Emotions: from brain to robot

Michael A. Arbib[1] and Jean-Marc Fellous[2]

[1]Computer Science, Neuroscience and USC Brain Project, University of Southern California, Los Angeles, CA 90089-2520, USA
[2]Biomedical Engineering Department and Center for Cognitive Neuroscience, Duke University, Durham, NC 27708-0281, USA

Some robots have been given emotional expressions in an attempt to improve human–computer interaction. In this article we analyze what it would mean for a robot to have emotion, distinguishing emotional expression for communication from emotion as a mechanism for the organization of behavior. Research on the neurobiology of emotion yields a deepening understanding of interacting brain structures and neural mechanisms rooted in neuromodulation that underlie emotions in humans and other animals. However, the chemical basis of animal function differs greatly from the mechanics and computations of current machines. We therefore abstract from biology a functional characterization of emotion that does not depend on physical substrate or evolutionary history, and is broad enough to encompass the possible emotions of robots.

Interest in the creation of robots with emotions is fourfold. First, current technology already shows the value of providing robots with 'emotional' expressions (e.g. computer tutors) and bodily postures (e.g. robot pets) to facilitate human–computer interaction. Second, this raises the question of the possible value (or inevitability) of future robots not only simulating emotional expression but actually 'having emotions'. Third, this in turn requires us to re-examine the neurobiology of emotion to generalize concepts first developed for humans and then extended to animals so that the question of robot emotions becomes meaningful. And fourth, this suggests in turn that building 'emotional robots' could also provide a novel test-bed for theories of biological emotion.

This article samples the state of the art on current robot technology, and examines recent work on the neurobiology of emotions, to ground our suggestions for a scientific framework in which to approach robot 'emotions'. The question of 'emotional robots' being used to test theories of biological emotion is of great interest, but beyond the scope of this article.

Different kinds of emotions

There is a wide spectrum of feelings, from the 'motivation' afforded by drives such as the search for food afforded by hunger [1] to 'emotions' in which, at least in humans, cognitive awareness might be linked to feeling the 'heat' of love, sorrow or anger, and so on. But as we have no criterion for saying that a robot has 'feelings', we will seek here to understand emotions in their functional context, noting that not all emotions need be like human emotions. We analyze emotion in two main senses:

(1) Emotional expression for communication and social coordination.

(2) Emotion for organization of behavior (action selection, attention and learning).

The first concerns 'external' aspect of emotions; the second 'internal' aspects. In animals, these aspects have co-evolved. How might they enter robot design? Both robots and animals need to survive and perform efficiently within their 'ecological niche' and, in each case, patterns of coordination will greatly influence the suite of relevant emotions (if such are indeed needed) and the means whereby they are communicated.

A key function of emotion is to communicate simplified but high impact information. A scream is extremely poor in information (it says nothing about the cause for alarm), but its impact on others is high. Moreover, neurobiology shows that simplified but high impact information is communicated between brain areas, through the very different 'vocabulary' of neuromodulation.

The similarity in facial expressions between certain animals and humans prompted classic evolutionary analyses [2], which support the view that mammals (at least) have emotions (although not necessarily the same as human emotions), and work reviewed below explores their (neuro)biological underpinnings. What of robots? Robots are mechanical devices with silicon 'brains', not products of biological evolution. But as we better understand biological systems we will extract 'brain operating principles' that do not depend on the physical medium in which they are implemented. These principles might then be instantiated for particular robotics architectures to the point where we might choose to speak of robot-emotions.

Many researchers (see for example [3,4]) have proposed explicit functions for emotions: coordinating behavioral response systems; shifting behavioral hierarchies; communication and social bonding; short-cut cognitive processing; and facilitating storage and recall of memories. However, emotions are not always beneficial [5] – if one is caught in a traffic jam, anger can easily set in, but anger in this case has no apparent usefulness. How does the brain maximize the benefits of emotion yet minimize its occasional inappropriateness? And how would the understanding of such tradeoffs affect our ideas about robot designs?

From ethology to robot motivation and emotion

We advance our discussion by reviewing work that has added emotion-like features to robotic systems [6], some

Corresponding authors: Michael A. Arbib (arbib@pollux.usc.edu),
Jean-Marc Fellous (fellous@duke.edu).

work inspired by ethology (the study of animal behavior), and we then sample attempts to analyze the role of emotion in general 'cognitive architectures' at the interface between artificial intelligence and cognitive science.

Robot ethology

In the 'ethological robots' reviewed here, each drive, perceptual releaser, motor-, and emotion-related process is modeled as a different 'specialist' [7] or 'schema' [8]. Each schema computes on the basis of its inputs to update its 'activation level' and internal state, and sends on appropriate outputs. A robot controller based on the ethology of the preying mantis [9] provided four different motivated visuomotor behaviors: prey acquisition, predator avoidance, mating and chantlitaxia (coined by Rolando Lara, from the Nahuatl word *chantli* for 'shelter' and the Latin *taxia* for 'attraction'). For prey acquisition, hunger is the primary motivator; for predator avoidance, fear serves similarly; for mating the sex drive dominates. The behavioral controller was implemented on a small hexapod robot.

In Bowlby's [10] theory of attachment, infants view certain individuals as sources of comfort, with the 'comfort zone' depending on the circumstances. Likhachev and Arkin [11] extrapolated these ideas to produce useful behavior in autonomous robots, rather than to model the human child. The 'comfort zone' ensures that the robot does not stray from a given task or area; it can also provide a basis for creating a robot pet that can relate to a particular human being.

Studies of canine ethology support work on human–robot bonding for Sony's AIBO [12], a speaking dog-like entertainment robot. Ekman's model [13], with its six basic emotional states, has been influential in work on emotional expression in robots. The Kismet robot [14], for example, can communicate an emotive state and other social cues to a human through facial expressions, gaze, and quality of voice. The computations needed to communicate an 'emotional state' to a human might also improve the way robots function in the human environment.

In order to interact with another agent, it is essential to have a good conceptual model for how this agent operates [15]. As complexity of environment and interactions increases, the social sophistication of a robot interacting with humans must be scaled accordingly. Some would argue that this entails that the robot 'has' emotions, but others would distinguish having a model of emotions of the other agent from having emotions oneself. This in some sense reverses the simulation theory of human empathy [16,17]: in this theory, there is no question that the human 'has' emotion, and the proposal is that the system for expressing one's own emotions drives the ability to recognize those of others.

Cognitive architectures

We now turn to general 'cognitive architectures', in which the role of emotion can be situated at several levels. Ortony *et al.* [18] analyze the interplay of affect (value), motivation (action tendencies), cognition (meaning), and behavior at three levels of information processing:

(1) Reactive: a hard-wired releaser of fixed action patterns and an interrupt generator. This level has only the most rudimentary affect.

(2) Routine: the locus of unconscious well-learned automatized activity and primitive and unconscious emotions.

(3) Reflective: the home of higher-order cognitive functions, including metacognition, consciousness, self-reflection, and full-fledged emotions.

Ortony *et al.* address the design of emotions in computational agents (these include 'softbots' as well as embodied robots) that must perform unanticipated tasks in unpredictable environments. They argue that such agents, if they are to function effectively, must be endowed with curiosity and expectations, and a 'sense of self' that reflects parameter settings that govern the agent's functioning.

Sloman [19] also offer a three-level view of central processes:

(1) Reactive: producing immediate actions. When inconsistent reactions are simultaneously activated one may be selected by a competitive mechanism.

(2) Deliberative: using explicit hypothetical representations of alternative possible predictions or explanations, comparing them and selecting a preferred option.

(3) Meta-management: allowing internal processes to be monitored, categorized, evaluated and controlled or modulated.

Sloman also notes the utility of an 'alarm' system – a reactive component that gets inputs from and sends outputs to all the other components and detects situations where rapid global redirection of processing is required.

To reconcile the two schemes, we suggest using four levels: reactive, routine, reflective–deliberative, and reflective–meta-management.

Finally, recent work in multi-agent teamwork suggests that in virtual organizations, agents that simulate emotions would be more believable to humans and could anticipate human needs by appropriate modeling of others [15,20].

The neurobiological roots of emotion

We turn now from outlining a functional view of robots to review research linking brain and emotion [3,21–23], before proposing a functional framework for synthesis. We shall see that: (i) emotion is not computed by a centralized neural system; (ii) emotions operate at many time scales and at many behavioral levels; and (iii) there is no easy separation between emotion and cognition.

Behavioral control columns

An animal comes with a set of basic 'drives' that provide the 'motor' (motivation) for behavior. Most of the neural circuitry underlying these drives involves specific nuclei of the hypothalamus. Swanson [24] introduced the notion of the 'behavioral control column' (Figure 1), comprising interconnected sets of nuclei in the hypothalamus underlying specific survival behaviors: spontaneous locomotion, exploration, ingestive, defensive and reproductive behaviors. The hypothalamus sends this information to higher centers such as the amygdala and the orbitofrontal cortex.
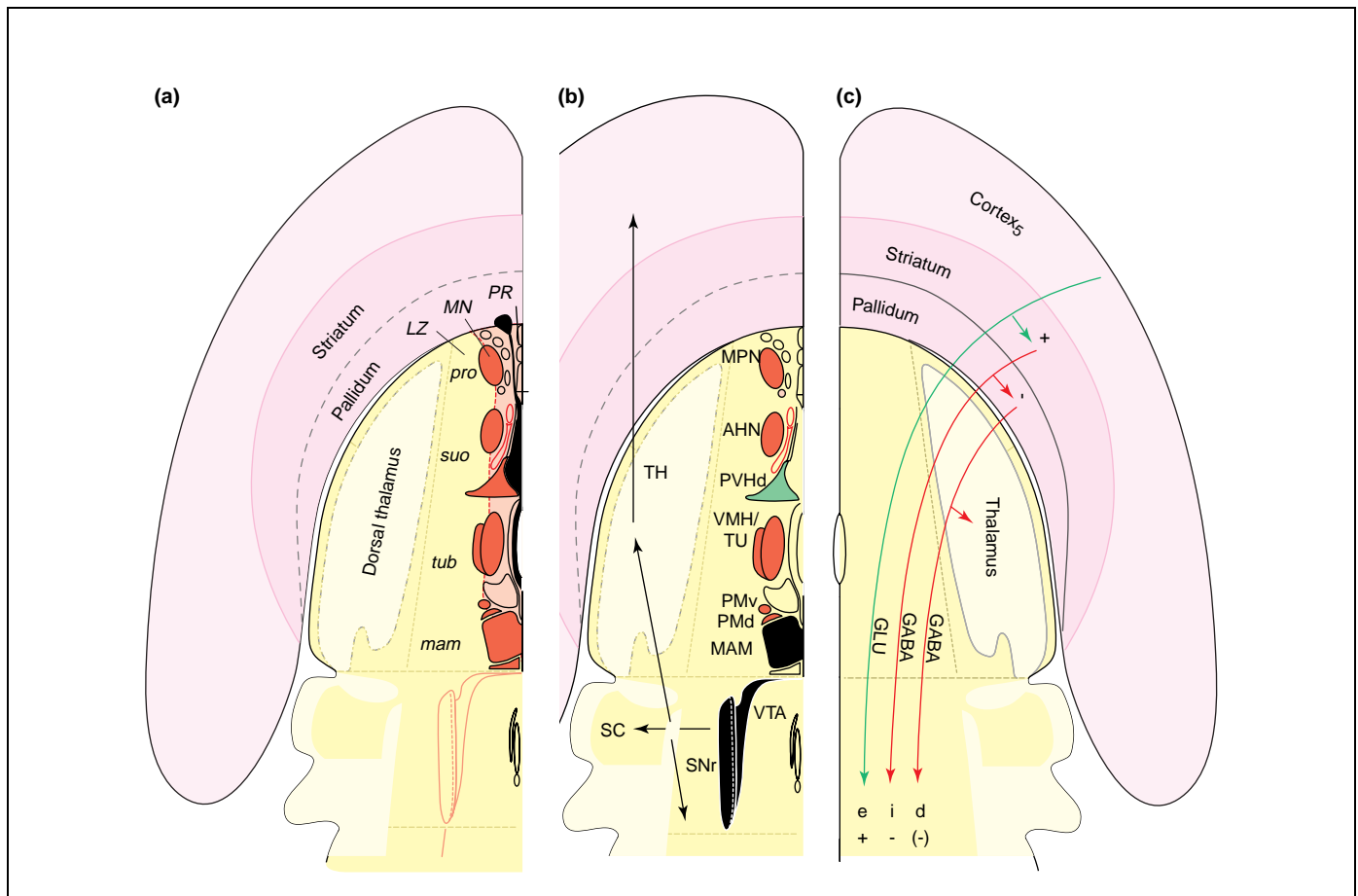
Figure 1. Major features of interactions of the behavioral control column and cerebral cortex in regulation of motivated behavior, as seen on flatmaps of the rat central nervous system. **(a)** The neuroendocrine motor zone shown in black and three subgroups of hypothalamic nuclei: the periventricular region (PR) which contains visceromotor pattern generators; the medial nuclei (MN); and the lateral zone (LZ). **(b)** Almost all nuclei in the behavioral control column generate a dual projection, descending to the motor system and ascending to thalamocortical loops. **(c)** The embedding of the column in cortical computations. This prototypical circuit element consists of an excitatory projection from cortex with a collateral to the striatum (the input system for the basal ganglia which play a key role in the sequencing and interleaving of actions). The striatum then generates an inhibitory projection to the motor system with a collateral to the pallidum (the output system for the basal ganglia). Finally, the pallidum generates an inhibitory projection to the brainstem motor system, with a collateral to the dorsal thalamus (which projects back to cerebral cortex). This pallidal projection is disinhibitory because it is inhibited by the striatal input. (Adapted from Swanson [24], Figs 8,10,14, respectively, whose captions explain abbreviations not needed in this article).

Amygdala and cerebral cortex

The human ability to plan behaviors on the basis of future possibilities rather than present contingencies alone has been linked to the increased size and differentiation of cerebral cortex [25,26]. Kelley stresses that feedforward hypothalamic projections provide the motivational network with access to associative and cognitive cortical areas [27]. The amygdala can influence cortical areas via feedback from proprioceptive, visceral or hormonal signals, via projections to various 'arousal' networks, and through interaction with the medial prefrontal cortex [28] (Figure 2a). The prefrontal cortex, in turn, sends distinct projections back to several regions of amygdala, allowing elaborate cognitive functions to regulate the amygdala's roles in emotion. For example, our ability to create heuristics and general rules from our everyday experiences has been shown to depend on prefrontal cortices [29] and their ability to bias activity in target structures [30]. Because of the tight interactions between amygdala and prefrontal cortex, it is likely that our ability to generalize and abstract is directed by (and influences, in turn) some aspects of our emotional state. How this is done, and how

robots could take advantage of it remains an open question. Some functional connections between the amygdala, thalamus and cortical areas allow for both fast elicitation of emotion and more refined emotion control based on memory and high-level sensory representations (Figure 2a). The role of the amygdala in the experience and expression of fear has received particular study [21,31]. Stimuli that elicit fear reactions can be external (e.g. a loud noise) or internal, from the behavioral control columns or from memory structures such as hippocampus or prefrontal cortex. Decision-making ability in emotional situations is also impaired in humans with damage to the medial prefrontal cortex, and abnormalities in prefrontal cortex might predispose people to develop fear and anxiety disorders [32].

Activation of the human amygdala can be produced by observing facial expressions, and lesions of the human amygdala can cause difficulty in the identification of some such expressions [33,34]. The amygdala and prefrontal cortices are therefore involved in social as well as internal aspects of emotion, and together play a crucial role in the regulation of emotion, a key
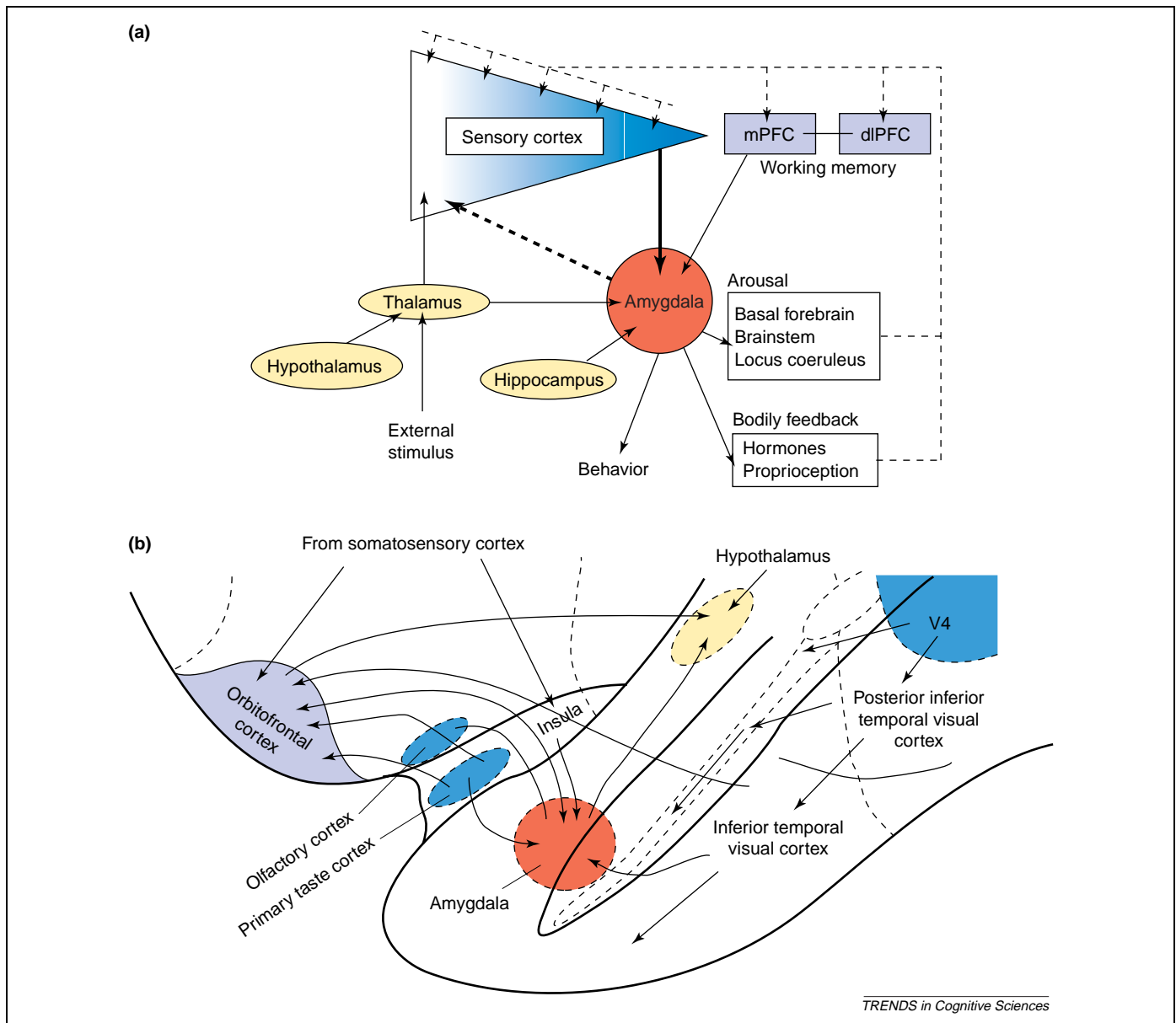
Figure 2. (a) Interaction of the amygdala with cortical areas in the mammalian brain (adapted from [32]). **(b)** Lateral view of part of the macaque monkey brain emphasizing how the orbitofrontal cortex (involved in social emotions) links to amygdala, and to sensory cortices. V4 is visual area 4 (adapted from [3]).

component of affective style, affective disorders [35] and social interactions [36].

Figure 2b provides a view of how orbitofrontal cortex links to the amygdala in macaque monkey. Damage to monkey caudal orbitofrontal cortex produces emotional changes that include the tendency to respond inappropriately. Orbitofrontal neurons also serve as part of a mechanism that evaluates whether a reward is expected, and different subregions of the prefrontal cortex are selectively involved during positive rewards or punishments [37]. Dolcos *et al.* have shown that different subregions of the medial temporal lobe memory system are selectively and differentially involved for emotional and neutral stimuli in human, and that this area was strongly correlated with amygdala activations during emotional stimuli [38]. Many other brain areas have been involved in the experience and expression of

emotions in humans, including the anterior cingulate cortex, insula, hippocampus and fusiform gyrus [39,40].

The mirror system, language and empathy

Going beyond the hypothalamo-amygdala-cortical inter-actions, we note that language plays a unique role in shading and refining human emotions. It is therefore interesting that recent research suggests that 'mirror neurons' might provide the substrate for both 'empathy' – the ability to recognize the emotional dispositions of others – and communication through language.

In monkey, parietal area AIP [41] processes visual information concerning objects to extract possibilities for manual interaction and is reciprocally connected with the so-called 'canonical neurons' of ventral premotor area F5 [42] whose discharge correlates with more-or-less specific hand actions. Certain F5 neurons, called mirror

neurons [43,44], discharge when the monkey observes the experimenter make a gesture similar to one that, when actively performed by the monkey, involves activity of that neuron. PET experiments in humans showed that superior temporal sulcus (STS), the inferior parietal lobule, and the inferior frontal gyrus (area 45) in the left hemisphere were significantly activated for both grasp execution and grasp recognition [45,46]. Area 45 is part of Broca's area, a major component of the human brain's language mechanisms. F5 is considered to be the monkey homologue of Broca's area.

These findings grounded the hypothesis that the mirror system provided the basis for the evolution of human language via intermediate stages involving 'complex' imitation (acquiring novel sequences of abstract actions in a few trials), protosign (manually-based communication, enabled by freeing action from praxis to be used in pantomime and then conventionalized communication) and protospeech (vocally-based communication exploiting the brain mechanisms that support protosign) [47,48].

However, mirror neurons have also been implicated in empathy – but with the emphasis now on recognizing facial expressions instead of manual actions. Adolphs [49] and Ochsner [36] stress the important role of social interaction in the forming of emotions. Clearly, human emotions are greatly influenced by our ability to empathize with the behavior of other people [50]. Indeed, some have suggested that mirror neurons can contribute not only to 'simulating' other people's actions as the basis for imitation [51], but also 'simulating' other people's feelings as the basis for empathy [16,17,52].

Neuromodulation

We now switch structural levels, turning from specific brain structures to systems of neuromodulation. Neuromodulation refers to the action on nerve cells of endogenous substances called neuromodulators. These are released by a few specialized brain nuclei that have somewhat diffuse projections throughout the brain and receive inputs from brain areas that are involved at all levels of behavior from reflexes to cognition. Each neuromodulator typically activates specific families of receptors in neuronal membranes. The receipt of its own neuromodulator by a receptor has very specific effects on the neuron at various time scales, from a few milliseconds to minutes and hours [53]. Each neuron has its own mixture of receptors, depending on where it is located in the brain.

Kelley's [27] analysis of motivation and emotion emphasizes three neuromodulatory systems – those for dopamine, serotonin and opioids (Box 1). Strikingly, although these three neuromodulatory systems seem to be distinct from each other in their overall functionalities, they each exhibit immense diversity and synergism of behavioral consequences. The different effects depend on both molecular details (the receptors) and global arrangements (the circuitry within the modulated brain region, and the connections of that region within the brain).

Neuromodulation thus provides a simple but high-impact signal that can fundamentally change the way single neurons and synapses 'compute' and in this sense is

> **Box 1. Three main neuromodulatory systems involved in emotion**
>
> **Dopamine**
> In the mammalian brain, dopamine appears to play a major role in motor activation, appetitive motivation, reward processing and cellular plasticity, and might be important in emotion. Dopamine is contained in two main pathways that ascend from the midbrain to innervate many cortical regions. Dopamine neurons in the monkey have been observed to fire to predicted rewards [67,68]. Moreover, dopamine receptors are essential for the ability of prefrontal networks to hold neural representations in memory and use them to guide adaptive behavior. Therefore, dopamine plays essential roles all the way from 'basic' motivational systems to working memory systems essential for linking emotion, cognition and consciousness.
>
> **Serotonin**
> Serotonin has been implicated in behavioral state regulation and arousal, motor pattern generation, sleep, learning and plasticity, food intake, mood and social behavior [69]. The cell bodies of serotonergic systems are found in midbrain and pontine regions in the mammalian brain and have extensive descending and ascending projections. Serotonin plays a crucial role in the modulation of aggression and in agonistic social interactions in many animals. In crustaceans, serotonin plays a specific role in social status and aggression; in primates, with the system's expansive development and innervation of the cerebral cortex, serotonin has come to play a much broader role in cognitive and emotional regulation, particularly control of negative mood or affect. The serotonin system is the target of many widely used anti-depressant drugs.
>
> **Opioids**
> The opioids, which include endorphins, enkephalins and dynorphins, are found particularly within regions involved in emotional regulation, responses to pain and stress, endocrine regulation and food intake. Increased opioid function is associated with positive affective states such as relief of pain, and feelings of euphoria, well-being or relaxation. Activation of opioid receptors promotes maternal behavior in mothers and attachment behavior and social play in juveniles. Separation distress, exhibited by archetypal behaviors and calls in most mammals and birds, is reduced by opiate agonists and increased by opiate antagonists in many species [70]. Opiates can also reduce or eliminate the physical sensation induced by a painful stimulus, as well as the negative emotional state it induces. Opioids and dopamine receptors are two major systems affected by common drugs of abuse.

akin to the alarm system of Sloman [19]. Earlier, we said that reconciling the cognitive architectures of Ortony *et al.* and Sloman led us to consider four architectural levels. Fellous [54] has also produced a four-level hierarchy, but this time on the basis of a review of data on hypothalamus, amygdala and orbitofrontal cortex, and the suggestion that the neural basis for emotion involves both computations in such structures and their current state of neuromodulation (see Figure 3, and [55,56]). Others have suggested that neuromodulation might be a key to meta-learning [57].

Towards a functional view of emotions

Emotions are, of course, far more complex than a few brain structures and three ascending modulatory systems that interact with them. We can only outline the lessons that neurobiology offers for the study of robotic emotions, not provide all the details. However, we stress that the richness of human emotion is in part due to the linkage
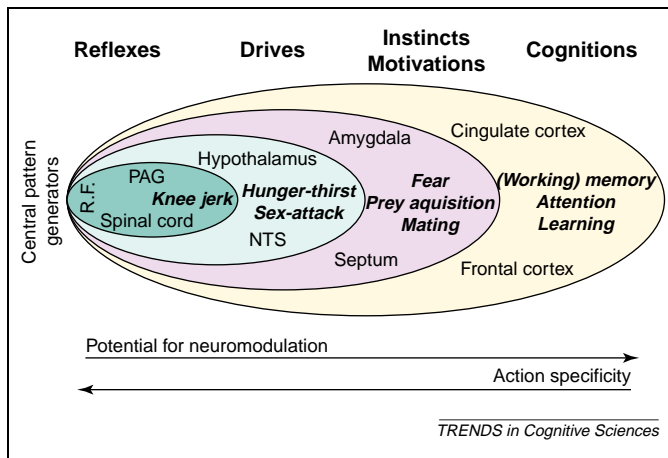
Figure 3. Organization of behavior with respect to potential for neuromodulation and action specificity. Reflexes are fixed, highly specific, motor patterns, the neural substrate of which undergoes few neuromodulations. At the other end of the scale, 'cognitions' are unspecific with respect to sensory stimuli but heavily susceptible to neuromodulation. The colored ellipses represent zones of neural recruitment during emotional expression and experience. It is argued that the neural substrate of emotion is intimately linked to that of neuromodulation. The diagram includes the mapping of brain structures to Reflexes, Drives, Instincts and Motivations, and Cognitions. Abbreviations: PAG: Periaqueductal gray; R.F.: reticular formation; NTS: nucleus of the solitary tract. See [54] for details.

of emotion with cognitive processes of great subtlety. The neural substrate of emotion involves many interacting structures, from the hypothalamus to the orbitofrontal cortex to the diverse mirror systems, which have clear 'cognitive' as well as emotional roles. Having said this, we turn from specific lessons of neurobiology to the question of how to realize these lessons in robots that do not share the biology of living creatures. For example, the effects of neuromodulation might be simulated in robots by large-scale broadcast signals that have specific and local effects on their computational targets, and some of these effects could yield behavioral states analogous to emotional states [58].

Appraisal theory [59] develops a catalogue of human emotions and seeks to provide a computational account of the 'appraisals' that lead one emotion to be evoked over another, but without the 'heat' provided by emotion's biological underpinnings in humans. Appraisal theory is therefore a good candidate for algorithms that enable a robot to simulate the appearance of human emotional behavior. However, a truly fruitful theory of robot emotions must address the fact that many robots will have no human–computer interface. A more abstract view of emotion is required.

In biology, the four Fs (feeding, fighting, fleeing and reproduction) [60] are paramount. However, robot design will normally be based on the availability of a reliable power supply, and the reproduction of robots will be left to factories. Thus the 'ecological niche' of a robot will tend to be defined in terms of the set of tasks it is to perform. S-RETIC [61] modeled the brain's reticular formation as a stack of modules, each with a slightly different selection of inputs, but each module trying to decide to which four-F-like mode to commit the organism. They would communicate back and forth, competing and cooperating until reaching a consensus that would switch the organism's mode. Within any mode of behavior, many different acts

are possible: if the cat should flee, will it take the mouse first or leave it, climb a tree or skirt it? The notion is that a hierarchical structure that computes modes, and then acts within modes, might in some sense be better than one that tries to determine successive acts directly.

Robot emotions?
Arbib [5] has suggested that these modes might be akin to 'robot emotions' (or, at least, motivational systems). Consider a robot with a set of basic functions, each with appropriate perceptual schemas and access to various motor schemas. Each perceptual schema evaluates the current state to come up with an 'urgency level' for activating various motor schemas, as well as determining appropriate motor parameters. A competition mechanism can adjudicate between these processes. A motivational system can be defined as a state-evaluation process that can adjust the relative weighting of the different functions, raising the urgency level for one system while lowering the motivation system for others, depending on the context.

One can then generalize 'modes' to abstract groups of tasks, with many strategies gathered into a small number of modes. When faced with a problem, it will then in general be more efficient to select an appropriate mode first and then select a strategy from within that mode. The catch is that there might be cases in which rapid commitment to one group of strategies would preclude finding the most appropriate strategy – possibly with disastrous consequences on occasions. Effective robot design would therefore have to balance this 'fast commitment' process against more subtle evaluative processes that can check the suitability of a chosen strategy before committing to it completely (and thus we return to the varied levels of 'cognitive architectures'). We might then liken 'motivation' to biases that favor one strategy group over another, and 'emotion' to the way in which these biases interact with more subtle computations. One way emotions might become maladaptive, then, is when biases favoring rapid commitment to one mode overwhelm more cautious analysis of available strategies.

Earlier, we introduced two 'views' of emotion: (1) emotional expression for communication and social coordination; and (2) emotion for the organization of behavior. The very utility of 'modes' in the robot's own decision making (sense 2 of emotion) might well make fast communication of modal commitment an effective way to coordinate robots or foster human–robot integration (sense 1) [15].

Interactions between emotion and cognition
Emotion changes the operating characteristics of cognition and action selection. There is neuroscientific evidence that affect is a prerequisite for establishing long-term memories [62]. It also has important consequences for the allocation of attention. Fear tends to focus attention on local details, whereas under conditions of positive affect, people tend to focus on the bigger picture [63,64].

Operating-system-like tasks such as resource mobilization and prioritization of behaviors rely on compact signals that have high-impact on the functioning of an

autonomous agent. The underlying hypothesis is that hierarchical assessment of task priorities can help robots cope with the complexity and unreliability of the (external or internal) environment – mirroring the rich interaction of cognitive and emotional processing in the mammalian brain. Adaptive mechanisms can adjust the hierarchy and grouping of tasks to match the constraints imposed by time, physical limits and energy resources. In this regard, note that human emotion becomes maladaptive not only when rapid commitment processes overwhelm more cautious analysis, but also when worry or other related emotions that favor more cautious analysis overwhelm adaptive rapidity. We leave the reader to think through computational variants on this theme.

The stage seems set for dramatic progress in integrating brain and society in theories of emotion in animals and humans, and for linking solo tasks, robot–human interaction and teamwork in further exploring the question of whether an autonomous robot would need emotions. Building on the work of the past few years (see overviews in [65,66]), the study of brains will not only continue to inform our analysis of robots, but the precision required to make explicit the computational strategies of robots will enrich the vocabulary for the study of motivation and emotion in animal and human.

References

1 Watts, A.G. (2003) Motivation: neural substrates. In *The Handbook of Brain Theory and Neural Networks* (2nd edn) (Arbib, M.A., ed.), pp. 680–683, MIT Press
2 Darwin, C. (1872) *The Expression of Emotions in Man and Animals*, Julian Friedman, London
3 Rolls, E.T. (1999) *The Brain and Emotion*, Oxford University Press
4 Levenson, R.W. (2003) Blood, sweat, and fears: the autonomic architecture of emotion. *Ann. N. Y. Acad. Sci.* 1000, 348–366
5 Arbib, M.A. Beware the passionate robot. In *Who Needs Emotions? The Brain Meets the Robot* (Fellous, J-M. and Arbib, M.A., eds), Oxford University Press (in press)
6 Picard, R.W. (1997) *Affective Computing*, MIT Press
7 Minsky, M. (1988) *The Society of Mind*, Simon & Schuster
8 Arkin, R.C. (1989) Motor schema-based mobile robot navigation. *Int. J. Robotics Res.* 8, 92–112
9 Arkin, R.C. *et al.* (2000) Behavioral models of the praying mantis as a basis for robotic behavior. *J. Robotics Auton. Syst.* 32, 39–60
10 Bowlby, J. (1969) *Attachment and Loss. I: Attachment*, Hogarth Press
11 Likhachev, M. and Arkin, R.C. (2000) Robotic comfort zones. In *Proc. SPIE Sensor Fusion and Decentralized Control in Robotic Systems III*, pp. 27–41, SPIE
12 Arkin, R. *et al.* (2003) An ethological and emotional basis for human–robot interaction. *Rob. Auton. Syst.* 42, 3–4
13 Ekman, P. (2003) Darwin, deception, and facial expression. *Ann. N. Y. Acad. Sci.* 1000, 205–221
14 Breazeal, C. and Scassellati, B. (2002) Robots that imitate humans. *Trends Cogn. Sci.* 6, 481–487
15 Nair, R. *et al.* The role of emotions in multiagent teamwork. In *Who Needs Emotions? The Brain Meets the Robot* (Fellous, J-M. and Arbib, M.A., eds), Oxford University Press (in press)
16 Jeannerod, M. How do we decipher other's mind. In *Who Needs Emotions? The Brain Meets the Robot* (Fellous, J-M. and Arbib, M.A., eds), Oxford University Press (in press)
17 Gallese, V. (2003) The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology* 36, 171–180
18 Ortony *et al.* Affect and proto-affect in effective functioning. In *Who Needs Emotions? The Brain Meets the Robot* (Fellous, J-M. and Arbib, M.A., eds), Oxford University Press (in press)
19 Sloman, A. (2002) How many separately evolved emotional beasties live within us? In *Emotions in Humans and Artifacts* (Trappl, R. *et al.*, eds), pp. 35–114, MIT Press
20 Murphy, R. *et al.* (2002) Emotion-based control of cooperating heterogeneous robots. *IEEE Transactions on Robotics and Automation* 18, 744–757
21 LeDoux, J. (1996) *The Emotional Brain*, Simon & Schuster
22 Damasio, A.R. (1994) *Descartes' Error: Emotion, Reason and the Human Brain*, Putnam
23 Dalgleish, T. (2004) The emotional brain. *Nat. Rev. Neurosci.* 5, 583–589
24 Swanson, L.W. (2000) Cerebral hemisphere regulation of motivated behavior. *Brain Res.* 886, 113–164
25 Striedter, G. (2004) *Principles of Brain Evolution*, Sinauer Associates
26 Butler, A.B. and Hodos, W. (1996) *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation*, John Wiley & Sons
27 Kelley, A.E. Neurochemical networks encoding emotion and motivation: an evolutionary perspective, In *Who Needs Emotions? The Brain Meets the Robot* (Fellous, J-M. and Arbib, M.A., eds), Oxford University Press (in press)
28 Holland, P.C. and Gallagher, M. (2004) Amygdala-frontal interactions and reward expectancy. *Curr. Opin. Neurobiol.* 14, 148–155
29 Miller, E.K. *et al.* (2002) The prefrontal cortex: categories, concepts and cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357, 1123–1136
30 Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202
31 Pare, D. and Quirk, G.J. (2004) New vistas on amygdala networks in conditioned fear. *J. Neurophysiol.* 92, 1–9
32 LeDoux, J.E. (2000) Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184
33 Rolls, E.T. (2000) Neurophysiology and functions of the primate amygdala, and the neural basis of emotion. In *The Amygdala: Second Edition. A Functional Analysis* (Aggleton, J.P., ed.), pp. 447–478, Oxford University Press, Oxford
34 Adolphs, R. (2003) Is the human amygdala specialized for processing social information? *Ann. N. Y. Acad. Sci.* 985, 326–340
35 Davidson, R.J. (2003) Darwin and the neural bases of emotion and affective style. *Ann. N. Y. Acad. Sci.* 1000, 316–336
36 Ochsner, K.N. (2004) Current directions in social cognitive neuroscience. *Curr. Opin. Neurobiol.* 14, 254–258
37 Schultz, W. (2000) Multiple reward signals in the brain. *Nat. Rev. Neurosci.* 1, 199–207
38 Dolcos, F. *et al.* (2004) Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* 42, 855–863
39 Dolan, R.J. (2002) Emotion, cognition, and behavior. *Science* 298, 1191–1194
40 Hamann, S. and Canli, T. (2004) Individual differences in emotion processing. *Curr. Opin. Neurobiol.* 14, 233–238
41 Taira, M. *et al.* (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp. Brain Res.* 83, 29–36
42 Rizzolatti, G. *et al.* (1988) Functional organization of inferior area 6 in the macaque monkey II. Area F5 and the control of distal movements. *Exp. Brain Res.* 71, 491–507
43 Rizzolatti, G. and Craighero, L. (2004) The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192
44 Rizzolatti, G. *et al.* (1995) Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3, 131–141
45 Rizzolatti, G. *et al.* (1996) Localization of grasp representations in humans by positron emission tomography: 1. Observation versus execution. *Exp. Brain Res.* 111, 246–252

46 Grafton, S.T. *et al*. (1996) Localization of grasp representations in humans by PET: 2. Observation compared with imagination. *Exp. Brain Res.* 112, 103–111

47 Rizzolatti, G. and Arbib, M.A. (1998) Language within our grasp. *Trends Neurosci.* 21, 188–194

48 Arbib, M.A. From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci*. (in press)

49 Adolphs, R. Could a robot have emotions? Theoretical perspectives from social cognitive neuroscience. In *Who Needs Emotions? The Brain Meets the Robot* (Fellous, J-M. and Arbib, M.A., eds), Oxford University Press (in press)

50 Hatfield, E. *et al*. (1993) Emotional contagion. *Current directions in psychological sciences* 2, 96–99

51 Heiser, M. *et al*. (2003) The essential role of Broca's area in imitation. *Eur. J. Neurosci.* 17, 1123–1128

52 Gallese, V. and Goldman, A. (1998) Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* 2, 493–501

53 Hasselmo, M.E. (1995) Neuromodulation and cortical function: modeling the physiological basis of behavior. *Behav. Brain Res.* 67, 1–27

54 Fellous, J.M. (1999) Neuromodulatory basis of emotion. *The Neuroscientist* 5, 283–294

55 Lane, R.D. (2000) Neural correlates of conscious emotional experience. In *Cognitive Neuroscience of Emotion* (Lane, R.D. and Nadel, L., eds), pp. 345–370, Oxford University Press

56 Fellous, J.M. (2004) From human emotions to robot emotions. In *Architectures for Modeling Emotion: Cross-Disciplinary Foundations* (Hudlicka, E., ed.), pp. 37–47, American Association for Artificial Intelligence

57 Doya, K. (2002) Metalearning and neuromodulation. *Neural Netw.* 15, 495–506

58 Avila-Garcia, O. and Cañamero, L. (2004) Using hormonal feedback to modulate action selection in a competitive scenario. In *From Animals to Animats: Proc. 8th Int. Conf. on Simulation of Adaptive Behavior (SAB'04)*, pp. 243–252, MIT Press

59 Ortony, A. *et al*. (1988) *The Cognitive Structure of Emotions*, Cambridge University Press, Cambridge

60 Pribram, K.H. (1960) A review of theory in physiological psychology. *Annu. Rev. Psychol.* 11, 1–40

61 Kilmer, W.L. *et al*. (1969) A model of the vertebrate central command system. *Int J. Man-Machine Studies* 1, 279–309

62 McGaugh, J.L. *et al*. (2000) Brain systems and the regulation of memory consolidation. In *Brain, Perception, Memory: Advances in Cognitive Neuroscience* (Bolhuis, J.J., ed.), pp. 233–251, Oxford University Press

63 Pessoa, L. and Ungerleider, L.G. (2004) Neuroimaging studies of attention and the processing of emotion-laden stimuli. *Prog. Brain Res.* 144, 171–182

64 Gasper, K. and Clore, G.L. (2002) Attending to the big picture: mood and global versus local processing of visual information. *Psychol. Sci.* 13, 34–40

65 Trappl, R. *et al*. (2002) *Emotions in Humans and Artifacts*, MIT Press

66 Fellous, J-M. and Arbib, M.A., eds *Who Needs Emotions? The Brain Meets the Robot*, Oxford University Press (in press)

67 Schultz, W. (2004) Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr. Opin. Neurobiol.* 14, 139–147

68 Fellous, J-M. and Suri, R. (2003) Roles of dopamine. In *The Handbook of Brain Theory and Neural Networks* (2nd edn) (Arbib, M.A., ed.), pp. 361–365, MIT Press

69 Abrams, J.K. *et al*. (2004) Anatomic and functional topography of the dorsal raphe nucleus. *Ann. N. Y. Acad. Sci* 1018, 46–57

70 Panksepp, J. (1998) *Affective Neuroscience*, Oxford University Press