, and K. Kato (2000, Nov). Calcium stores regulate the polarity and input specificity of synaptic modification. *Nature 408*(6812), 584–588.

Niv, Y., N. Daw, D. Joel, and P. Dayan (2006, Oct). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*.

Nolfi, S. (2006). Behaviour as a complex adaptive system: on the role of self-organization in the development of individual and collective behaviour. *Complexus 2*(3-4), 195–203.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *J Math Biol 15*(3), 267–273.

O'Keefe, J., N. Burgess, J. G. Donnett, K. J. Jeffery, and E. A. Maguire (1998, August). Place cells, navigational accuracy, and the human hippocampus. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 353*(1373), 1333–1340.

O'Reilly, R. C., M. J. Frank, T. E. Hazy, and B. Watz (2007, Feb). PVLV: the primary value and learned value Pavlovian learning algorithm. *Behav Neurosci 121*(1), 31–49.

Otani, S., H. Daniel, M.-P. Roisin, and F. Crepel (2003, Nov). Dopaminergic modulation of long-term synaptic plasticity in rat prefrontal neurons. *Cereb Cortex 13*(11), 1251–1256.

Pan, W.-X. and B. I. Hyland (2005, May). Pedunculopontine tegmental nucleus controls conditioned responses of midbrain dopamine neurons in behaving rats. *J Neurosci 25*(19), 4725–4732.

Paré, D., G. J. Quirk, and J. E. LeDoux (2004, Jul). New vistas on amygdala networks in conditioned fear. *J Neurophysiol 92*(1), 1–9.

Parisi, D. (2004, December). Internal robotics. *Connect Sci 16*(4), 325–338.

Parkinson, J. A., T. W. Robbins, and B. J. Everitt (2000, Jan). Dissociable roles of the central and basolateral amygdala in appetitive emotional learning. *Eur J Neurosci 12*(1), 405–413.

Parkinson, J. A., P. J. Willoughby, T. W. Robbins, and B. J. Everitt (2000, Feb). Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs Pavlovian approach behavior: further evidence for limbic cortical-ventral striatopallidal systems. *Behav Neurosci 114*(1), 42–63.

Pascucci, T., R. Ventura, E. C. Latagliata, S. Cabib, and S. Puglisi-Allegra (2007, Dec). The medial prefrontal cortex determines the accumbens dopamine response to stress through the opposing influences of norepinephrine and dopamine. *Cereb Cortex 17*(12), 2796–2804.

Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford University Press.

Peciña, S., K. S. Smith, and K. C. Berridge (2006, Dec). Hedonic hot spots in the brain. *Neuroscientist 12*(6), 500–511.

Petrovich, G. D. and M. Gallagher (2007, Jul). Control of food consumption by learned cues: a forebrain-hypothalamic network. *Physiol Behav 91*(4), 397–403.

Petrovich, G. D., P. C. Holland, and M. Gallagher (2005, Sep). Amygdalar and prefrontal pathways to the lateral hypothalamus are activated by a learned cue that stimulates eating. *J Neurosci 25*(36), 8295–8302.

Petrovich, G. D., B. Setlow, P. C. Holland, and M. Gallagher (2002, Oct). Amygdalo-hypothalamic circuit allows learned cues to override satiety and promote eating. *J Neurosci 22*(19), 8748–8753.

Peyron, C., J. M. Petit, C. Rampon, M. Jouvet, and P. H. Luppi (1998, Jan). Forebrain afferents to the rat dorsal raphe nucleus demonstrated by retrograde and anterograde tracing methods. *Neuroscience 82*(2), 443–468.

Pfeifer, R. (1993). Studying emotions: Fungus eaters. In *Proceedings of the first european conference on artificial life*, pp. 916–927.

Phelps, E. A. (2004, Apr). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr Opin Neurobiol 14*(2), 198–202.

Phelps, E. A. and J. E. LeDoux (2005, Oct). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron 48*(2), 175–187.

Phillips, A. G., G. Vacca, and S. Ahn (2008, Aug). A top-down perspective on dopamine, motivation and memory. *Pharmacol Biochem Behav 90*(2), 236–249.

Pickens, C. L., M. P. Saddoris, M. Gallagher, and P. C. Holland (2005, Feb). Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behav Neurosci 119*(1), 317–322.

Pickens, C. L., M. P. Saddoris, B. Setlow, M. Gallagher, P. C. Holland, and G. Schoenbaum (2003, Dec). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J Neurosci 23*(35), 11078–11084.

Pitkänen, A., E. Jolkkonen, and S. Kemppainen (2000). Anatomic heterogeneity of the rat amygdaloid complex. *Folia Morphol 59*(1), 1–23.

Pitkänen, A., V. Savander, and J. E. LeDoux (1997, Nov). Organization of intra-amygdaloid circuitries in the rat: an emerging framework for understanding functions of the amygdala. *Trends Neurosci 20*(11), 517–523.

Pitkänen, A., L. Stefanacci, C. R. Farb, G. G. Go, J. E. LeDoux, and D. G. Amaral (1995, May). Intrinsic connections of the rat amygdaloid complex: projections originating in the lateral nucleus. *J Comp Neurol 356*(2), 288–310.

Porr, B. and F. Wörgötter (2003). Isotropic sequence order learning. *Neural Computation 15*, 831–864.

Prescott, T. J., F. M. Gonzalez, M. Humphries, and K. Gurney (2003). Towards a methodology for embodied computational neuroscience. In *Proceedings of the Symposium on Scientific Methods for the Analysis of Agent-Environment Interaction (AISB2003)*. AISB Press.

Price, J. L. (2003, Apr). Comparative aspects of amygdala connectivity. *Ann N Y Acad Sci 985*, 50–58.

Redgrave, P. and K. Gurney (2006, Dec). The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci 7*(12), 967–975.

Redgrave, P., T. J. Prescott, and K. Gurney (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience 89*(4), 1009–1023.

Reynolds, S. M. and K. C. Berridge (2001, May). Fear and feeding in the nucleus accumbens shell: rostrocaudal segregation of GABA-elicited defensive behavior versus eating behavior. *J Neurosci 21*(9), 3261–3270.

Reynolds, S. M. and K. C. Berridge (2002, Aug). Positive and negative motivation in nucleus accumbens shell: bivalent rostrocaudal gradients for GABA-elicited eating, taste ”liking”/”disliking” reactions, place preference/avoidance, and fear. *J Neurosci 22*(16), 7308–7320.

Richter-Levin, G. (2004, Feb). The amygdala, the hippocampus, and emotional modulation of memory. *Neuroscientist 10*(1), 31–39.

Roesch, M. and G. Schoenbaum (2006). From associations to expectancies: orbitofrontal cortex as gateway between the limbic system and representational memory. In D. H. Zald and S. L. Rauch (Eds.), *The orbitofrontal cortex*, pp. 199–236.

Rolls, E. T. (1999). *The brain and emotion*. Oxford: Oxford University Press.

Rolls, E. T. (2000, Apr). Précis of the brain and emotion. *Behav Brain Sci 23*(2), 177–91; discussion 192–233.

Rolls, E. T. (2005, Jan). Taste and related systems in primates including humans. *Chem Senses 30 Suppl 1*, i76–i77.

Rolls, E. T. and R. P. Kesner (2006, May). A computational theory of hippocampal function, and empirical tests of the theory. *Prog Neurobiol 79*(1), 1–48.

Romanelli, P., V. Esposito, D. W. Schaal, and G. Heit (2005, Feb). Somatotopy in the basal ganglia: experimental and clinical evidence for segregated sensorimotor channels. *Brain Res Brain Res Rev 48*(1), 112–128.

Romanski, L. M., M. C. Clugnet, F. Bordi, and J. E. LeDoux (1993, Jun). Somatosensory and auditory convergence in the lateral nucleus of the amygdala. *Behav Neurosci 107*(3), 444–450.

Rosen, J. B. (2004, Mar). The neurobiology of conditioned and unconditioned fear: a neuro behavioral system analysis of the amygdala. *Behav Cogn Neurosci Rev 3*(1), 23–41.

Saddoris, M. P., M. Gallagher, and G. Schoenbaum (2005, Apr). Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. *Neuron 46*(2), 321–331.

Sah, P., E. S. L. Faber, L. D. M. Armentia, and J. Power (2003, Jul). The amygdaloid complex: anatomy and physiology. *Physiol Rev 83*(3), 803–834.

Sander, D., D. Grandjean, and K. R. Scherer (2005, May). A systems approach to appraisal mechanisms in emotion. *Neural Networks 18*(4), 317–352.

Schachter, S. and J. Singer (1962, Sep). Cognitive, social, and physiological determinants of emotional state. *Psychol Rev 69*, 379–399.

Schembri, M., M. Mirolli, and G. Baldassarre (2007). Evolution and learning in an intrinsically motivated reinforcement learning robot. In *Advances in Artificial Life*, Volume 4648, pp. 294–303. Springer Berlin / Heidelberg.

Scherer, K. R., A. Schorr, and T. Johnstone (2001). *Appraisal processes in emotion: theory, methods, research*. Oxford University Press.

Schoenbaum, G., M. P. Saddoris, and T. A. Stalnaker (2007, Dec). Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann N Y Acad Sci 1121*, 320–335.

Schoenbaum, G., B. Setlow, S. L. Nugent, M. P. Saddoris, and M. Gallagher (2003). Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learn Memory 10*(2), 129–140.

Schoenbaum, G., B. Setlow, M. P. Saddoris, and M. Gallagher (2003, Aug). Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron 39*(5), 855–867.

Schultz, W. (2002, Oct). Getting formal with dopamine and reward. *Neuron 36*(2), 241–263.

Schultz, W., P. Dayan, and P. R. Montague (1997, Mar). A neural substrate of prediction and reward. *Science 275*(5306), 1593–1599.

Schultz, W. and A. Dickinson (2000). Neuronal coding of prediction errors. *Annu Rev Neurosci 23*, 473–500.

Sejnowski, T. J., C. Koch, and P. S. Churchland (1988, Sep). Computational neuroscience. *Science 241*(4871), 1299–1306.

Sejnowski, T. J. and G. Tesauro (1989). The hebb rule for synaptic plasticity: algorithms and implementations. In J. H. Byrne and W. O. Berry (Eds.), *Neural Models of Plasticity*, Chapter 6, pp. 94–103. Academic Press.

Semba, K. and H. C. Fibiger (1992, Sep). Afferent connections of the laterodorsal and the pedunculopontine tegmental nuclei in the rat: a retro- and antero-grade transport and immunohistochemical study. *J Comp Neurol 323*(3), 387–410.

Setlow, B., M. Gallagher, and P. C. Holland (2002, Jun). The basolateral complex of the amygdala is necessary for acquisition but not expression of CS motivational value in appetitive Pavlovian second-order conditioning. *Eur J Neurosci 15*(11), 1841–1853.

Setlow, B., P. C. Holland, and M. Gallagher (2002, Apr). Disconnection of the basolateral amygdala complex and nucleus accumbens impairs appetitive pavlovian second-order conditioned responses. *Behav Neurosci 116*(2), 267–275.

Shi, C. and M. Davis (1999, Jan). Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *J Neurosci 19*(1), 420–430.

Simon, H. A. (1996, October). *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Skinner, B. F. (1938). *The Behavior of Organisms*. New York, NY: Appleton Century Crofs.

Smith, K. S. and K. C. Berridge (2005, Sep). The ventral pallidum and hedonic reward: neurochemical maps of sucrose "liking" and food intake. *J Neurosci 25*(38), 8637–8649.

Sørensen, E., B. Bjorvatn, and R. Ursin (2000, Sep). Sleep-wake effects following the selective 5-ht(1a) receptor antagonist p-mppi in the freely moving rat. *Behav Brain Res 114*(1-2), 31–38.

Spratling, M. W. and M. H. Johnson (2002). Pre-integration lateral inhibition enhances unsupervised learning. *Neural Comput 14*(0), 2157–2179.

Surmeier, J. D., J. Ding, M. Day, Z. Wang, and W. Shen (2007, May). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci 30*(5), 228–235.

Sutton, R. S. and A. Barto (1998). Reinforcement learning i: Introduction. Course notes.

Sutton, R. S. and A. G. Barto (1981, March). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev 88*(2), 135–170.

Sutton, R. S. and A. G. Barto (1990). Time-derivative models of pavlovian reinforcement. In M. Gabriel and J. Moore (Eds.), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, pp. 497–537. The MIT Press.

Sutton, R. S., A. G. Barto, and C. J. C. H. Watkins (1989). Learning and sequential decision making. In M. Gabriel and J. W. Moore (Eds.), *Learning and Computational Neuroscience*, pp. 539–602. The MIT Press.

Thompson, R. F., R. Swain, R. Clark, and P. Shinkman (2000, Jun). Intracerebellar conditioning–Brogden and Gantt revisited. *Behav Brain Res 110*(1-2), 3–11.

Thorndike, E. L. (1911). *Animal Intelligence*. Rutgers, NJ: Transaction Publishers.

Tindell, A. J., K. S. Smith, S. Peciña, K. C. Berridge, and W. J. Aldridge (2006, Nov). Ventral pallidum firing codes hedonic reward: when a bad taste turns good. *J Neurophysiol 96*(5), 2399–2409.

Tolman, E. C. (1932). *Purposive Behavior in Animals and Men*. Century, New York.

Trullier, O., S. I. Wiener, A. Berthoz, and J.-A. Meyer (1997, April). Biologically based artificial navigation systems: review and prospects. *Prog Neurobiol 51*(5), 483–544.

Usuda, I., K. Tanaka, and T. Chiba (1998, Jun). Efferent projections of the nucleus accumbens in the rat with special reference to subdivision of the nucleus: biotinylated dextran amine study. *Brain Res 797*(1), 73–93.

Venditti, A., M. Mirolli, D. Parisi, and G. Baldassarre (2009). A neural-network model of the dynamics of hunger, learning, and action vigor in mice. In R. Serra (Ed.), *Proceeding of the Italian Workshop on Artificial Life and Evolutionary Computatoin (WIVACE08)*. World Scientific. In press.

Verhagen, J. V., M. Kadohisa, and E. T. Rolls (2004, Sep). Primate insular/opercular taste cortex: neuronal representations of the viscosity, fat texture, grittiness, temperature, and taste of foods. *J Neurophysiol 92*(3), 1685–1699.

Voorn, P., B. Jorritsma-Byham, V. C. Dijk, and R. M. Buijs (1986, Sep). The dopaminergic innervation of the ventral striatum in the rat: a light- and electron-microscopical study with antibodies against dopamine. *J Comp Neurol 251*(1), 84–99.

Weidenfeld, J., M. E. Newman, A. Itzik, E. Gur, and S. Feldman (2002, Aug). The amygdala regulates the pituitary-adrenocortical response and release of hypothalamic serotonin following electrical stimulation of the dorsal raphe nucleus in the rat. *Neuroendocrinology 76*(2), 63–69.

Weng, J., J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen (2001). Autonomous mental development by robots and animals. *Science 291*, 599–600.

Winstanley, C. A., D. E. H. Theobald, R. N. Cardinal, and T. W. Robbins (2004, May). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J Neurosci 24*(20), 4718–4722.

Woodin, M. A., K. Ganguly, and M.-M. Poo (2003, Aug). Coincident pre- and postsynaptic activity modifies gabaergic synapses by postsynaptic changes in cl- transporter activity. *Neuron 39*(5), 807–820.

Wörgötter, F. and B. Porr (2005, February). Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Computation 17*(2), 245–319.

Wyvell, C. L. and K. C. Berridge (2000, Nov). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. *J Neurosci 20*(21), 8122–8130.

Yeterian, E. H. and D. N. Pandya (1995, Feb). Corticostriatal connections of extrastriate visual areas in rhesus monkeys. *J Comp Neurol 352*(3), 436–457.

Yin, H. H. and B. J. Knowlton (2006). The role of the basal ganglia in habit formation. *Nat Rev Neurosci 7*, 464–476.

Yin, H. H., B. J. Knowlton, and B. W. Balleine (2004, Jan). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci 19*(1), 181–189.

Zahm, D. S. (2000, Jan). An integrative neuroanatomical perspective on some subcortical substrates of adaptive responding with emphasis on the nucleus accumbens. *Neurosci Biobehav Rev 24*(1), 85–105.

Zlatev, J. and C. Balkenius (2001). Introduction: Why epigenetic robotics? In C. Balkenius, J. Zlatev, H. Kozima, K. Dautenhahn and, and C. Breazeal (Eds.), *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, Volume 85, pp. 1–4.