Künstl Intell (2011) 25:235–244
DOI 10.1007/s13218-011-0106-y

FACHBEITRAG

A Neuroscientific View on the Role of Emotions in Behaving Cognitive Agents

Julien Vitay · Fred H. Hamker

Received: 1 February 2011 / Accepted: 2 May 2011 / Published online: 15 May 2011

Abstract While classical theories systematically opposed emotion and cognition, suggesting that emotions perturbed the normal functioning of the rational thought, recent progress in neuroscience highlights on the contrary that emotional processes are at the core of cognitive processes, directing attention to emotionally-relevant stimuli, favoring the memorization of external events, valuating the association between an action and its consequences, biasing decision making by allowing to compare the motivational value of different goals and, more generally, guiding behavior towards fulfilling the needs of the organism. This article first proposes an overview of the brain areas involved in the emotional modulation of behavior and suggests a functional architecture allowing to perform efficient decision making. It then reviews a series of biologically-inspired computational models of emotion dealing with behavioral tasks like classical conditioning and decision making, which highlight the computational mechanisms involved in emotional behavior. It underlines the importance of embodied cognition in artificial intelligence, as emotional processing is at the core of the cognitive computations deciding which behavior is more appropriate for the agent.

Keywords Emotion · Cognition · Behavior · Computational neuroscience · Cognitive agents

J. Vitay · F.H. Hamker (✉)
Fakultät für Informatik, Technische Universität Chemnitz, Straße der Nationen 62, 09107 Chemnitz, Deutschland
e-mail: fred.hamker@informatik.tu-chemnitz.de

J. Vitay
e-mail: julien.vitay@informatik.tu-chemnitz.de

1 Introduction

Although the successes of artificial intelligence (AI) have long relied on abstract forms of cognition, like problem-solving, planning or sensorimotor learning, the interest for a computational study of emotions has recently grown up [2]. A large body of emotional research in AI is performed under the label "Affective Computing" [6, 29]. It is strongly influenced by the psychology of emotion research and deals with the recognition and subsequent processing of emotional responses in humans by artificial agents. The modeling of emotion takes place at a rather descriptive and functional level independent of particular brain areas involved. We here focus on another direction of emotional research in AI which concerns the incorporation and simulation of emotional responses in the artificial agents themselves by relying on neuroscientific observations. Neuroscience research in the last 15 years has produced a huge wealth of data which can guide research in AI towards brain-like approaches to intelligence. Contrary to the classical view which states that emotions are competing with cognitive processes to express behavior [21], it is clearly suggested that emotions are integrant parts of cognition as they contribute to valuate external events, evaluate competing goals and infer mental states [34]. Agents or cognitive robots whose cognitive abilities incorporate an emotional dimension could therefore interact more efficiently with humans and learn useful tasks in a complex and weakly predictable environment.

Emotions have acquired through evolution a crucial role in regulating an organism's behavior. The emotional response to a stimulus can be used by the brain to guide sensory processing towards the relevant features and focuses attention on the emotional object: visual search is for example faster and discrimination more successful when the objects have an emotional value [26]. Emotions have an im-

portant role in facilitating the formation of long-term memory: emotionally-charged events are better remembered [8]. Emotion is also a motivator for behavior, not only because motor actions are more vigorous towards affective objects, but also because these objects are more likely to become the goal of an action. In goal-directed learning, the consequence of an action, the outcome, has an emotional value which is associated through learning to the action itself and becomes itself the goal of the action [3]. The outcome gains therefore an incentive value that elicits action in order to experience a given emotion, forming the basis of motivation. This incentive value can moreover be used to decide between two possible options: the one associated with the greatest incentive value will be chosen to guide behavior. Emotion acts therefore as a common currency that helps to evaluate and decide between alternatives, even when they lead to physically different outcomes. Interestingly, it has been shown that money activates the same brain regions as primary rewards like food or sex, and the relative incentive value of different outcomes is computed and compared by the same population of neurons, the ventromedial prefrontal cortex (vmPFC) [9].

This predictive nature of emotions can also be used in planning processes and decision making. The somatic marker hypothesis of Damasio [5] proposes that the emotional state of the organism (including visceral signals) is represented in vmPFC and can be used as an emotional emulator by cognitive processes in order to predict the emotional consequences of a planned action and chose the better option. This emotional prediction can further provoke the bodily responses corresponding to the prediction, although this can be controlled and reduced with practice. In other words, executive structures elaborate different plans in a given context and the emotional system tells if they will lead to a satisfying situation or not from the organism's viewpoint. The same system is activated when observing other people's emotions, suggesting an emulation of others' emotions in order to infer their mental state, forming the basis of empathy through an emotional mirror system [4].

Despite the accumulation of knowledge about the role of emotion in cognition and behavior, it is still largely neglected by AI models. One famous exception is the field of reinforcement learning (RL), which use the delivery of a primary reward (comparable to an object procuring a positive emotion) to influence behavior and learning through expectation [39]. Although emotions are hugely simplified and exist only in terms of positive or negative value, RL has the merit to introduce emotional processing and its expectation into cognitive functions. The technical success of this approach and its relative biological plausibility (see Sect. 3.4) paves the way for a deeper understanding of the brain mechanisms linking emotions and behavior, and for the design of more flexible cognitive, emotional and autonomous agents [44].

The first part of this review presents the main brain areas that are thought to be involved in emotional processes and influence behavior, by describing their functional roles and relationships. We propose a functional network of these areas that focuses on decision making, although emotions have a much broader role in brain processes, e.g. in memory, attention or perception. The second part presents a set of biologically plausible computational models of emotion in various behavioral tasks such as classical conditioning[1] (either aversive or appetitive), reinforcement learning, reversal learning[2] and decision making. The interested reader would benefit from reading the excellent review by Levine on "Neural network modeling of emotion" which spans a larger array of emotional computational models [23].

2 Biological Groundings of Emotion Processing

The nervous system is anatomically composed of several interconnected assemblies of neurons (called structures, nuclei or areas) that are specialized for specific aspects but which interact together by electrically propagating information (in the form of action potentials), or by releasing neuromodulators that modify neural properties (input/output function) on a large scale. These structures can be grouped into four major categories: the brainstem, controlling basic homeostatic functions such as breath or cardiac rhythm; the cerebellum, involved in motor learning; subcortical structures such as the hippocampus (episodic memory, place coding), the thalamus (sensory relay, multimodal integration), the hypothalamus (sleep/wake cycles, consumatory behaviors, hormonal regulation), the amygdala (fear processing, emotional attention) or the basal ganglia (action selection, decision making, reinforcement learning); and the cerebral cortex, the most phylogenetically advanced structure (characteristic of mammals), which is the outermost surface of the brain. The cerebral cortex has many functional roles depending of the considered area, ranging from perception in its posterior part to movement generation, planning, anticipation or socialization in its anterior part (the prefrontal cortex).

Although each structure can be described as involved in some particular functions, they do not compute them as iso-

[1] *Classical conditioning* (or Pavlovian conditioning) is the learned pairing of an unconditioned stimulus (US, e.g. food) that produces an unconditioned response (UR, e.g. salivation) with a conditioned stimulus (CS, e.g. a bell) presented a certain amount of time before the US. After sufficient learning, the appearance of the CS produces a conditioned response (CR) that is similar to the UR.

[2] *Reversal learning* is a form of operant conditioning where the consequences of two well-learned actions are reversed: if the action A systematically led to reward and B to punishment, the experimenters look how fast the animal adapts its behavior when A suddenly leads to punishment and B to reward.

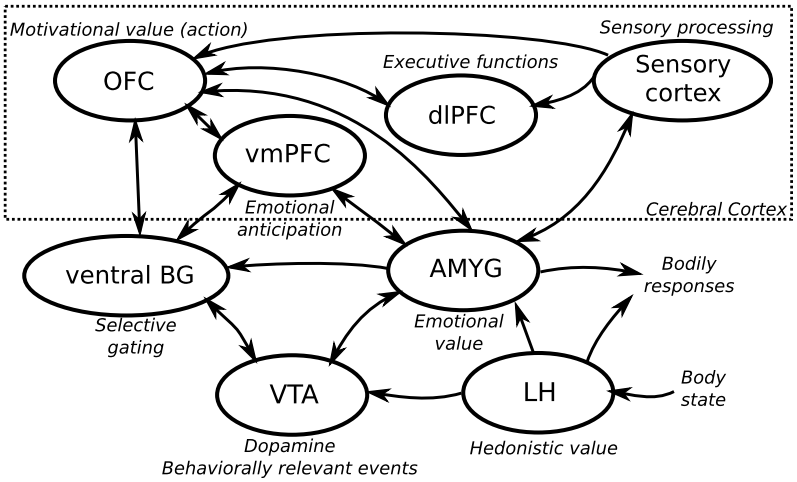

Fig. 1 Simplified connection diagram of the main areas involved in emotion and behavior. *OFC*: orbitofrontal cortex; *vmPFC*: ventromedial prefrontal cortex; *dlPFC*: dorsolateral prefrontal cortex; *ventral BG*: ventral basal ganglia; *VTA*: ventral tegmental area; *AMYG*: amygdala; *LH*: lateral hypothalamus. The term "sensory cortex" gathers here various sensory areas of the cortex, in all modalities (vision, audition, smell, taste, touch) and at various levels of computation (from basic feature extraction to object recognition)

lated independent modules, but rather within an interconnected, looped and dynamic network of structures that cooperate to perform the considered function, one structure being able to participate to several networks depending on the context. The several cortical areas are connected with each other along a complex hierarchy, but also receive subcortical information (relayed and integrated in the thalamus) from different structures. In the case of emotions, several structures (called the limbic system) have been identified as participating mainly in emotional processes, but further researches have extended the emotional brain to various areas that were previously thought as purely motor or cognitive. It is know admitted that emotions influence a majority of brain structures, including perceptive ones, at various levels. Figure 1 represents a simplified schematic diagram of the major emotional brain areas that play a role in the generation of behavior. For clarity, it omits many structures and known connections that play nevertheless an important role in emotions. This section will present the role of the most relevant areas and present a tentative overview of the neural network that underly the influence of emotion on behavior and cognition.

2.1 Lateral Hypothalamus

The lateral hypothalamus (LH) is the place where the needs of the organism are evaluated. It maintains homeostasis by regulating blood pressure, heart rate, and temperature and continuously scrutinize different bodily parameters such as the glucose level in blood for hunger or osmolite concentration for thirst. It releases a variety of hormones that urge the organism (other brain areas or functional organs) on reestablishing its equilibrium by expressing the appropriate behavior. Its cells particularly respond to food consumption and its predictors: they respond similarly to the deprivation of a metabolite (e.g. a low glucose level), its taste (the ingestion of sugar) or the associated conditioned stimulus (a bell that predicts the delivery of sugar) [27]. Some cells respond selectively for appetitive or aversive gustatory inputs (salt, fat, sugar, umami, bitter...), when others shows an opposite pattern of activation: they are excited by appetitive stimuli and inhibited by aversive, and *vice versa*. Interestingly, these opposite cells show a rebound activation when their non-preferred stimulus stops: cells that respond for appetitive (resp. aversive) stimuli can become transiently activated when a painful (resp. rewarding) stimulation stops. This may form the basis of relief and frustration and can be modeled by gated dipoles (see Sect. 3.2). Another interesting property of LH cells is that gustatory responses are modulated by the satiety of the organism, which means that a sugared aliment stops being appetitive when the glucose blood level is high. In short, LH represents the hedonistic value of objects relative to the state of the organism and is able to signal this to other brain areas which in turn can select the kind of behavior that would be beneficial for the organism.

2.2 Amygdala

The amygdala is the most studied area in emotional neuroscience. It was long associated to fear conditioning (a form of classical conditioning where the US is painful, see Sect. 3.3) and fear processing because of its increased activation when viewing fearful objects or faces, but more recent studies have shown that it is also responsive for appetitive stimuli, although the appetitive and aversive populations of neurons are distinct (but not spatially segregated) [24]. Its output nuclei (CE) has direct projections to LH and other brainstem nuclei and is necessary to exhibit conditioned response such as freezing, increased blood pressure or heartbeat rate and can provoke the release of various neuromodulators such as serotonin, adrenalin or dopamine. Its input nuclei (BLA) receives connections from virtually the whole cerebral cortex (especially high-level sensory areas and prefrontal cortex) as well as from the thalamus and LH.

It receives information about both conditioned (CS) and unconditioned (US) stimuli and can learn the association between the two through long-term potentiation (LTP) of its synapses. It is also a major place where the sensory features of a stimulus are combined with its hedonistic value (e.g. taste in LH), leading to a representation of the emotional value of a stimulus in the context of the organisms's needs and its cognitive expectations.

On top of its connections to CE that can generate bodily responses to the conditioned stimuli, BLA projects back to the cerebral cortex. Its connections to the visual cortex allows a sharper and faster processing of emotionally-relevant stimuli and creates an emotional focus of attention [43]. Its projections to the hippocampus also enhance long-term memory formation by dynamically modulating the learning of emotionally-charged events through the process of *emotional tagging* [32]. Furthermore, it heavily projects to the prefrontal cortex, especially to the orbitofrontal cortex (OFC) and the ventromedial prefrontal cortex (vmPFC). The role of these connections is to update (rather than to store) the motivational value of expected outcomes. Once this value has been learned, the amygdala is no longer necessary, but animals are not able to update it when the contingency changes without the amygdala [24]. The amygdala has therefore an important role in computing the emotional valence of a stimulus in relation to its sensory features, and in using it to modulate motivation and bias action selection.

2.3 Ventral Basal Ganglia and Dopamine

The basal ganglia (BG) is considered as a central structure for appetitive behavior and reward processing. The BG is a set of nuclei in the basal forebrain which creates a functional loop with the cerebral cortex through its massive cortical afferents and its projections to the thalamus. It can be functionally divided into three main domains depending on its cortical afferents: associative (dorsolateral prefrontal cortex dlPFC), motor (premotor, motor and somatosensory cortices) and limbic (OFC and vmPFC). These three domains are largely segregated, but a transfer of information from limbic to motor exists [18]. We will focus here on the limbic BG (or ventral BG), but the interested reader can refer to our review of cognitive and motor models of the BG [41].

The ventral BG (with the nucleus accumbens—NAcc-as an input nucleus) is critical for learning flexible behaviors [19]. It receives connections from the amygdala (emotional value), the hippocampus (contextual information) and prefrontal areas like OFC and vmPFC (motivational values, emotional state, goals). Its architecture and its central position in the emotional system confer it a role of selection of behavioral strategies in accordance with the emotional needs and expectations of the organism. One particularity of NAcc is its strong innervation by dopaminergic cells in the ventral tegmental area (VTA) whose firing have been shown to reflect the delivery of primary rewards, but also the appearance of the corresponding CS after conditioning [36]. Dopamine (DA) is considered as the neuromediator of pleasure, being released in rewarding situations, but also as a reward prediction error teaching signal thanks its ability to modulate synaptic learning in prefrontal cortex and BG (see Sect. 3.4).

NAcc is crucial to learn the association between an action and its outcome which forms the basis of instrumental conditioning and goal-directed learning. It transfers the emotional value of the outcome (learned by classical conditioning processes in the amygdala) into the valuation of the preceding action: NAcc is critical for these Pavlovian-to-Instrumental transfer (PIT) processes. It receives information about the potential goals (or reachable outcomes) that are represented in OFC and updates their motivational value thanks to the information in the amygdala. Due to its funneling architecture, the ventral BG selects only one possible goal with the higher motivational value and feeds it back to OFC, influencing the final decision towards that particular goal.

2.4 Orbitofrontal Cortex

The orbitofrontal cortex (OFC) is an important area for representing the motivational value of expected outcomes or goals. It is central in decision making and goal-directed learning. Interestingly, there is an anatomical segregation in OFC depending on the value of the reward association: the medial OFC monitors the association between a stimulus and a rewarded response, while the lateral OFC deals with punishing outcomes or outdated rewarded associations (reversal learning). There is also evidence that OFC cells show sustained activation to maintain stimulus-reward associations in working memory [33]. OFC-lesioned patients are unable to update their action-outcome associations once it has been learned (reversal and extinction learning[3]). They show abnormal perseverance in bad solutions and are unable to show the cognitive flexibility necessary to solve complex non-verbal cognitive tasks like the Iowa gambling task [5]. It has been shown that the connection from the amygdala to OFC is crucial for this updating function and that both VTA and OFC are necessary in order to learn to obtain new rewarding outcomes [40]. This emphasizes the role of the functional network composed of OFC, the ventral BG, VTA and the amygdala in flexibly evaluating the desirableness of an outcome and in selecting the most appropriate behavior. The OFC is in strong interaction with the dorsolateral prefrontal cortex (dlPFC), which is involved in executive functions like visual working memory and planning and is able

[3] *Extinction learning* is a form of classical conditioning where a learned CS is no longer associated to an US (or an action to it outcome).

to learn and execute complex plans in order to achieve the goal that is selected in the OFC-BG-amygdala network.

2.5 Ventromedial Prefrontal Cortex

The ventromedial prefrontal cortex (vmPFC) is a key structure in the cognitive control and regulation of emotion [25]. Its inhibitory projections to the amygdala have been shown to be necessary for the extinction of fear learning: after an aversive CS-US association has been learned, the successive presentation of the CS alone should lead to a decrease in the fear-related bodily responses generated by the amygdala. If vmPFC is lesioned, these responses do not vanish, showing a role of vmPFC in the anticipation of emotional events and in the updating of emotional responses by the cognitive context. Interestingly, hypoactivity in vmPFC is associated to post-traumatic stress disorder (PTSD—the uncontrollable recall of traumatic events), chronic stress and phobias [20]. The vmPFC is also considered the place of the self (individual preferences, self-reflection). It represents the emotional state with respect to the individual's past experience with a common currency: physiologically different emotional events (like food and money deliveries) can be compared on an unified basis [9]. It is also involved in the social communication of emotion and to the internal representation of others' emotions (empathy). Lesions in vmPFC lead both to impaired social skills and to the inability to decide for advantageous alternatives.

2.6 Summary

Although we take the risk to oversimplify the functional architecture underlying the role of emotions in cognition and behavior, we can stress a few sentences to summarize these processes, as represented in Fig. 1. The hedonistic value of an outcome is evaluated in LH (and other nuclei) by comparing the physical properties of the outcome with the biological needs of the organism and releasing various hormones and neuromodulators (including DA) that can influence behavior. It is associated in the amygdala with the sensory features of the outcome and eventually to the conditioned stimuli that predicted it, provoking anticipatory bodily emotional responses. This associative learning of the emotional value of an outcome in the amygdala is rather automatic in classical conditioning, but instrumental learning (goal-directed) requires the production of the adequate action in order to obtain that outcome. This action-outcome association (motivational value of the outcome) is represented in OFC and updated by the amygdala. The functional loop between OFC and the ventral BG, modulated by the amygdala, allows to select the action that leads to the maximum amount of reward and instruct dlPFC to perform the sequence of movements or cognitive operations necessary to obtain the corresponding outcome. Additionally, the functional loop between vmPFC and the amygdala allows to express the emotional value of the different possible outcomes into a common framework that depends on the individual's preferences and social interactions. We now review a few computational models of emotional processing that have minimal biological plausibility by using artificial neural networks. We first focus on models of conditioning, both appetitive and aversive, as we saw that conditioning is a key process in emotional processing, allowing to understand many involved mechanisms. We then focus on more elaborated models of decision making, including reinforcement learning approaches.

3 Computational Models of Classical Conditioning

3.1 Associative Models of Conditioning

Early theories of classical conditioning considered that the animal had learned directly the stimulus-response (S-R) association between the CS and the UR, as if the animal thought the CS was indeed the US. Later theories have rather suggested that the animal only learns that the CS predicts the US (stimulus-stimulus association S-S) and that this learned association strength generates the CR. Rescorla and Wagner proposed a functional model of S-S association, in which the associative strength V_t is updated through learning proportionally to the difference between the former association strength V_{t-i} and the conditioning strength of the US Λ (let us say the amplitude of the UR, or its emotional value) [31]. This simple error-correction model was already able to explain several experimental results, such as blocking (a stimulus presented simultaneously to an already conditioned stimulus does not gain association strength), extinction (the CS is no longer associated to the US) or generalization (a stimulus sharing enough sensory features with a CS can elicit the same CR). Although it does not deal with temporal representations of stimuli and is not specific regarding the CR, this model influenced a bunch of subsequent models, covering the remaining observation on classical conditioning but lacking biological plausibility [35, 38].

Grossberg proposed in [14] an alternative functional model that has been later transposed into a biologically-motivated neural network [16]. In these models, both CS and US are represented in a sensory layer of neurons that compete with each other through shunting inhibition so that only one neuron is active at a time (attentional effect). Their projections to an associational area (corresponding to OFC) and to a "drive" neuron in the amygdala (which represents the emotional value associated to the US as well as the need of the organism for that reinforcer) are learned so that a CS can at the same time activate the emotional representation

of its US and become preferentially the target of an action. On top of its higher biological plausibility, this model has the advantage to explain classical conditioning not in terms of S-R or S-S associations, but rather by the emotional value of the US that is transformed into a motivational value for responding to the US. In other terms, Pavlov's dog does not salivate because he thinks the bell is eatable, but because the bell prepares it to the action of consuming something pleasant.

3.2 Gated Dipoles

As the previous model worked with either appetitive or aversive conditioning, but not both at the same time, Grossberg proposed a model of affective opponent processing called *gated dipoles* [15]. If one consider the situation where an animal has to perform a certain action in order to stop a painful stimulation, the question that arises is: how can an absence of external stimulation become the goal of an action? He proposed a neural network where each emotional response is computed by two opponent ON and OFF channels. These channels compete with each other and adapt to their inputs by the depletion of the corresponding neurotransmitter. They also both receive an arousal input. When the appetitive (resp. aversive) stimulation is present, the ON channel inhibits the OFF channel but progressively loses its efficiency. When the stimulation stops, the OFF channel receives more arousal than the ON signal because of the neurotransmitter depletion and wins the competition against the ON channel, what produces a phasic antagonistic rebound of the opposite emotional response without stimulation. This explains the sensation of relief after the cessation of a painful stimulation, or the frustration when a pleasant stimulation suddenly stops. This mechanism is in accordance with the general opponent-process theory of motivation but, more interestingly, also provides an account for the observed pattern of rebound activation in LH. This model was combined with the previous to explain both primary and secondary, appetitive and aversive, conditioning [16]. The same author later added time delays to cope with different intervals between the CS and the US, and used this Cog-Em (cognitive-emotional) network to explain a variety of tasks, including the observed emotional and cognitive deficits in autism [17].

3.3 Fear Conditioning in the Amygdala

The most neurobiologically studied form of emotional conditioning in the brain is fear conditioning in the amygdala, mostly thanks to the work by LeDoux [22]. He identified the anatomical pathways supporting a particular form of conditioning where an auditory stimulus (CS) is followed by an electric shock (US). As shown in Fig. 2, he observed that

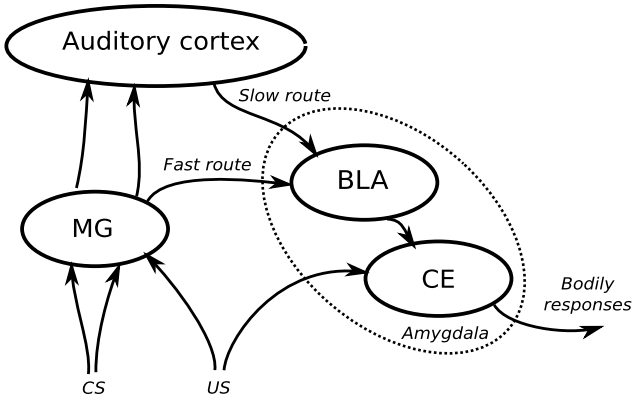

Fig. 2 Schematic diagram of fear conditioning in the amygdala. *MG*: medial geniculate body of the thalamus; *BLA*: basolateral amygdala; *CE*: central nucleus of the amygdala

information reached the amygdala through different pathways: the nociceptive information about the US reaches directly the output nucleus of the amygdala CE to generate bodily responses, but also the medial geniculate body of the thalamus (MG) where it is further relayed to the basolateral amygdala (BLA). Auditory information about the CS also reaches MG, forming a "fast" route of CS information to the amygdala. However, auditory information projects further to the auditory cortex, where it also reaches BLA through the "slow" route. The auditory cortex has finer discrimination properties than the auditory thalamus, allowing a better evaluation of the exact nature of the signal. The BLA is then able to associate information about the CS and the US, and its projections to CE can generate the CR. This has been simulated in a computational model which replicated several known properties of fear conditioning as well as neural firings, and made several predictions [1]. A lot of issues remain nevertheless open. Information about the CS and the US are already integrated in MG, which raises the question of what is really learned in BLA. In fear conditioning, the US has only a nociceptive dimension, but no visual one like in the food delivery in Pavlov's experiment: what are the pathways in that case? It would be also interesting to know if the appetitive conditioning in the amygdala follow the same mechanism and how it interacts with aversive conditioning. Further experimental and modeling work is necessary to better understand the conditioning processes in the amygdala.

3.4 Reinforcement Learning with Dopamine

The Rescorla-Wagner rule of conditioning was further improved by Sutton and Barto [38] by adding memory traces for the representation of stimuli that explained that the optimal time interval between the CS and the US is around one second in several experiments. This lead to the development of the temporal difference algorithm (TD) that relied on a temporal derivative of the discounted expected reward

and paved the way for the field of reinforcement learning [39]. The reward prediction error signal of the TD algorithm shows similarities with the firing of dopaminergic cells in VTA in appetitive conditioning [36], what lead to the development of efficient models of DA and BG functioning [37]. In these models, the ventral BG together with VTA is thought to play the role of the critic in actor/critic architectures by evaluating the discrepancy between the expected and the actual reward, the associative and motor parts of the BG playing the role of the actor.

We reviewed in [41] some of the criticisms made to this analogy, the most important of them being that the reward prediction error should progressively shift back in time through learning from the US to the appearance of the CS, when a simultaneous increase of CS-related and decrease of US-related DA activation in VTA is observed. Moreover, it is only able to learn one fixed time interval between the CS and the US. Other models have since been developed to cope with these problems [7, 28, 42] by separating the pathways that generate DA firing for the CS and the US. In particular, the model proposed in [28] can successfully modulate the learning of working memory tasks. Nevertheless, as long as no biological plausibility is required, TD is still a widely used model to explain DA firing because of its computational simplicity.

4 Computational Models of Decision-Making

There is a growing number of computational models of decision-making at various levels of abstraction. We focus here on models that claim for a minimal biological plausibility. We separate them into two categories depending on the emphasis they put on cognitive or emotional processes, although the frontier is blurry.

4.1 Cognitive TD-Based Models

The analogy between the TD algorithm and DA firing, associated with the fact that DA strongly modulates processing in prefrontal areas that are involved in decision-making, led to the development of cognitive models of decision-making that rely on RL techniques. In [10], Daw et al. make a distinction between model-free learning (by associating to each state or action a value corresponding the expected sum of reward, like in the TD algorithm) and model-based learning (by computing on the fly this value through the descent of a Markovian tree that reflects the expected states encountered by the system). They assign the former to the basal ganglia and dopaminergic system, while the latter should be constructed in the prefrontal cortex. The two systems compete for the final decision by signaling their uncertainty: the model-free system estimates itself the correctness of its reward prediction, while the model-based one sums the uncertainty of the decision tree's states. Although this model provides an interesting account of behavioral data, it is not subject to learning (the tree is hard-coded) and to exploration. Moreover, it predicts that the prefrontal cortex learns faster than BG the changes in action-outcome contingencies, while the contrary has been experimentally observed.

In [30], Rao elaborates further this Bayesian inference principle by relying on partially observable Markov decision processes (POMDP). Instead of evaluating individually all states of the environment (which are never fully accessible), they compute in the cerebral cortex a posterior probability distribution called the belief state. Similarly to the actor-critic models of BG, they assume that the motor domain of the BG selects the most appropriate action in this belief context while the ventral BG should "critic" the current belief. Interestingly, the actor can also decide to sample more deeply its environment when the uncertainty is too high before taking a decision and risking negative reinforcement. This behavior decreases with successful learning. The model is able to learn a random dots discrimination task[4] and provide an interesting account of the cooperation between cortical and subcortical structures, as well as of the role of the dopaminergic signal in uncertain situations. Although the mapping of the different algorithms onto brain structures is fairly coherent, it remains to map this model onto realistic neural networks.

4.2 Emotional Models

Deco and Rolls have provided a detailed spiking neurons model of reversal learning in OFC [11]. They noticed that such action-outcome reversal occurs sometimes within a single trial, which is too fast to be accounted for by long-term synaptic changes. They designed an attractor-network model of OFC composed of different pools of neurons (or modules): sensory, associative, rewarding and rule-specific pools. The different stimulus-reward associations are hard-coded in the associative pool (combinatory representation) and biased by the rule pool: for example, rule 1 states that object A is rewarding and not B, while rule 2 states the contrary. Each time an expected reward is omitted (which occurs at each reversal), the currently active associative representation is inhibited, the rule is updated and a new stimulus-reward association is selected. Unsurprisingly, such an architecture leads to combinatory explosion when the number

[4] *Random dots discrimination* task consists in the visual presentation of an array of randomly moving dots, biased on average towards a particular direction (e.g. left or right). The subject has to guess this direction of movement. Reaction times typically increase with the difficulty (or uncertainty) of the task.

of stimulus and rewards increases, but this model gives already insights on the mechanisms involved in cognitive flexibility. As the network does not learn connections, the model does not explain how the stimulus-reward associations are formed in OFC, but it suggests how cognitive rules (presumably maintained in dlPFC) or self-monitoring of performance (vmPFC) can influence the motivational values computed in OFC.

Frank and Claus provided a computational model of the close cooperation between OFC and the ventral BG in flexible decision making [13]. They build upon a previous model of action selection in the BG which was able to learn probabilistically rewarded stimulus-response association, but which failed at learning long-term strategies such as in gambling tasks (e.g. frequent small losses and rare large gains that are on average beneficial). This is explained by the nature of the dopaminergic signal: its stereotypic phasic activation can not represent adequately the respective importances of gains and losses. They added to this model a strategical module composed of the OFC (with the medial part representing the current outcome-response expectancies and the lateral part representing the previously experienced reinforcements) and of BLA (representing the magnitude of the reward). The model successfully learn the Iowa gambling task, reversal learning, and devaluation paradigms. It explains various aspects of decision making deficits in OFC damaged patients or Parkinsons patients, and explains risk aversion. This dichotomy between the evaluation of the frequency of reward delivery by the DA system, and the valuation of behavioral strategies in OFC is of particular biological relevance in order to understand the role of emotions in decision-making.

Dranias et al. proposed an anatomically realistic model of motivation called MOTIVATOR that successfully addresses classical conditioning, visual discrimination, devaluation, extinction and reversal learning [12]. Its architecture contains most of the structures depicted on Fig. 1. It combines visual and gustatory inputs to the sensory cortex with the organism's needs (or drive, like hunger or thirst) in LH to compute the motivational value of outcomes and to influence behavior. The network composed of the amygdala and LH combines sensory and internal information to learn CS-US associations and provides the emotional value of the CS to the OFC (again with its lateral and medial parts) which represents the relative motivational value of responding to one of the multiple simultaneously present objects. This motivational value is modulated by the ventral BG which tracks the discrepancy between the expected and actual emotional values. When they do not match, the DA signal updates motivational values in OFC. Dues to its biological plausibility, this model reproduces a lot of experimental evidence, both on neuronal firing patterns or behavioral measurements like saccadic response time or blood pressure changes, and provides several testable predictions. It studies extensively the effect of satiety on behavior, or more generally the role of the needs of the organism in the formation of choices, what is neglected in most models: they suppose that any rewarding object is incentive for action.

5 Conclusion

Neurobiological computational models are very useful to understand brain processes and human behavior but also to give valuable insights in order to design intelligent agents that could interact with humans. The models we reviewed here demonstrate that the dichotomy between emotion and cognition does not lead to efficient cognitive strategies if they are thought as opponent processes. On the contrary, they interact together with the common goal to improve the usefulness of behavior. The computational study of classical conditioning, that may seem too simple for AI, has led to a better understanding of the role of the prediction of the emotional value of an object in motivated behavior. It highlights that potential goals are evaluated in terms of emotional significance before being processed by cognitive structures and has led to efficient improvement of RL techniques.

The same procedure may lead to a better understanding of flexible cognitive abilities. Let us consider for example the case of risk-taking: in order to reach a highly valuable object, a robot may need to take the risk to go through a series of intermediary actions that could probably harm it. Depending on the value of the reward and on the probability to be injured (that can depend on its own estimation of performance), the robot should deliberate to decide whether it engages the action towards the reward, wait for help or even try to build another strategy. This difficult cognitive ability to estimate trade-offs between long-term rewards and potentially immediate punishments in a sequential strategy may be implemented by a computational model derived from the anatomical areas depicted on Fig. 1. LH and the amygdala can learn to assign to both rewarding and punishing events a value that can compared on the same basis. The recurrent structure of the loop between OFC and the ventral BG can compute the motivational value of the reward by decomposing the sequence of actions required to obtain it (planned by dlPFC) and estimating the associated risks. Self-confidence about its own performance can be monitored by vmPFC and inhibit the amygdala, leading to a decreased aversive value for the intermediary steps. Such a computational model would allow to design a cognitive agent that bases its decisions not only on a model of the external world, but also on its own beliefs about its needs and performance, which should improve its communicative abilities and "human-like" way of thinking.

The expression of an adequate behavior can not solely rely on the cognitive manipulation of symbols, but the different events occurring in a dynamical environment should

rather be emotionally valued in order to: (1) incite the system to act, either by responding or by foraging, and (2) prioritize the available options. This emotional valuation is only possible when the cognitive system is not considered in isolation from its surrounding organism, but rather when it is remembered that its ultimate function is to maintain the body's homeostasis at different levels: finding food, water or a sexual partner, searching for pleasant stimulations and avoiding painful ones, ameliorating the self-esteem of the individual and its social integration. This emphasizes particularly the importance of embodied cognition, whether in virtual reality or in robotics.

Acknowledgements The authors are in part supported by the German research foundation (Deutsche Forschungsgemeinschaft) grant (DFG HA2630/4-2) "The cognitive control of visual perception and action selection".

References

1. Armony JL, Servan-Schreiber D, Cohen JD, Ledoux JE (1997) Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. Trends Cogn Sci 1(1):28–34
2. Bach J (2009) Principles of synthetic intelligence PSI: an architecture of motivated cognition. Oxford University Press, Oxford
3. Balleine BW, Liljeholm M, Ostlund SB (2009) The integrative function of the basal ganglia in instrumental conditioning. Behav Brain Res 199(1):43–52
4. Bastiaansen JACJ, Thioux M, Keysers C (2009) Evidence for mirror systems in emotions. Philos Trans R Soc Lond B, Biol Sci 364(1528):2391–2404
5. Bechara A, Damasio H, Damasio AR (2000) Emotion, decision making and the orbitofrontal cortex. Cereb Cortex 10(3):295–307
6. Becker-Asano C, Wachsmuth I (2010) Affective computing with primary and secondary emotions in a virtual human. Auton Agents Multi-Agent Syst 20(1):32–49
7. Brown J, Bullock D, Grossberg S (1999) How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. J Neurosci 19(23):10 502–10 511
8. Cahill L, Babinsky R, Markowitsch HJ, McGaugh JL (1995) The amygdala and emotional memory. Nature 377(6547):295–296
9. Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. J Neurosci 29(39):12315–12320
10. Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8(12):1704–1711
11. Deco G, Rolls ET (2005) Synaptic and spiking dynamics underlying reward reversal in the orbitofrontal cortex. Cereb Cortex 15(1):15–30
12. Dranias MR, Grossberg S, Bullock D (2008) Dopaminergic and non-dopaminergic value systems in conditioning and outcome-specific revaluation. Brain Res 1238:239–287
13. Frank MJ, Claus ED (2006) Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychol Rev 113(2):300–326
14. Grossberg S (1971) On the dynamics of operant conditioning. J Theor Biol 33(2):225–255
15. Grossberg S (1972) A neural theory of punishment and avoidance, i: Qualitative theory. Math Biosci 15:39–67
16. Grossberg S, Levine DS (1987) Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, interstimulus interval, and secondary reinforcement. Appl Opt 26(23):5015–5030
17. Grossberg S, Seidman D (2006) Neural dynamics of autistic behaviors: cognitive, emotional, and timing substrates. Psychol Rev 113(3):483–525
18. Haber S (2008) Parallel and integrative processing through the basal ganglia reward circuit: lessons from addiction. Biol Psychiatry 64(3):173–174
19. Humphries MD, Prescott TJ (2010) The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Prog Neurobiol 90(4):385–417
20. Koenigs M, Grafman J (2009) Posttraumatic stress disorder: the role of medial prefrontal cortex and amygdala. Neuroscientist 15(5):540–548
21. Lazarus RS (1999) The cognition-emotion debate: A bit of history. In: Dalgleish T, Power MJ (eds) Handbook of cognition and emotion. Wiley, New York, pp 3–19
22. LeDoux J (2003) The emotional brain, fear, and the amygdala. Cell Mol Neurobiol 23(4–5):727–738
23. Levine DS (2007) Neural network modeling of emotion. Phys Life Rev 4(1):37–63
24. Murray EA (2007) The amygdala, reward and emotion. Trends Cogn Sci 11(11):489–497
25. Ochsner KN, Gross JJ (2005) The cognitive control of emotion. Trends Cogn Sci 9(5):242–249
26. Ohman A, Flykt A, Esteves F (2001) Emotion drives attention: detecting the snake in the grass. J Exp Psychol Gen 130(3):466–478
27. Ono T, Nakamura K, Nishijo H, Fukuda M (1986) Hypothalamic neuron involvement in integration of reward, aversion, and cue signals. J Neurophysiol 56(1):63–79
28. O'Reilly RC, Frank MJ, Hazy TE, Watz B (2007) Pvlv: the primary value and learned value Pavlovian learning algorithm. Behav Neurosci 121(1):31–49
29. Picard R (ed) (1997) Affective Computing. MIT Press, Cambridge
30. Rao RPN (2010) Decision making under uncertainty: a neural model based on partially observable Markov decision processes. Front Comput Neurosci 4:146
31. Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: Variation in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF (eds) Classical Conditioning II: Theory and Research. Appleton Century Crofts, New York
32. Richter-Levin G (2004) The amygdala, the hippocampus, and emotional modulation of memory. Neuroscientist 10(1):31–39
33. Roesch MR, Olson CR (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. Science 304(5668):307–310
34. Scherer KR, Shorr A, Johnstone T (eds) (2001) Appraisal processes in emotion: theory, methods, research. Oxford University Press, Canary
35. Schmajuk NA, DiCarlo JJ (1992) Stimulus configuration, classical conditioning, and hippocampal function. Psychol Rev 99(2):268–305
36. Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275(5306):1593–1599
37. Suri RE, Schultz W (2001) Temporal difference model reproduces anticipatory neural activity. Neural Comput 13(4):841–862
38. Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88(2):135–170
39. Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. MIT Press, Cambridge. A Bradford Book

40. Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G (2009) The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. Neuron 62(2):269–280
41. Vitay J, Fix J, Beuth F, Schroll H, Hamker FH (2009) Biological models of reinforcement learning. Künstl Intell 3:12–18
42. Vitay J, Hamker FH (2010) A computational model of basal ganglia and its role in memory retrieval in rewarded visual memory tasks. Front Comput Neurosci 4:1–18
43. Vuilleumier P (2005) How brains beware: neural mechanisms of emotional attention. Trends Cogn Sci 9(12):585–594
44. Ziemke T (2008) On the role of emotion in biological and robotic autonomy. Biosystems 91:401–408

Julien Vitay is associate lecturer at the Department of Computer Science of the Chemnitz University of Technology. He received an Engineer degree from the Supélec engineering school in Rennes (France) in 2002 and a PhD in Computer Science from the University of Nancy in 2006. He then made a postdoctoral internship at the Department of Psychology of the Westf. Wilhelms-University Münster.

Fred H. Hamker is professor of Artificial Intelligence at the Department of Computer Science of the Chemnitz University of Technology. He received his diploma in Electrical Engineering from the University of Paderborn in 1994 and his PhD in Computer Science at the TU Ilmenau in 1999. He was a Post-Doc at the University of Frankfurt and the California Institute of Technology (Pasadena, USA). In 2008 he received his venia legendi from the Department of Psychology at the Westf. Wilhelms-University Münster.