Adaptive Behavior

http://adb.sagepub.com/

Robots that *have* emotions

Domenico Parisi and Giancarlo Petrosino

Adaptive Behavior published online 15 November 2010
DOI: 10.1177/1059712310388528

Original Paper

Adaptive Behavior
0(00) 1–17
© The Author(s) 2010
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1059712310388528
adb.sagepub.com

Robots that *have* emotions

Domenico Parisi and Giancarlo Petrosino

Abstract

Animals have many different motivations and they must decide which of their different motivations they should try at any given time to satisfy with their behavior. Emotions are states of an individual's body and brain that allow the motivational decision mechanism of the individual to function more properly, that is, in ways that increase the individual's survival and reproductive chances. The article describes five different simulated robots which, unlike current "emotional" robots, can be said to *have* emotions. The robots may need to eat and drink, eat and fly away from a predator, eat and find a mating partner, eat and take care of their offspring, or eat and rest in order to heal from physical damage. We show that adding a special *emotional circuit* to the neural network controlling the robots' behavior leads to better motivational decisions, and therefore to higher fitness, and we describe how in many circumstances the robots endowed with an emotional circuit behave differently than those lacking the circuit. We conclude by indicating various directions of future research.

Keywords

Emotion, motivation, robots

1 Introduction

Current "emotional" robots can *express* emotions or can *recognize* our emotional expressions but they cannot be said to *have* emotions because emotions do not play any functional role in their behavior. In this article we advance an hypothesis about emotions by describing robots that *have* emotions because emotions can be shown to play a well-identified function in what they do. Our hypothesis has two parts: (1) to construct robots that have emotions it is first necessary to construct robots that have many different motivations that compete with one another for the control of the robot's behavior so that the robot has to decide which motivation should control its behavior at any given time; (2) the robot will have emotions if the neural network that controls the robot's behavior includes a special *emotional circuit* which allows the robot to take faster and more correct motivational decisions that allow the robot to live longer and generate more offspring.

In Section 2 we introduce the distinction between a strategic or motivational level and a tactical or cognitive level of functioning of an organism. In Section 3, we describe our hypothesis about emotions as states of an organism that allow the organism's brain to make more correct, faster, and generally better motivational decisions. In Section 4 we discuss current work on "emotional" robots and we explain why these robots cannot be said to have emotions. In Section 5 we describe five different robots that incorporate the two levels of functioning, motivational and cognitive, in their "brain" (neural network) and in Section 6 we add to the robots' neural network a special emotional circuit that we show leads to an increase in the survival/reproductive chances of the robots. In Section 7 we discuss various aspects of the behavior of our emotional robots and, finally, in Section 8 we summarize the results and indicate some directions of future research.

2 The motivational level and cognitive level of behavior

Behavior tends to be explained as caused by stimuli. The brain processes information provided by the senses and tells the effectors what to do. In cognitive

---

Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy.

Corresponding author:
Giancarlo Petrosino, Institute of Cognitive Sciences and Technologies, National Research Council, 44 Via S. Martino della Battaglia, 00185 Rome, Italy
Email: giancarlo.petrosino@istc.cnr.it

models the processing may be very complex and may be based on *internal models* but stimuli still appear to be the main determinants of behavior.

This picture of behavior is incomplete. Imagine an animal that is perceiving a piece of food. The animal will respond to the sensory stimuli from food by approaching and eating the food only if the animal is hungry. If it is not hungry, the animal will simply ignore food. This indicates that sensory stimuli from the environment are not the only causes of behavior. An explanation of the animal's behavior requires that we take into consideration the motivation which currently controls the behavior. It is the currently active motivation, together with the sensory stimuli, that makes it possible to predict what the animal will do. Stimuli are not enough.

Animals have many different motivations. To survive and reproduce even simple animals need to eat, drink, escape from predators, and find mating partners. However, their behavior can generally be controlled by only one motivation at a time. For example, to survive an animal may need to both eat and drink but, if food and water are located in different places, at any given time the animal must decide whether to look for food or water. This requires that animals possess a mechanism for deciding which one of their different motivations should control their behavior at any given time. (We use the verb ''to decide'' but motivational decisions are the result of low-level processes and, except in some cases in humans, do not require higher order cognitive abilities such as making explicit predictions and evaluations.) We call this level of functioning of the animal the strategic or motivational level. Once one particular motivation has been chosen at the strategic level, the animal will execute the behavior that satisfies the motivation. We call this second level of functioning the tactical or cognitive level. If an animal has to survive and reproduce both levels must function properly. If an individual takes the appropriate decision at the strategic level but is unable to execute the appropriate behavior that will allow the individual to satisfy the motivation chosen at the strategic level, the probability that the individual will survive and reproduce will be reduced. Conversely, if an individual is good at executing the behavior that makes it possible to satisfy the motivation chosen at the strategic level but not so good at choosing the appropriate motivation to be pursued at that particular time, the individual's fitness will also be decreased. The ability to map stimuli into the appropriate behaviors constitutes the cognitive level of behavior but we need to postulate another level of functioning of the animal, the motivational level.

We assume a simple, implicit, mechanism for deciding which particular motivation will control the robot's behavior at any given time: all the robot's motivations have a quantitative level of intensity and the motivation which wins the competition is the motivation which currently has the highest level of intensity (Parisi, 2004; Ruini, Petrosino, Saglimbeni, & Parisi, 2010). The current level of intensity of a motivation may depend on a number of factors such as the particular environment in which the robot has evolved, the intrinsic importance of the motivation for the robot's overall adaptive pattern, the current state of the robot's body, and the current stimuli from the environment.

The idea that motivation in animals (and robots) should be studied starting from the competition among different motivations is not usual in the scientific literature on motivation which prefers to study single motivations and dedicates little attention to the existence in animals of many different motivations that cannot all be satisfied at the same time and therefore compete for the control of the animal's behavior (see, e.g., Colgan, 1989; for a detailed empirical investigation of one single motivation, foraging, see Altmann, 1998). *Motivational systems* by Toates (1986), which is one of the few books which has a (short) chapter dedicated to the interaction among different motivations, recognizes this limitation and attributes it to the great complexity of each individual motivation and to the fact that in the experimental laboratory it is easier to study single motivations rather than multiple competing motivations. This is certainly true but the fact remains that until we are able to analyze and explain competition among different motivations we cannot be said to have really understood motivation.

We have distinguished two levels of functioning in animals, the strategic or motivational level and the tactical or cognitive level. (We use ''level'' in a purely functional sense. As we will see, nothing corresponds to the two levels in the neural network which controls the behavior of our robots and which has only a single layer of internal units.) The motivational level might be considered as more adaptively important than the tactical level for at least three reasons. One reason is that functioning badly at the motivational level may more directly and more seriously imperil the animal's survival and reproductive chances than functioning badly at the cognitive level. If an animal is hungry but it pursues another motivation rather than the motivation to eat, the animal will quickly and inexorably die. On the other hand, not being very good at the cognitive level is generally less critical. If an animal is hungry and it correctly looks for food, the animal may not be particularly good at finding food but it will probably survive, although not well. A second reason why the motivational level can be considered as more critical for an animal's survival than the cognitive level is that while the cognitive level of functioning may be improved through learning this is less true for the motivational level, although of course new motivations that are added to existing motivations can be learned. An animal can learn to produce better behaviors to satisfy its motivations but its ability to decide more effectively which motivation to pursue at any given time is more difficult to improve. And, finally, learning depends on receiving rewards or punishments, and it is the motivational level that provides rewards and punishments.

Motivational decisions should not be confused with action selection (Brooks, 1986; Canamero, 2003, 2005; Maes, 1990) if action selection is choosing the behavior which is most appropriate to satisfy the animal's current motivation. Motivational decisions are super-ordinate with respect to action selection and, in fact, motivational decisions belong to the strategic level of behavior while action selection belongs to the tactical level. (Anil Seth's notion of action selection is closer to our notion of motivational decision, as indicated by the simulations he describes; cf. Seth, 2007.)

3 Emotions

As we have said, deciding correctly which motivation should control the animal's behavior at any given time is critical for an animal's survival and reproductive success. Our hypothesis about emotions is that emotions (or emotional states; we use the two expressions as synonymous) are *(felt) states of an animal's body/brain that increase the correctness and effectiveness of the motivational decisions of the animal by influencing the current intensity of the different motivations*. The motivational level may not function well and, as we have said, its "mistakes" may be critical for the animal's survival. If an animal has to choose among a limited number of different motivations, or if the factors that have to be taken into consideration to make the appropriate decision are restricted in number, or if the decision process is such that errors and slowness do not compromise the individual's survival and reproductive chances, motivational decisions may not need emotions. Emotions are part of the strategic level at which alternative motivations compete for the control of the animal's behavior but motivational decisions need not be necessarily accompanied by emotional states. An animal can shift from looking for food to looking for water with no particular emotional state accompanying the decision to shift. Emotions tend to emerge in more complex animals that have to choose among a large number of different motivations, or have to take many different factors into consideration in order to decide correctly, or have to decide quickly, or have to either persist in pursuing a motivation which they find difficult to satisfy or abandon a motivation they cannot satisfy. These animals can make errors in their motivational decisions or they can be slow in deciding, and this may compromise their survival and reproductive chances. Our hypothesis is that emotional states are an evolved mechanism for making the strategic level of behavior more effective, less subject to errors, and faster. Using a metaphor we might say that emotional states allow a motivation to "speak louder" than other motivations in order to win the competition with other motivations. What we will try to do in this article is try to go beyond the metaphor and outline a more operational (robotic) model of emotions.

Our hypothesis about emotions is related to the work of Niko Frijda (Frjida, 1987; see also Canamero, 2005). However, Frijda believes that emotions arise from the value of stimuli as rewards and punishments while we think that the motivational level of behavior confers reward and punishment value to stimuli and that emotions are a specific submechanism of the motivational level which only enters in action to make motivational decisions more correct and faster. An animal can find an experience rewarding without associating any emotional state to the experience.

The terms *motivation* and *emotion* tend to be confused together. For example, Edmund Rolls' important book *The Brain and Emotion* (Rolls, 1999), contrary to its title, is a book about motivations, not emotions. Although emotions belong to the motivational level of behavior they are a specialized submechanism of the motivational level and, as we have said, motivations can be studied without studying emotions. That the motivational mechanism operates at a very basic level is indicated by the fact that even a very simple organism such as the small worm *Caenorhabditis elegans*, with only about 300 neurons and a total of about 1000 cells in its body, has to decide whether to look for food or to bend its body when a stimulus touches its body. But it is not clear that one can attribute emotions to such a simple creature. Our robots are not much more complex than *C. elegans* and in fact they can survive and reproduce even without an emotional circuit in their brain. However, the results of our research show that if an emotional circuit is added to the neural network that controls their behavior they reach higher levels of fitness. Therefore, we suppose that, given the appropriate preconditions and pre-adaptations, they would spontaneously evolve such a circuit.

Although motivation and emotion are distinct phenomena, they both belong to the strategic level of behavior, and an account of emotions and their role in behavior is impossible if it does not include motivations and competition among different motivations. Contrary to this, another important book, Jaak Panksepp's *Affective Neuroscience* (Panksepp, 1998), offers a treatment of the neural bases of emotions with no reference to motivation or to competition among different motivations; and a recent book edited by Peter Ellison and Peter Gray (Ellison & Gray, 2009) contains a very useful treatment of the hormonal bases of social motivations and emotions without mentioning the terms ''motivation'' and ''emotion'' in its index of words.

4 Robots that *have* emotions

Can we construct robots that have emotions, that is, robots for which emotions play a clearly identifiable functional role in their behavior?

There is much work devoted to robots and emotions but most of this work is aimed at constructing robots that *express* emotions but cannot be said to *have* emotions, or to robots that can understand our expression of emotions as a purely perceptual task, without sharing our emotions (Adolphs 2005; Canamero, 2005; Dautenhahn et al., 2009; Picard, 2000, 2003). (For some attempts at understanding the functional role of emotions in behavior, see Avila-Garcia & Canamero, 2004; Ziemke, 2008; Ziemke & Lowe, 2009.) The expression of emotions makes the emotional states of an individual accessible to other individuals and it is an important phenomenon and an important topic of research. However, in this article we are not interested in the expression of emotions but in what appears to be the most basic and general function of emotional states which is to make the organism's motivational decision mechanism function more effectively.

The reason why current ''emotional'' robots (Arbib & Fellous, 2004; Fellous & Arbib, 2005) are robots that can express emotions but cannot be said to have emotions, is that current robots do not have motivations and therefore do not have to decide among different motivations to select the specific motivation that must control their behavior at any particular time. As we have said, emotions are a submechanism that exists in order to help the motivational decision mechanism to function more properly. Therefore, if current robots do not have motivations, they cannot have emotions. Current robots possess only a tactical or cognitive level of functioning (they respond to stimuli with responses) but they do not have a strategic or motivational level. The reason why current robots do not have a motivational level of functioning is that the ''motivation'' that has to control their behavior is decided by us, their users. (In this sense ''autonomous robots'' appears to be a misnomer. For an attempt at constructing motivationally autonomous robots controlled by neural networks see Lowe et al., 2010. Dörner (2001) describes an artificial system controlled by a variety of symbolic and neural systems which has many different motivations and whose behavior is modulated by emotional states.) The reason why the motivations of robots are not decided by the robots themselves but by us is that most of the work done in robotics is oriented toward practical applications, and for practical applications it does not make much sense, and can be counter-practical or even dangerous, that robots autonomously decide which motivation to pursue and have emotional states that we do not control.

Another difference between our robots that have emotions and most work done on emotional robots is that the behavior of our robots is controlled by an artificial neural network, which is a (very) simplified model of the brain, while current emotional robots tend to be controlled by symbolic or rule-based systems such as those used in the research described in Breazeal and Brooks (2005). This also applies to the few robots that have been constructed which can be said to have a motivational decision mechanism (Brooks, 1986; Maes, 1990). These systems assume the existence of different behaviors such as eating, drinking, avoiding predators, and so forth, and they include a set of rules for deciding which behavior should be executed by the robot at any given time. Dörner's (2001) PSI model has a number of boxes each containing a different motivation, and the model describes how each box gains control of the system's behavior at any given time. Our approach is different. We use the expressions ''motivations'' and ''motivational decisions'' but these are only useful, descriptive terms. A neural network is made up only of units (neurons) and connections between units (synapses between neurons) and the only thing that takes place in a neural network is that activation patterns cause other activation patterns. Therefore, literally there is nothing like a ''motivation'' or a ''behavior'' or a ''decision mechanism'' in a neural network. The difference between rule-based approaches and our approach is that behaviors and motivational decisions in rule-based robotics tend to be explicitly represented while they are only implicit in neural network approaches. (The distinction is discussed in Seth, 1998, 2007, which describe robots that, like our robots, take motivational decisions and are controlled by neural networks.)

In the next two Sections we describe some robots that have two different motivations which compete with one another for the control of the robot's behavior and we contrast two types of robots: robots whose neural network includes an emotional circuit and robots whose neural network lacks the emotional circuit.

5 Robots that have to take motivational decisions

Our robots are simulated Khepera robots. Khepera robots have a cylindrical body of 55 mm diameter and

30 mm height, light, infrared, and ground sensors, and two DC motors with incremental encoder controlling the two wheels that make the Khepera robot able to move. Our robots are simulated in a computer and we have used Evorobot, a simulation tool developed for the Khepera robot (Nolfi & Gigliotta, 2010). We have run five different simulations. We specify here what is common to all five simulations while we will add specific details when we will describe the different simulations.

The environment in which the robots live is a grid of pixels with a size which varies in the different simulations. The robot's body has a diameter of 75 pixels. The environment is circumscribed by a wall but, except for our robots that have to eat and drink, we have not used the Khepera's infrared sensors and when the robot happens to touch the wall its body is rotated in a randomly chosen new direction. The only external sensors of our robots are visual sensors. Khepera has eight light sensors but we have used only four of these sensors, two on the left and two on the right. In some of our robots each of these sensors is replicated twice to allow the robot to see two differently colored objects that can appear in different positions in its visual field. For one of our robots we have used not Khepera's visual sensors but its ground sensors (see below).

The robots' behavior is controlled by a neural network with two sets of input units. One set of input units encodes the perceptual properties of the different objects which are present in the robot's environment and another set of input units encodes different states of the robots' body (hunger, thirst, pain). The number of the input units varies in the different simulations and will be specified later. Both sets of input units are connected to an intermediate set of four internal units which in turn send their connections to two motor units encoding the speed of Khepera's two wheels. The neural network is a standard network with activation level between 0 and 1 for all the network's units and a sigmoid activation function for the internal units and the motor units. Each internal unit and each motor unit has a "bias" that defines a spontaneous level of activation of the unit which is added to the activation arriving from other units to determine the effective level of activation of the unit at any given time. The bias is always operative, which means that the internal units will be activated even when no input is arriving from the input units and will always influence the robot's motor units. The activation level of the motor units specifies the speed of the robot's two wheels and therefore the robot's displacements in the environment.

To develop the neural network's connection weights we have used a genetic algorithm (Mitchell, 1998), with a population of 100 robots that reproduce selectively based on their individual fitness and with random changes in the inherited genotypes (Nolfi & Floreano, 2000). Each individual robot lives alone in its own copy of the environment. The 20 individuals of each generation which have the highest fitness generate five offspring each and the $20 \times 5 = 100$ offspring constitute the next generation. All simulations last for 1000 generations. The criterion for ranking the 100 individuals, that is, what constitutes fitness, varies with the simulation and will be described below.

The robots' inherited genotype encodes the connection weights of the neural network that controls their behavior and the biases of the network's units. At the beginning of the simulation the robots' neural networks have randomly assigned connection weights and biases within the range -5.0 and $+5.0$ and, if a robot reproduces, its offspring inherit the connection weights of their (single) parent with random mutations in the range $-5.0/+5.0$ for both the connection weights and the biases. The genotype is a sequence of bits and each bit has a probability of mutation of 2%, except for the robots that have to eat and drink for which the probability of mutation is 4%. (For more details on the simulations, see Petrosino, Da Rol, Zotti, & Parisi, 2010; Ruini et al., 2010; and Saglimbeni & Parisi, in press.)

To survive and reproduce all our robots have to satisfy two different motivations. They can be robots that have to look for food *and* for water, robots that have to look for food *and* try to escape from a predator, robots that have to look for food *and* for a mating partner, robots that have to look for food *and* take care of their offspring, or robots that have to look for food *and* rest when their body incurs some physical damage and has to heal from the damage. Since a robot cannot pursue both motivations at the same time, at any given time the robot has to choose between its two different motivations and produce the behavior that will make it possible to satisfy the motivation which has been chosen.

Our first type of robot needs both energy and water to survive. The environment (1000×1000 pixels) contains food tokens (providing energy) and water tokens (both with a diameter of 30 pixels) and the two types of tokens have different colors so that the robots can distinguish between them. The robots' neural network has four visual input units, with one unit encoding the presence of food tokens and one unit encoding the presence of water tokens on the left side of the robot's visual field, and two other units for the right side. In addition there are two body input units, one encoding the current level of energy (hunger) and the other one the current level of water (thirst) in the robot's body. The robots' environment is a seasonal one, with seasons with more abundant food (5 tokens) and little water (1 token) followed by seasons with little food (1 token) and abundant water (5 tokens). When the centre of the robot's body happens to be inside a food or water token, the token disappears (is eaten or drunk) and a new token of the same type appears in a randomly chosen location in the environment, while the energy or water in the robot's body is increased by some fixed quantity. The robot's body includes two internal stores having the same size, one for energy and the other one for water, and the current level of energy and water in the body is communicated to the robot's brain through the two body input units, one for hunger and the other one for thirst. These units have an activation level of 1 when the corresponding store is completely full and a level of 0 when it is empty. A robot's life lasts for 10 epochs, each of a maximum of 1500 time steps, with one epoch with more food than water followed by an epoch with more water than food. At each time step (input/output cycle of the robot's neural network) the same quantity of energy and water is consumed to keep the robot alive, and if either one of the two stores reaches the zero level, the robot dies. The robot periodically generates one offspring. Hence, the robot's fitness is equivalent to the length of its life. The behavior that evolves in the robots can be described as "Look for food if you feel more hungry than thirsty and look for water if you feel more thirsty than hungry."

In a variant of the simulation the robots live in an environment in which food is always more abundant than water (5 food tokens and 1 water token). In this environment the motivation to drink is intrinsically stronger than the motivation to eat and this is reflected in the robots' behavior: the robots always tend to go toward water rather than food, unless food is very close or they are very hungry. (For a detailed description of the two simulations, see Saglimbeni & Parisi, 2009, in press).

For the robots that live in the seasonal environment, the information that allows a robot to decide whether to look for food or water is conveyed by the internal stimuli that arrive to the robot's brain from within the robot's body (hunger and thirst). However, the sensory input that originates in the external environment (seeing food or water) can also play a role in deciding the current intensity level of the two motivations and in this way influence the decisions taken at the strategic level. For example, if a food token is very near the robot may go and eat the food token even if it is more thirsty than hungry.

A better example of the role of external stimuli in deciding which motivation wins the competition with other motivations is our next robot which to survive has to both eat food and avoid being killed by a predator. The environment is 1500 × 1500 pixels in size and it contains 24 food patches of 70 pixel diameter. When the centre of the robot's circular body is inside a food patch, the robot eats some of the food in the patch, the size of the patch's diameter is decreased by 2 pixels, and the patch disappears when its diameter is less than 10 pixels. A hardwired predator appears at random intervals in a random location in the environment; for a randomly decided number of time steps before disappearing it chases the robot, and if it reaches the robot, the robot dies. (The predator's speed is equivalent to the maximum speed of the robot.) The robot has a neural network with two input units encoding the location of the nearest food patches, two input units encoding the presence and location of the predator when the predator appears, and one input unit encoding the level of energy in the robot's body (hunger). When the predator is absent, the motivation to eat controls the robot's behavior and the robot looks for food because if the energy store in its body reaches the zero level, the robot dies. But when the predator appears and is perceived by the robot, the motivation to avoid being eaten by the predator wins the competition with the motivation to eat: the robot ignores food and tries to escape from the predator. A robot's life is made up of five epochs of maximum 2000 time steps and in this simulation the fitness criterion is the quantity of food eaten during life.

Like the preceding simulation in which food was always more abundant than water, and therefore the motivation to drink was intrinsically more important that the motivation to eat, this is another example showing that the intrinsic importance of different motivations is a factor in deciding which motivation will control the robot's behavior. Finding food is important but flying away from the predator is more important. However, the problem may not be one of wrongness versus correctness in the decision taken but of time: the motivational decision may be the correct one but the strategic level takes too much time to decide. When the predator is absent, the robot should approach and eat the food. But when the predator appears, the robot should not only switch from looking for food to flying away from the predator but it should make the switch as rapidly as possible. The results show that at the end of the simulation most robots are able to do so. (For a more detailed description of these simulations, see Petrosino et al., 2010.)

All the robots we have described so far reproduce nonsexually. A robot generates an offspring with no need for a mating partner. Our next robots reproduce sexually, that is, they need a mating partner to reproduce. They need to eat in order to remain alive because if they die they cannot reproduce. But remaining alive is not enough. To transmit its genes to the next generation, a robot must find mating partners. In fact, the fitness of these robots is the number of mating events, where a mating event occurs when a robot touches a mating partner.

The robot's life lasts for three epochs each with a maximum number of 3000 time steps. The environment is 1000 × 1000 pixels and it contains four food patches with a diameter of 40 pixels and a single mating partner. Mating partners are not robots themselves but they are nonmoving objects with a color that distinguishes them from food tokens. When the robot reaches the mating partner, a mating event occurs, and after the mating event the mating partner disappears, and another potential mating partner appears in a new, random, location in the environment. To leave its genes to the next generation the robot must both reach and eat the food tokens and reach and mate with mating partners. If the robot always looks for mating partners and ignores food, the robot will die and cannot reproduce. If the robot always looks for food but ignores mating partners, the robot will live a long life but it will not leave its genes to the next generation.

The robot's neural network has two input units encoding the location of food tokens, two units encoding the location of the mating partner, and one body input unit encoding the level of energy in the robot's body, that is, hunger. By using this information, at the end of the simulation most robots are able to choose appropriately at any given time between the motivation to eat and the motivation to find a mating partner.

For many animal species generating offspring is not a sufficient strategy to leave one's genes to the next generation. In the early stages of their life the offspring may not be able to take care of themselves and therefore their parents have to take care of them if the offspring have to survive so that their parents' genes can be transmitted to future generations. Our fourth type of robots live for four epochs of a maximum of 2000 time steps in an environment (500 × 500 pixels) containing two circumscribed zones, a food zone and an offspring-care zone, both of 60 pixel diameters. For each time step spent in the food zone the energy contained in the robot's body is increased by some fixed quantity. For each time step spent in the offspring-care zone, the survival chances of the offspring already generated increase by some fixed quantity. Like our robots living in the environment with food and water and in the environment with food and predators, these robots reproduce nonsexually and they generate one offspring at regular intervals. However, unlike the preceding robots, their fitness is not the length of their life or the quantity of food eaten or the number of mating episodes but the time spent in the offspring-care zone. A robot has to both eat (time spent in the food zone) and take care of its offspring (time spent in the offspring-care zone) and it must find the appropriate balance between these two motivations. The robot's neural network includes, in addition to the usual body internal unit for hunger, two local (not distance, as in the preceding simulations) sensors located under the robot's body (Khepera's ground sensors) and informing the robot's brain when the robot is inside the food zone or inside the offspring-care zone. In other words, these robots do not perceive the food zone or the offspring-case zone from a distance but only when they happen to be inside one of the two zones. They have to explore the environment without external sensory input when they are outside the two zones and to react appropriately when they enter the two zones. The results of the simulations show that the robots evolve the capacity to decide appropriately between staying in the food zone in order to eat and remain alive and leaving the food zone and reaching the offspring-care zone in order to take care of their offspring and insure their survival.

Our final type of robots live in an environment of 1000 × 1000 pixels with four food patches of 40 pixel diameter which they have to eat in order to survive and they reproduce nonsexually at regular intervals. However, these robots have an additional problem. At irregular intervals their body may incur some physical damage which can vary in severity and which lasts for a certain (random) number of time steps. When this happens, the robots should rest or at least reduce their speed in order for the physical damage to heal (for resting as an appropriate strategy in response to physical damage, see Wall, 2002). If they do not do so, their fitness is correspondingly reduced because of continuing physical damage (for a detailed description of this simulation, see Acerbi & Parisi, 2007). Their neural network is informed of the physical damage and of its severity by a continuously activated pain sensor (with an activation of 1 for maximum physical damage/maximum pain and an activation of 0 for zero physical damage/no pain) and, like our preceding robots, of the level of energy in their body by a hunger sensor. The reduction in fitness due to physical damage is proportional to the severity of the physical damage and to how the robot responds to pain by reducing its speed of movement. The robot's life lasts for five epochs, each of a maximum of 2000 time steps.

These robots too have to take motivational decisions. When they feel pain, they have to decide whether to look for food or to rest or at least reduce their speed. Of course they cannot do both things at the same time. If they decide to look for food, they have to move in their environment and to move as quickly as possible, ignoring pain. If they decide to respond to pain they have to rest and cannot look for food. At the end of the simulation the robots appear to be able to take the appropriate motivational decisions.

6 Adding an emotional circuit to the robots' neural network

The robots we have described in the preceding section can be said to have both a strategic/motivational level and a cognitive level of functioning because at any given time they have to decide which one of their two different motivations to pursue and they have to produce the behavior that allows them to satisfy the chosen motivation. The best robots tend to be good at doing both things, less good robots may be good at one but not very good at the other, and the worst robots are good at neither. However, these robots cannot be said to have emotions. There is nothing in the robots that can be described as an emotion or an emotional state and there is no proof of the functional role of emotions or emotional states in the robots' behavior. In this section we describe robots that can be said to have emotional states and we show how these emotional states increase the fitness of the robots.

For each of the five types of robots we have described in the preceding section we compare two different populations of robots. In one population the neural network of the robots has the architecture we have described in the preceding section: input units encoding information from the external environment (food, water, predator, mating partner, offspring-care zone) and from inside the body (hunger, thirst, pain), internal units, and motor output units (Figure 1a). In the other population we add an emotional circuit to the robots' neural network. The emotional circuit is made up of one or two emotional units to which some of the input units send their activation and which in turn send their activation either to the internal units (Figure 1b) or directly to the motor units (Figure 1c). (We will discuss the results obtained with these two different emotional circuits later on.) The robots that have to both eat and drink have two emotional units to which the hunger and thirst sensors send their activations, while all the other robots have a single emotional unit. The emotional unit receives activation from an appropriate source: from the predator's sensors for robots that have to escape from the predator; from the mating partner's sensors for the robots that have to find a mating partner; from the offspring-care zone sensors for the robots that have to take care of their offspring; and from the pain sensor for the robots that have to stop moving when they feel pain.

The emotional units have a number of properties that distinguish them from the "standard" internal units. First, they have no bias, which means that when no activation arrives to the unit from the input units, the unit is not activated and its does not influence the robot's behavior. Second, they have an activation threshold, which implies that if not enough activation arrives to an emotional unit from the input units, the unit remains inactive. Third, unlike the standard internal units, the emotional units are not only active in the cycle in which activation arrives from the input units but their activation persists in subsequent cycles. This temporally extended activation is controlled by three parameters: (1) a parameter specifying how the activation of the unit increases in a succession of cycles; (2) a parameter specifying the maximum possible activation; (3) a parameter specifying how the unit's activation decreases after reaching its maximum level. Figure 2 shows the formulas that define the "increase" and "decrease" phases of the emotional unit's activation and the temporal shape of the activation curve.

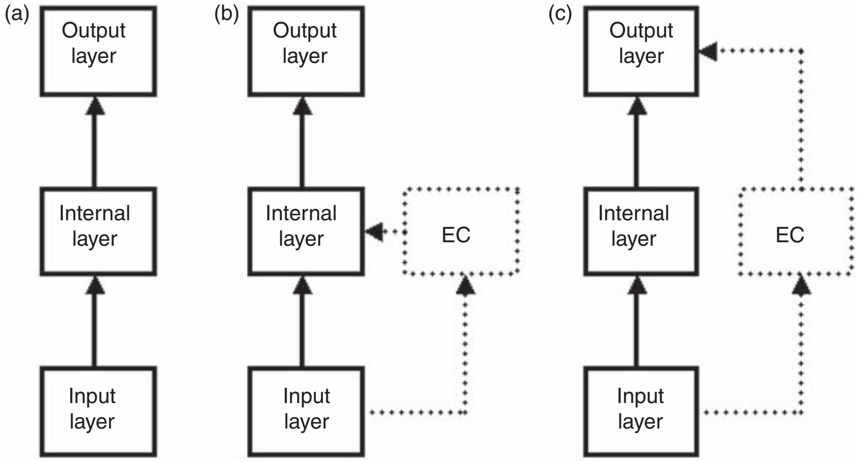


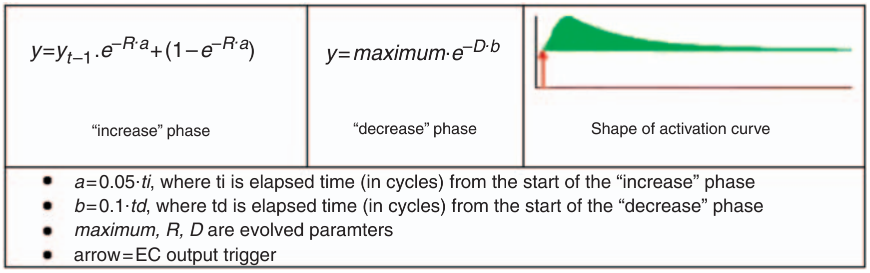
Figure 1. Neural network architecture for robots without an emotional circuit (a) and for robots with an emotional circuit linked to either the internal units (b) or the motor units (c).

The unit's activation threshold and the three parameters are encoded in the robot's inherited genotype together with the connection weights of the robot's neural network, and therefore they all have evolved values.

If we compare the fitness of the two types of robots at the end of the simulation we find that the robots with an emotional circuit have higher fitness than the robots without the emotional circuit. The robots are able to take the appropriate motivational decisions and to execute the behavior which allows them to satisfy the motivation which has been decided even without the emotional circuit. However, possessing an emotional circuit leads to more effective behavior and higher level of fitness, and this is true for all five types of robots: those that have to choose between looking for food and looking for water, those that have to choose between looking for food and escaping from the predator, those that have to decide between looking for food and looking for a mating partner, those that have to decide between looking for food and taking care of their offspring, and those that have to decide between looking for food and resting to heal from physical damage (Figure 3). The robots endowed with an emotional circuit can be said to *have* emotions or emotional states. Their emotional states are the activation states of their emotional circuit and these emotional states have a functional (beneficial) role in their behavior.

One might object that adding an emotional circuit is only adding more computational power to the robots' neural network and this explains why the robots with the emotional circuit have a higher fitness. To answer this objection we have compared the performance of the robots with the emotional circuit with the performance of robots in which we have simply added one or two (according to the robots) internal units to their neural network. What we have found is that the robots with a greater number of internal units do not behave any better than the robots with the basic neural network described in the preceding section. Hence the emotional circuit has a beneficial effect on the robot's behavior not because it simply increases the computing power of the robots' neural network but because of its special properties (described above).

As we have said, for each of the five types of robots endowed with an emotional circuit we have compared two different neural architectures. In one architecture the emotional units send their activation to the internal units while in the other architecture they send their activation directly to the motor units (cf. Figure 1b and 1c). Which architecture provides the best results? The answer to this question depends on the different robots. The architecture with the emotional circuit directly connected to the motor output units turns out to be better for the food/predator, food/mating partner, and food/offspring-care zone robots, while the opposite is true for the food/water and food/pain robots. In other words, the emotional circuit directly connected to the motor units gives better results when the emotional circuit is activated by external (environmental) input: perceiving the predator, perceiving a mating partner, or perceiving the offspring-care zone. On the other hand, the emotional circuit connected to the internal units, and therefore only indirectly to the motor units, gives better results when the emotional circuit is activated by internal input from the body: hunger, thirst, or pain. Why? We do not have a clear answer to this question but we note two differences between external and internal input to the brain. The first difference is that external input to the brain, that is, sensory input which originates in the environment outside the organism's body, is to some extent under control of the organism in that the organism can change this sensory input by moving its body while this is less true for internal input, that is, for sensory input which originates inside the organism's body (Parisi, 2004). The second difference is that sensory input from the external environment tends to be richer in (cognitive) information than sensory input from the internal environment. Perceiving the predator, the mate, or the offspring-care zone provides the organism's brain not only with motivationally relevant information but also with cognitive information (for example, where is the predator, the mate, or the offspring-care zone). In contrast,



$$y=y_{t-1}\cdot e^{-R\cdot a}+(1-e^{-R\cdot a})$$

$$y=maximum\cdot e^{-D\cdot b}$$

- $a=0.05\cdot ti$, where ti is elapsed time (in cycles) from the start of the "increase" phase
- $b=0.1\cdot td$, where td is elapsed time (in cycles) from the start of the "decrease" phase
- *maximum, R, D* are evolved paramters
- arrow = EC output trigger

Figure 2. The two formulas that define the "increase" and "decrease" phases of the emotional unit's activation and the temporal shape of the activation curve.

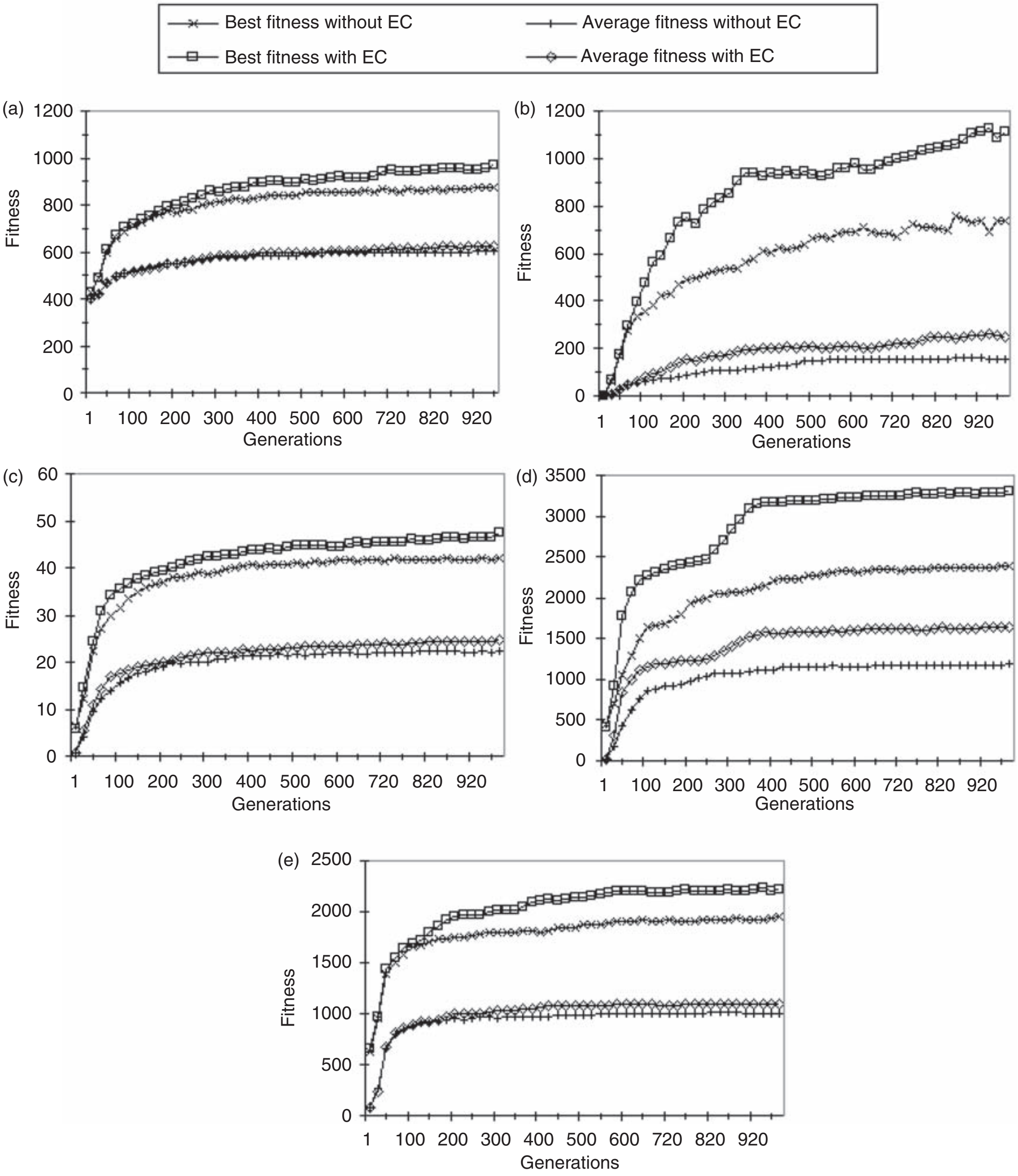


Figure 3. Average and best fitness for our five different robots with and without emotional circuit: (a) Food and water; (b) Food and predator; (c) Food and mating partner; (d) Food and offspring care; (e) Food and pain.

sensory input from within the body is generally poorer in cognitive information and it contains prevalently motivationally relevant information (how strong is the hunger, thirst, or pain). It might be that, in the case of the more cognitively rich external input, the emotional circuit can directly go to the motor units without passing through the cognitive part of the brain (the internal units), while when it is the cognitively poorer body input which activates the emotional circuit (feeling hunger, thirst, or pain) the emotional circuit must first pass through the cognitive part of the brain so that the motor units (motor response) can be more cognitively informed.

7 Analyzing our emotional robots

7.1 Looking at the robots' behavior

Let us now look more closely at the behavior of the robots that possess an emotional circuit. We will only discuss the robots that have to eat and escape from the predator, both because of space limitations and because these robots offer a particularly clear demonstration of the role of emotional states in behavior.

As we have seen, the robots with the emotional circuit have more fitness than the robots that do not have the emotional circuit, where fitness depends on two factors: the robot's ability to find food and its ability to escape from the predator when the predator appears. These are both cognitive or tactical abilities. But what is critical from the point of view of a robot's fitness is the robot's ability to shift appropriately from one activity to the other, which is a motivational or strategic ability. In fact we find evidence for differences in all three abilities (two cognitive abilities, approaching food and flying away from the predator, and one motivational ability, shifting from searching for food to flying away from the predator) between the two populations of robots. If we test individual robots in an ''experimental laboratory'' which either contains only food but no predator or only a predator but no food, we find that not only in their ecological (evolutionary) environment but also in these two artificial (laboratory) conditions, the robots with the emotional circuit are better than the robots without the emotional circuit. They eat more food than the robots without the emotional circuit when food is their only problem (1568 vs 845 food units; average of 10 best individuals of last generation) and they are better able to escape from the predator when their only problem is the predator. This seems to indicate that the presence of an emotional circuit not only helps the brain to take more correct motivational decisions but it leads to a better division of labor between the cognitive and the motivational/emotional parts of the brain, with advantages also for the cognitive part. Furthermore, when we test the robots either in a completely empty environment or in an environment which contains only the predator, we find that the robots with the emotional circuit have lower speed when the predator is absent than when the predator is present, while the robots without emotional circuit have the same high speed (Figure 4). This seems to imply that the robots with the emotional circuit look more carefully for food when the predator is absent.

What is also interesting is how in their ecological environment the two types of robots react when the predator appears. When the predator appears the robots *without* the emotional circuit do not react to the predator for a while but continue to approach and eat food. Only when the predator approaches do they cease to look for food and fly away from the predator. The behavior of the robots endowed *with* the emotional circuit is different. As soon as they perceive the predator these robots immediately cease looking for food and run in a randomly chosen direction. Only when the predator approaches them, they run in a specific direction, opposite to the direction from which the predator is approaching. This indicates that the robots with the emotional circuit are able to shift from one activity (looking for food) to another activity (reacting to the predator) more quickly than the robots without the emotional circuit. Their immediate reaction to the predator is not ''rational'' but is very effective.

We find support for this analysis if we examine what happens inside the neural network of the robots endowed with the emotional circuit. The emotional circuit tends to be activated only when the predator first appears and therefore it controls the robot's behavior only at that time. After the predator has appeared and it is approaching the robot, the emotional circuit ceases to be activated and the control of the robot's behavior

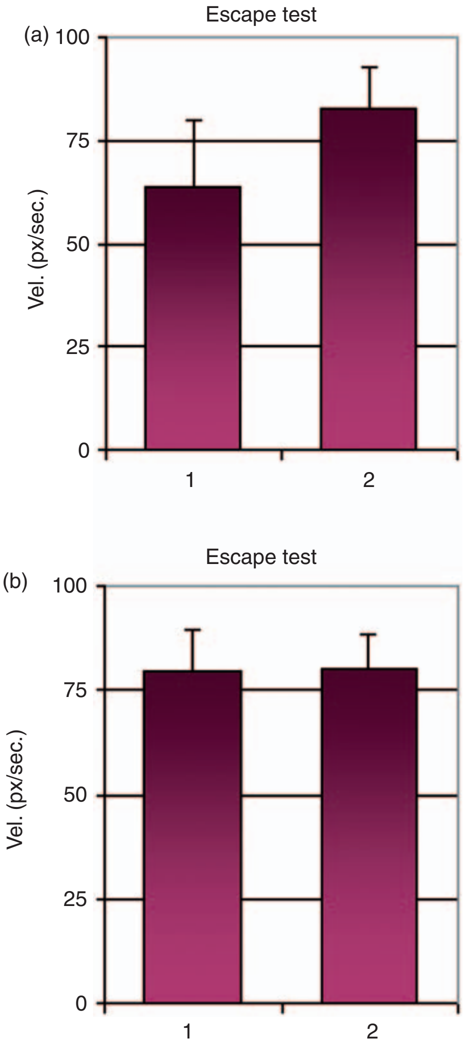


Figure 4. (a) Speed of robots with emotional circuit in an empty environment (1) and in an environment with only a predator (2). (b) Same for robots without emotional circuit.

shifts to the cognitive internal units which process the more specific information concerning the direction from which the predator is approaching the robot, allowing the robot to go in the opposite direction. As will be remembered, the emotional circuit of these robots is made of a single emotional unit. If we analyze the activation level of this unit we find that the unit typically encodes, that is, has an activation level which co-varies with, the distance of the predator, not the direction from which the predator arrives, and its activation is highest when the predator first appears and is still distant while it decreases when the predator approaches. This seems to indicate that the emotional unit controls the robot's behavior when the predator first appears and the robot has to rapidly shift from looking for food to escaping from the predator. The unit determines a quick, although imprecise, response (running in a randomly chosen direction) to the first appearance of the predator. On the other hand, when the predator approaches, the emotional unit has a lower activation level and the control of the robot's behavior shifts from the emotional circuit to the cognitive circuit which encodes the direction from which the predator is approaching the robot. Based on this more specific information the motor units produce a movement in the opposite direction with respect to the predator. This is in accord with our suggested explanation (see above) that when the emotional circuit is activated by external (environmental) sensory input, the emotional circuit gives better results if it is directly connected to the motor output. Even if the sensory input units encoding the presence and location of the predator are connected to both the internal units and the emotional unit, the emotional circuit appears to process only the motivationally relevant content of the input while their cognitive content is processed by the internal units.

7.2 Attention

Another way to understand what difference it makes for a robot to have an emotional circuit is examine what the robots pay attention to. Motivation and attention are clearly related phenomena. When the behavior of an organism is controlled by one specific motivation, the organism will attend to the stimuli that are relevant to satisfy the currently active motivation and it will ignore the other stimuli that arrive to the organism's sensors. Attention can be described in purely behavioral terms but our robots make it possible to define attention in neural terms and to determine what a robot is attending to by examining the robot's neural network.

We have done another experiment with our robots. We expose a robot to a single stimulus for a single time step and then to the same stimulus together with another stimulus, and we measure the difference between the activation level of the two motor units of the robot's neural network in response to the first stimulus alone and to the first stimulus together with the other stimulus. If there are no differences, we are entitled to say that, when both stimuli are presented, the robot is attending to the first stimulus and ignoring the second stimulus, whereas if there are differences, the robot is paying some attention to the second stimulus. This gives us a quantitative measure of attention: the quantitative difference between the activation levels of the two motor units in response to the first stimulus alone and in response to the first stimulus together with the second stimulus. In the experiments we can also vary the position of the stimuli with respect to the robot and other parameters such as the state of the robot's body (hunger).

In our robots which have to both eat and avoid being killed by the predator the emotional circuit influences the robot's attention by causing the attention to immediately shift from food to predator when the predator appears. We have put individual robots in two different controlled environments, one with the predator and a single food patch and the other with only the predator and no food, and we have measured the difference between the activation levels of the robots' motor output units in the two conditions. The results indicate that for the robots without the emotional circuit their motor units have an activation level which is very different when food is present than when food is absent. In other words, the robots continue for a while to attend to food even if the predator has appeared. In contrast, the robots with the emotional circuit have more or less the same activation level in their motor units independently of the presence or absence of food. In other words, unlike the robots without the emotional circuit, the robots with the emotional circuit are able to ignore food, that is, to shift their attention from food to predator, as soon as the predator appears (Figure 5a).

But, as we have already mentioned, the emotional circuit can affect positively not only the motivational decisions of our robots but also their cognitive performance. This is shown by another test in which we compare the activation level of the robot's two motor units when the robot sees nothing and when it sees a food patch. (Notice that, because of their biases, the robot's two motor units have an activation level even if the robot sees nothing.) The results of the test show that the robots endowed with the emotional circuit pay much more attention to food (greater difference in the activation level of the motor units) than the robots without the emotional circuit (Figure 5b).

For our robots which have to both eat and find a mating partner, we have made a number of tests varying the distance of the mating partner and we have found that the robots with the emotional circuit pay attention to food when the mating partner is far away but tend to ignore food when the mating partner is closer, while the robots without the emotional circuit are unable to control their attention in the same effective way (Figure 6).

7.3 Lesioning the emotional circuit

If the robots have evolved with an emotional circuit which increases their fitness, lesioning the circuit should result in behavior that is even less fit than the behavior of the robots that never had an emotional circuit. To test this idea we have lesioned (eliminated) the emotional circuit of the robots that have to eat and avoid being killed by the predator and we have tested the lesioned robots in an environment that contains only the predator. The results show that the lesioned robots are even less able to avoid being reached by the predator than the robots that never had an emotional circuit. We conclude that the existence of an emotional circuit leads to a different overall organization of the robot's neural network and to a different distribution of tasks between the cognitive circuit constituted by the standard internal units and the emotional circuit with its special units (cf. our discussion above of the better division of labor between the cognitive and the motivational parts of the brain if the emotional circuit is present). This may explain why complex animals such as human beings that have a complex emotional regulation of their behavior may exhibit very unfit (pathological) behavior when their emotional apparatus does not function properly. (We will briefly return to pathological behavior in the next section.)

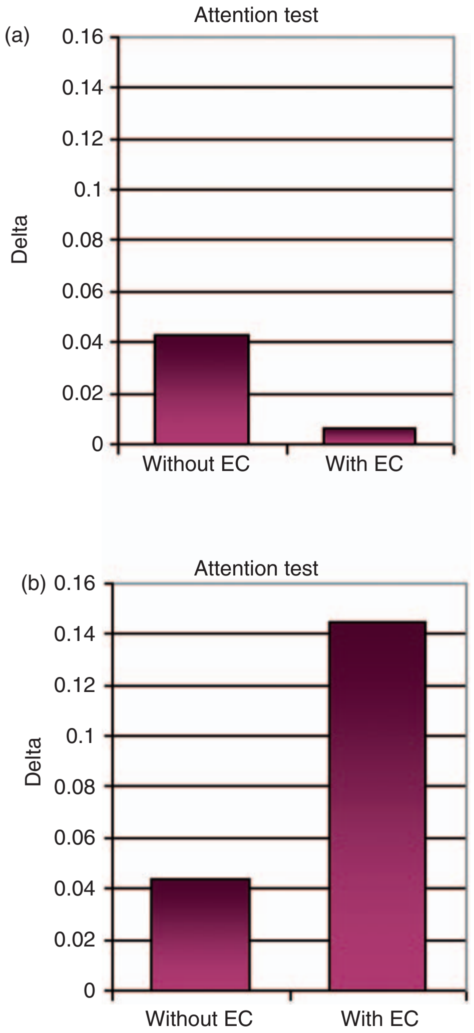


Figure 5. (a) Difference (delta) between the activation level of the motor units when the robot sees only the predator and when it sees both the predator and a food patch, for robots with and without emotional circuit (EC). (b) Same for when the robot sees food and when it sees nothing.

7.4 Motivationally conflicting stimuli

Our robots have conflicting motivations but the stimuli that arrive to their brain from the environment are each associated with one single motivation. Food is associated with the motivation to eat, water with the motivation to drink, and the predator with the motivation to fly away. In other words, stimuli are motivationally unambiguous. But consider another robot living in an environment with food and predator. In the new simulation the environment is somewhat different in that the predator does not appear at randomly chosen times but it appears just after the robot has eaten a food patch.

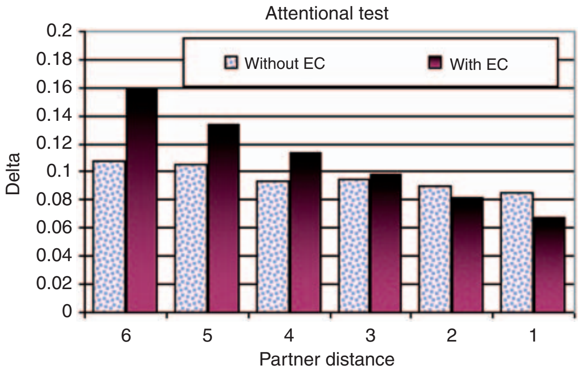


Figure 6. Difference between the activation level of the motor units when the robot sees only the mating partner and when it sees both the mating partner and a piece of food as a function of distance of the mating partner, for robots with and without emotional circuit.

For these robots food is motivationally ambiguous. It is associated with both the positive motivation to eat and the negative motivation to avoid being killed by the predator. In fact, if we examine the behavior of these robots, we see that these robots respond to food in a different way than our previous robots that live in an environment in which the predator appears at randomly chosen times. They approach food but when they are near to food they hesitate and they eat food only when their brain is informed that there is very little energy in their body, that is, when they are very hungry. What is more interesting is that for these robots adding an emotional circuit which is activated by the perception of food increases the robots' fitness more than adding an emotional circuit which is activated by the perception of the predator, as was the case for the robots described in Section 5. The predator is motivational unambiguous but food evokes conflicting motivations. Therefore it is no surprise that the robots are more helped in their motivational decisions by emotional reactions evoked by food rather than by the predator.

Motivationally conflicting stimuli can lead to pathological behavior, that is, to behavior which significantly decreases the survival and reproductive chances of the individual. In another variant of the simulation with food and predator, the predator now can only appear in one particular zone of the environment, the predator zone, and it never leaves that zone. However, the food which is present in the predator zone contains more energy than the food outside the zone. The robot is informed that it has entered the predator zone because its neural network has an additional sensory unit, a smell unit, which has an activation value of 1 when the robot is inside the predator zone and an activation value of 0 when it is outside the zone. The best robots penetrate the predator zone to eat the more energetic food which is found there and they are able to exit the zone before the predator reaches them, while less good robots either do not enter the predator zone or they enter the food zone but risk being killed by the predator when it appears. However, there are some robots that, as soon as they penetrate the predator zone, simply stop moving, that is, they "freeze," and therefore are easily killed by the predator. Clearly, the smell of the predator zone is a motivationally conflicting stimulus. It is a positive stimulus in that it is associated with better food and at the same time is a negative stimulus in that it is associated with the arrival of the predator. Some robots react to this conflict with a pathological behavior, freezing, which resembles a strong depressive state. What appears to be malfunctioning in these pathological robots is their emotional circuit because if we block (eliminate) this circuit, their behavior becomes less pathological.

8 Summary and directions of future research

To explain emotions we need to distinguish between a strategic or motivational level of behavior and a tactical or cognitive level. To survive and reproduce an animal has to be able to both decide appropriately which motivation to pursue at any given time (strategic or motivational level) and execute the behavior which allows the animal to satisfy the motivation decided at the motivational level (tactical or cognitive level). Emotions exist because they help the motivation decision mechanism to function better and therefore they increase the survival and reproductive chances of animals. This is the adaptive function of emotions and this explains why they have emerged evolutionarily.

To construct robots that have emotions it is necessary to work with robots that have more than one motivation and therefore have to autonomously decide which motivation should control their behavior at any given time. We have constructed five different types of simulated robots which have more than a single motivation and the results of our simulations show that robots whose neural network includes an emotional circuit with specific characteristics behave more effectively than robots that do not possess an emotional circuit. Our robots can be said to have emotions because it is possible to indicate the functional role that emotional states play in their behavior and to identify the particular part of the neural network controlling their behavior that makes it possible for emotional states to play their function.

We have called our robotic model of emotions an "hypothesis" about the adaptive role of emotions. The model is very simple and much work has to be done before we can consider the model as really contributing to our understanding of emotions and their role in behavior. What phenomena should our model be able to reproduce (explain)? As we have suggested, we should be able to show how our robots that have emotions can exhibit behavioral pathologies. Two other directions of research are (1) taking into account how emotional states are actually implemented in the brain/body and (2) how individuals can differ from one another in the role that emotions play in their behavior, which has an important role in interindividual differences in character or personality.

The neural network of our robots should be made more complex so that its structure and functioning will more closely match what we know about the emotional brain. Our emotional circuit is only very distantly related to biological circuits involving emotions, and this is a limitation which should be eliminated in future versions of the model (for some attempts in this direction, see Barto, 1995; Dayan & Yu, 2006;

Doya, 2002; Hestenes, 1992; Mirolli, Mannella, & Baldassarre, in press; Rolls & Treves, 1998). But the brain is not enough. Emotions result from the interaction between the brain and the rest of the body. As originally proposed by James and Lange (1922) and elaborated by Damasio (1994, 2004) and LeDoux (1996, 2000), felt emotional states are largely a product of these brain/body interactions. Our emotional circuit should send its outputs to different parts of the body, both internal organs and systems and the external body, and receive inputs from these parts, and this would allow the model to address the questions raised by Damasio's hypothesis of "somatic markers" (Damasio, 1994, 2004; Rolls, 1999).

Interindividual differences in motivation and emotion are another source of empirical evidence that our model should be able to address. To study interindividual differences in motivation and emotion it is necessary that we work not with single robots but with populations of individually different robots, and this has been an important reason in favor of using genetic algorithms in our simulations. In fact, some interesting interindividual differences have already emerged in our robots. We have found robots that tend to almost stop moving when they have both a high level of hunger and a high level of thirst and other robots that in these same circumstances more easily decide to look either for food or for water; robots that even when they are very hungry approach water if water is very close, in contrast to other robots that in the same circumstances ignore water and continue to look for food; and, in the simulations with food and predator, robots that are very "fearful" of the predator and simply stop moving as soon as the predator first appears, while other robots are still attracted by food when the predator appears but is still distant.

Other directions of research are the study of social motivations and emotions and the expression of emotions as a mechanism to facilitate social interactions, and the construction of "emotional artifacts" (e.g., religious or artistic artifacts) which activate the emotional circuit so that it can better regulate motivational decisions. More generally, our (any) robotic model of emotions should be able to answer the many questions about emotions that Klaus Scherer and Paul Eckman list in the Introduction of the book they edited more than 25 years ago (Scherer & Eckman, 1984).

Acknowledgments

Thanks to Stefano Nolfi, Marco Mirolli, and Gianluca Baldassarre for useful discussions on the topic of the article.

References

Acerbi, A., & Parisi, D. (2007). The evolution of pain. In F. Almeida e Costa et al (Eds), *Advances in artificial life* (pp. 816–824). Berlin: Springer.

Adolphs, R. (2005). Could a robot have emotions? Theoretical perspectives from social cognitive neuroscience. In J.-M. Fellous & M. A. Arbib (Eds.), *Who needs emotions? The brain meets the robot*. Oxford: Oxford University Press.

Altmann, S. (1998). *Foraging for survival*. Chicago: Chicago University Press.

Arbib, M. A., & Fellous, J.-M. (2004). Emotions: from brain to robot. *Trends in Cognitive Science*, *8*, 554–561.

Avila-Garcia, O., & Canamero, L. (2004). Using hormonal feedback to modulate action selection in a competitive scenario. In S. Schaal, A. J. Ijspeert, A. Billard, S. Vijayakumar, J. Hallam & J.-A. Meyer (Eds.), *From animals to animats 8* (pp. 243–252). Cambridge, MA: MIT Press.

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

Breazeal, C., & Brooks, R. (2005). Robot emotion. A functional perspective. In J.-M. Fellous & M. A. Arbib (Eds.), *Who needs emotions? The brain meets the robot* (pp. 271–310). Oxford: Oxford University Press.

Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, *2*, 14–23.

Canamero, L. (2003). Designing emotions for action selection in autonomous agents. In R. Trappl, P. Petta & S. Payr (Eds.), *Emotions in humans and artifacts* (pp. 115–148). Cambridge, MA: MIT Press.

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

Colgan, P. (1989). *Animal motivation*. New York: Chapman and Hall.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Grosset and Putnam.

Damasio, A. (2004). William James and the modern neurobiology of emotion. In D. Evans & P. Cruse (Eds.), *Emotion, evolution, and rationality* (pp. 3–14). Oxford: Oxford University Press.

Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Assif Mirza, N., & Blow, M. (2009). KASPAR: A minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics*, *6*, 369–397.

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: a neural interrupt signal for unexpected events. *Network*, *17*, 335–350.

Dörner, D. (2001). *Bauplan für eine Seele*. Rowohlt Taschenbuch.

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

Ellison, P. T. & Gray, P. B. (Eds.) (2009). *Endocrinology of social relationships*. Cambridge, MA: Harvard University Press.

Fellous, J.-M., & Arbib, M. A. (2005). *Who needs emotions? The brain meets the robot*. Oxford: Oxford University Press.

Frijda, N. (1987). *The emotions*. Cambridge: Cambridge University Press.

Hestenes, D. (1992). A neural network model of manic-depressive illness. In D. S. Levine & S. J. Leven (Eds.), *Motivation, emotion, and goal direction in neural networks* (pp. 209–258). Hillsdale, NJ: Erlbaum.

James, W., & Lange, C. (1922). *The emotions*. Baltimore: Williams and Wilkins.

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

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.

Lowe, R., Montebelli, A., Ieropoulos, I., Greenman, J., Melhuish, C., & Ziemke, T. (2010). *Towards and energy-mediated autonomous robot: A study of artificial metabolic constrained dynamics*. Manuscript submitted for publication.

Maes, P. (1990). A bottom-up mechanism for behavioural selection in an artificial creature. In J.-A. Meyer & S. W. Wilson (Eds.), *From animals to animats 1* (pp. 238–241). Cambridge, MA: MIT Press.

Mirolli, M., Mannella, F., & Baldassarre, G. (in press). The role of the amygdala in the affective regulation of body, brain, and behaviour. *Connection Science*.

Mitchell, M. (1998). *An introduction to genetic algorithms*. Cambridge, MA: MIT Press.

Nolfi, S., & Floreano, D. (2000). *Evolutionary robotics*. Cambridge, MA: MIT Press.

Nolfi, S., & Gigliotta, O. (2010). Evorobot: A tool for running experiments on the evolution of communication. In S. Nolfi & M. Mirolli (Eds.), *Evolution of communication and language in embodied agents* (pp. 297–301). Berlin: Springer.

Panksepp, J. (1998). *The foundation of human and animal emotions*. Oxford: Oxford University Press.

Parisi, D. (2004). Internal robotics. *Connection Science*, *16*, 325–338.

Petrosino, G., Da Rol, F., Zotti, F., & Parisi, D. (2010). *Adding an "emotional circuit" to a robot's neural network*. Manuscript in preparation.

Picard, R. W. (2000). *Affective computing*. Cambridge, MA: MIT Press.

Picard, R. W. (2003). What does it mean for a computer to "have" emotions? In R. Trappl, P. Petta & S. Payr (Eds.), *Emotions in humans and artifacts* (pp. 213–226). Cambridge, MA: MIT Press.

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

Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function*. Oxford: Oxford University Press.

Ruini, F., Petrosino, G., Saglimbeni, F., & Parisi, D. (2010). The strategic level and the tactical level of behaviour. In J. Gray & S. Nefti-Meziani (Eds.), *Advances in cognitive systems* (pp. 271–299). Herts: IET Publisher.

Saglimbeni, F., & Parisi, D. (in press). Input from the body and input from the external environment. *Darwin Meets Von Neumann, Proceedings of the 10th European Conference on Artificial Life*. New York: Springer.

Scherer, K. R. & Eckman, P. (Eds.) (1984). *Approaches to emotion*. Hillsdale, NJ: Erlbaum.

Seth, A. K. (1998). Evolving action selection and selective attention without action, attention, or selection. In R. Pfeifer et al (Eds), *Proceedings of the 5th Conference on the Simulation of Adaptive Behaviour* (pp. 139–146). Cambridge, MA: MIT Press.

Seth, A. K. (2007). The ecology of action selection: insights from artificial life. *Philosophical Transactions of the Royal Society B*, *362*, 1545–1558.

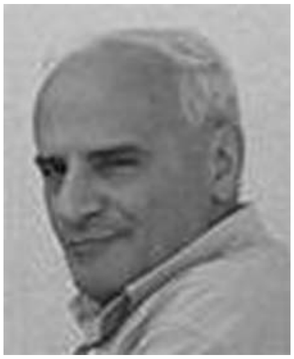
Toates, F. (1986). *Motivational systems*. Cambridge: Cambridge University Press.

Wall, P. (2002). *Pain*. New York: Columbia University Press.

Ziemke, T. (2008). On the role of emotion in biological and robotic autonomy. *BioSystems*, *91*, 410–408.

Ziemke, T., & Lowe, R. (2009). On the role of emotion in embodied cognitive architectures: from organisms to robots. *Cognitive Computation*, *1*, 104–117.

About the Authors



Domenico Parisi is past director of the Institute of Cognitive Sciences and Technologies of the National Researcher Council in Rome, where he currently does research on robotic models of animal and human behavior, with a special emphasis on both halves of the mind, cognitive and motivational/emotional. He is completing a book titled *Robots: Humanoid or human? Address*: Institute of Cognitive Sciences and Technologies, National Research Council, 44 Via S. Martino della Battaglia, 00185 Rome. *E-mail*: domenico.parisi@istc.cnr.it.

Giancarlo Petrosino is a member of the Laboratory of Autonomous Robotics and Artificial Life (LARAL) at the Institute of Cognitive Science and Technologies, National Research Council, Rome. He has a background in computer science and is interested in reproducing such phenomena as motivation, emotion, learning, behavioral pathologies, and individual differences in autonomous robots.