

## The Other Half of the Embodied Mind

Domenico Parisi

Laboratory of Artificial Life and Robotics, Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy

### Abstract

*Embodied theories of mind tend to be theories of the cognitive half of the mind and to ignore its emotional half while a complete theory of the mind should account for both halves. Robots are a new way of expressing theories of the mind which are less ambiguous and more capable to generate specific and non-controversial predictions than verbally expressed theories. We outline a simple robotic model of emotional states as states of a sub-part of the neural network controlling the robot's behavior which has specific properties and which allows the robot to make faster and more correct motivational decisions, and we describe possible extensions of the model to account for social emotional states and for the expression of emotions that, unlike those of current "emotional" robots, are really "felt" by the robot in that they play a well-identified functional role in the robot's behavior.*

### Introduction

In the Western cultural tradition the mind tends to be viewed as separated from the body and, in accordance with this tradition, the sciences of the mind try to understand the mind with no reference to the body. In the last few decades, however, this has changed. The cumulative and fast advances of the sciences of the body (neurosciences, evolutionary biology, genetics, the biological sciences more generally) make all attempts at studying the mind while ignoring the body less and less plausible. In fact, the idea that the mind is embodied and that to understand the mind it is necessary to take the body into consideration is being accepted by an increasing number of researchers and constitutes the premise of many important current investigations (Barsalou, 1999, 2008; Robbins and Aydede, 2009). The embodied view of the mind has led to a recognition of the importance of the actions with which the organism responds to the stimuli in determining how the world is represented in the organism's mind, in contrast to the traditional emphasis on mental representations as either entirely abstract or derived only from sensory input. This action-based view of the mind underlies a number of important ideas such as the grounding of symbols in the interactions of the organism with the physical environment (Harnad, 1990), the mental (neural) "simulation" of actions as a crucial component of all sorts of understanding (Gallese et al., 1996; Rizzolatti and Craighero, 2004), the mental representation of objects in terms not of their sensory properties but of the actions that the objects make possible (affordances; Gibson, 1977), the action-based nature of categories (Borghi et al., 2002; Di Ferdinando and Parisi, 2004). The embodied view of the mind is also reflected in computational models which reproduce the part of the body more directly linked to the mind, i.e., the brain (artificial neural networks) and, more recently, the entire body of the organism (robotics), and which have abandoned the disembodied view of the mind which is at the basis of artificial intelligence and of conceptions of the mind as symbol manipulation.

However, although the study of the mind can greatly benefit from an embodied conception of the mind, it still has to free itself from another tradition of Western culture which constitutes an obstacle to a complete understanding of the mind: the mind tends to be identified with cognition, that is, with knowing, reasoning, deciding, and acting. But cognition is only half of the mind. The

other half of the mind is its emotional half and, although the two halves of the mind continuously interact and behavior is a result of both halves, no satisfactory account of the mind can be provided if the science of the mind is only “cognitive” science. Today one speaks of “embodied cognition,” “grounded cognition,” and the mental representation of objects in terms of the actions with which organisms respond to them. But organisms, including humans, do not only have knowledge, goals, and the capacity to act. They also have motivations and emotional states which play a crucial role in their behavior. Current embodied views of the mind tend to be concerned with the cognitive half of the mind but an embodied account of the mind must be extended to the other half of the mind, its emotional half and, in fact, some psychologists and neuroscientists are trying to extend the embodied conception of the mind to emotions (see, for example, Gallese, 2008; Freina et al., 2009; Glenberg et al., 2009).

Even if one assumes that the mind generally does not contain anything which is unrelated to sensory input and motor output (which is what embodied theories assume), the cognitive and the emotional halves of the mind may not function in the same way. Input to the brain can be input from the environment but also input from inside the body, and output from the brain can be external motor output but also changes in the internal organs and systems of the body, and these different sensory inputs and motor outputs may have different characteristics and consequences. This is why we need models that capture both the cognitive half and the emotional half of the mind and their interactions. These models should explicitly indicate both similarities and differences between embodied cognition and embodied emotion.

This also applies to the study of the mind through the construction of computational models or robots. Robots are the most appropriate tools for exploring embodied theories of the mind because, although in a very simplified form, they reproduce the body of organisms and the physical organ that controls the organisms’ behavior (neuro-robots), and this is true for both physically realized robots and for robots which are simulated in a computer. However, current robots mostly try to reproduce the cognitive half of the mind but they ignore its emotional half. The robots displace themselves in the environment, move their arms and reach for objects, turn their eyes and their face, but they do not have emotions. Some current robots produce postures and movements of their bodies (mostly, the face) that in humans express emotions and they can recognize the expressed emotions of humans as a purely perceptual task, but they cannot be said to really have emotions and to really understand the emotions of others (Picard, 2000, 2003; Breazeal, 2002; Adolphs, 2005; Canamero, 2005; Dautenhahn et al., 2009; Robinson and el Kaliouby, 2009; cf. Arbib and Fellous, 2004; Fellous and Arbib, 2005). (For an attempt at understanding the functional role of emotions in behavior, see Ziemke, 2008.). The reason is quite simple. The cognitive half of the mind is the result of the interactions of the brain with the external environment or of processes self-generated inside the organism's brain (mental life). Current robots have artificial brains which interact with the external environment and, in some cases, can even self-generate inputs and respond to these self-generated inputs (Mirolli and Parisi, 2006, 2009; Parisi, 2007). But current robotics is an external robotics: robots reproduce the external morphology of an organism's body, the organism's sensory and motor organs, and the interactions of the organism's brain with the external environment. In contrast, the emotional half of the mind is the result of the interactions of the organism's brain with the organism's body and with the organs and systems that are inside the body. If we want to construct robots that can be said to really have emotions, what is needed is an internal robotics, that is, robots that have internal organs and systems with which the robot's brain can interact (Parisi, 2004). Only

an internal robotics can help us to better understand the emotional half of the mind and to construct a complete embodied theory of the mind.

Computational models and, more specifically, robotic models are important to understand the mind. Theories in psychology tend to be expressed verbally but verbally expressed theories have limitations because words often have different meanings for different people and because verbally expressed theories may be unable to generate specific, detailed, and non-controversial predictions. Robots are an alternative way of expressing theories. The theory is used to construct a robot and therefore, in a sense, it can be directly observed and it can contain no ambiguity because otherwise the robot cannot be constructed. Furthermore, the theory generates many specific, detailed, and uncontroversial predictions which are the behaviors exhibited by the robot. These predictions can be empirically validated by comparing them with all sorts of empirical facts: the results of behavioral experiments, data on the ecology and past evolutionary history of the organism, and data on the organism's body and brain.

As we have said, robotic models are especially appropriate for formulating embodied theories of the mind because, by definition, a robot has a body and the robot's behavior clearly depends on its having a body. Furthermore, since the brain is part of the body, to be consistent robots should be neuro-robots, that is, robots whose behavior is controlled by a system that resembles the structure and functioning of the brain, i.e., an artificial neural network. This has the advantage that it becomes possible to examine the internal representations contained in the robot's "brain" (the patterns of activation and successions of patterns of activation in the robot's neural network) and to determine if they are embodied or non-embodied representations, i.e., if they reflect the robot's actions and the reactions of the robot's internal organs and systems to sensory input rather than the sensory input. (This is more difficult to do with "emotional" robots which are not controlled by neural networks but by symbolic systems such as those of Breazeal and Brooks, 2005.)

What we will do in this paper is describe a number of simple robots that may help us to construct an entire theory of mind as made up of a cognitive half and an emotional half.

### **The Strategic and the Tactical Level of Functioning of Organisms**

The first step toward the construction of robots which have both a cognitive and an emotional mind is to construct robots that have many different motivations which cannot all be satisfied at the same time and therefore the robots have to decide at any given time which motivation to pursue with their behavior. Current robots tend to have only one motivation. In some cases their behavior is complex: for example, they may approach an object with their legs or wheels, reach the object with their arm, grasp the object with their hand, and put the object in their mouth. But behind this behavior there is a single motivation, say, the motivation to eat. In contrast, a robot's behavior may be very simple but the robot has many independent motivations. An example is a robot that has both a motivation to eat and a motivation to drink. The behavior which is needed to satisfy the motivation to eat or the motivation to drink may be very simple but the robot has to decide which motivation to satisfy if the two motivations cannot be both satisfied at the same time. Some current robots do have more than one motivation but it is their user which decides which motivation should control their behavior at any given time, i.e., what they must do, and in this sense current robots are not really autonomous. (For attempts at constructing robots that are motivationally autonomous or that take motivational decisions, cf. Brooks, 1986; Maes, 1990, although the last two references are

to symbolically controlled robots. The difference between neural network and rule-based approaches to motivations and motivational decisions is discussed in Seth, 1998, 2007.)

Real organisms are different. They have many different and largely independent motivations and one of the most important aspects of their behavior is that they have to autonomously decide which motivation to pursue at any given time. In fact, the behavior of real organisms has two levels of functioning, the strategic or motivational level and the tactical or cognitive level (Ruini et al., 2010). At the strategic level the organism has to decide which one of its many different motivations will control its behavior at any particular time. At the tactical level, the organism has to execute the behavior which is appropriate to satisfy the motivation which has been decided at the strategic level. Imagine an organism which lives in an environment containing both food and water and which to survive has to both eat and drink. Food and water are located in different places in the environment so that, at any given time, the organism has to decide whether to approach and reach food, and satisfy its motivation to eat, or to approach and reach water, and satisfy its motivation to drink. Once one of the two different motivations has been chosen at the strategic level, the tactical level of functioning of the organism generates the appropriate behavior which allows the organism to satisfy the motivation.

How is the decision taken at the strategic level? We assume that the decision is based on a simple mechanism. At any given time each motivation has an intensity which may vary from time to time, and the organism decides to pursue the motivation which currently has the highest intensity. We use the verb "to decide" but what we are talking about is implicitly expressed motivations and a purely physical mechanism that compares different motivational intensities, although in complex and verbal animals such as humans, the decision mechanism may involve talking to oneself, making explicit predictions, and producing explicit evaluations. The intensity of the different motivations is determined by two classes of factors: (1) the intrinsic intensity of the different motivations, which may depend on the overall adaptive pattern of the organism (e.g., the organism's body needs more food than water) or on the properties of the organism's environment (e.g., food is less abundant than water in the organism's environment), and (2) the current stimuli arriving to the organism's sensors from the external environment (e.g., the organism presently sees food rather than water) or from the organism's own body [e.g., the current level of nutrients (hunger) is lower than the current level of water (thirst)] or, in more complex animals such as humans, the stimuli which are self-generated (predicted, imagined, remembered) in the organism's brain. Once one motivation has been chosen at the strategic level, the organism produces the behavior that satisfies, or should satisfy, the motivation. Notice that since the intensity of the different motivations can change very rapidly because of the arrival of new stimuli and for other reasons, the organism should be able to shift very quickly from pursuing one motivation to pursuing another motivation, and this may happen even if the first motivation has not been satisfied or entirely satisfied.

That organisms function at two levels, the motivational and the cognitive level, is indicated by the fact that an individual may be good at the motivational level but not very good at the cognitive level while the opposite may be true for another individual. An individual may be good at choosing to satisfy its hunger rather than its thirst because its body needs energy rather than water but then it may not be very good at finding food. In contrast, another individual may be very good at finding both food and water but it makes the wrong motivational decisions, or is slow at deciding or, even more critically, is unable to decide and does nothing. Notice that to stay alive and possibly reproduce

an individual should be sufficiently good at both levels since both being unable to decide correctly and rapidly and being unable to do what is necessary to satisfy the motivation decided at the motivational level may reduce the organism's chances of surviving and reproducing or the organism's well-being.

This indicates that a crucial component of behavior is that the organism's motivational decision mechanism must function effectively and efficiently. It must take the correct motivational decisions and it must take these decisions at the appropriate time, which in many cases means quickly. This is where emotions come in. Emotions are states of the organism's body that allow the motivational decision mechanism to function more effectively and more efficiently. Emotional states influence the current intensity of the organism's motivations in such a way that the organism functions better at the strategic level. The sight of a potential mate might cause an emotional state in the organism that increases the probability that the organism will pursue the motivation to mate rather than other motivations. The sight of a present danger may induce an emotional state that increases the probability that the organism will stop pursuing other motivations and try to avoid the danger, and will do this fast. In more complex organisms such as humans, even the thought of a mate or of a danger may increase the importance of the motivation to mate or the motivation to avoid the danger, with the consequence that the organism will be less likely to pursue other motivations.

In the next section we describe some simple robots that have to decide among different motivations and we show that if the robots have emotions their motivational decisions are more correct and efficient.

### **Robots that Have Emotions**

We have constructed a robot which lives in an environment containing both food and a predator (Parisi and Petrosino, 2010; Ruini et al., 2010). To remain alive the robot has both to eat the food and to avoid being killed by the predator. The predator is not always present but when it appears the robot has to cease looking for food and escape from the predator. The neural network controlling the robot's behavior has sensory input units encoding the presence and location of both food and predator, motor output units encoding the movements that allow the robot to displace itself in the environment, and an intermediate layer of internal units. To eat the robot has to approach and reach the food elements while to avoid being killed by the predator the robot has to avoid physical contact with the predator. The connection weights of the robot's neural network are developed using a genetic algorithm with a population of robots that reproduce selectively and with the constant addition of random variations to the inherited connection weights (Mitchell, 1998). The robots are simulated Khepera robots (Nolfi and Gigliotta, 2010). They have a circular body with two wheels allowing the robot to displace itself in the environment and sensory organs allowing the robot to perceive what is in front of the robot within a given distance from the robot.

We compare two populations of robots. In one population the robots' neural network has the architecture we have already described. In the other population we add to the robot's neural network an "emotional circuit" made up of a certain number of units which receive activation from the sensory units encoding the presence of the predator and send their activation to either the internal units or the output units and therefore influence the robot's behavior. These emotional units have special properties compared to the other internal units of the robots' neural network. They have no "bias," they have an activation threshold, and an emotional unit is not only active in

the particular cycle in which activation arrives to the emotional unit from the input units but its activation may persist in subsequent cycles. The unit's activation threshold and the parameters that control how the activation of a unit persists in subsequent cycles all have evolved values. (For a more detailed description of the emotional circuit and of the results of the simulations, see Parisi and Petrosino, 2010.)

The results of the simulations show that the robots with the emotional circuit reach higher level of performance (they live longer) than the robots which do not have the emotional circuit in their brain. If we look at the robots' behavior, we see that they immediately cease looking for food and fly away when the predator appears, thereby reducing the probability of being reached and killed by the predator. The robots without the emotional circuit are less fast at shifting from the motivation to eat to the motivation to avoid being killed by the predator, and this leads to shorter lives.

In other simulations the robots have to make other motivational decisions: they have to decide whether to eat or drink, whether to eat or look for a mate, whether to eat or take care of their offspring, or whether to eat or rest when their body incurs some physical damage that can be healed by resting. The neural network of the robots has sensory units, motor units, and an intermediate layer of internal units and, in addition, it has internal input units encoding the current level of energy inside their body (hunger sensors) and, for the robots that have to both eat and drink, also internal input units encoding the level of water (thirst sensors), while the robots which can incur physical damage have internal input units encoding the presence of physical damage (pain sensors).

All these robots have to take motivational decisions, and their "fitness" depends on both their ability to take the appropriate motivational decisions with the required rapidity and their ability to produce the behavior which satisfies the motivation which has been chosen. The results of the simulations indicate that for all robots the possession of an emotional circuit allows them to take better and more rapid motivational decisions so that their fitness is higher than the fitness of the robots lacking the emotional circuit.

The robots we have described respond to the input from the external environment and/or from within their body not only with actions but also with emotional states, i.e., states of their emotional circuit, that make their motivational decisions more effective and efficient and therefore their behavior more "fit." The emotional circuit of these robots is very simple. The emotional units receive activation from the external environment (e.g., the sight of a predator) and/or from inside the body (e.g., the current level of energy in the robot's body) and they send their activation to the internal units or to the motor units of the robot's neural network, thereby influencing the robot's behavior. Real organisms are more complex. The equivalents of our robots' emotional units send their activation to various internal organs and systems such as the heart, the gut, the hormonal system, to the muscles of the face, and to other parts of the body, and receive activation from all these parts of the body, and it is this activation which influences the organism's motivational decisions and therefore the organism's behavior. In fact, as originally proposed by James and Lange (1922), and recently elaborated by Damasio (1994, 2004) and LeDoux (1996, 2000), felt emotional states are largely the product of these interactions. However, although very simple, our robots can be said to have emotions and to implement an embodied theory of the other half of the mind.

## **Emotions and Sociality**

The robots we have described in the preceding section do not have sociality. They live alone in their environment, and predators, mates, and offspring are not robots themselves but are only objects which are present in the environment and are perceived by the robots. Many animals, and especially humans, are social, which means that they live with conspecifics and a large portion of their adaptive pattern consists in being able to interact appropriately with conspecifics. What is the relation of emotions to sociality?

Emotions and sociality are linked in two distinct ways. Many motivations in social animals can be satisfied only with the participations of conspecifics, i.e., they are social motivations. For example, in species that reproduce sexually, an individual cannot satisfy its motivation to have offspring without the participation of an individual of the opposite sex. Since emotional states help the strategic level of functioning of the organism to take better motivational decisions, social animals tend to have social emotional states that allow them to take better motivational decisions when these decisions concern social motivations. The other way in which emotions and sociality are related is that emotional states are associated with postures and movements of the organism's body which can be perceived by other individuals and therefore can inform these individuals about the emotional states of the organism.

Examples of social motivations are the motivation to mate with another individual, the motivation to take care of one's offspring, and the motivation to be helped or to avoid being damaged by another individual. All these motivations enter into the motivational decisions of the organism which has to decide which motivation to pursue at any given time. Hence, social organisms tend to have social emotional states which regulate the current intensity of their social motivations and positively influence their motivational decisions that have to take into account social motivations.

It may not be too difficult to construct robots that have social emotions. In some recently completed simulations, the robots can be either female or male and to reproduce a robot has to mate with a robot of the opposite sex. The robots live in an environment with both food and other robots of the two sexes and to leave their genes to the next generation they have both to eat to remain alive and to approach and reach a robot of the opposite sex to reproduce. Female and male robots have different colors so that they can be recognized as females or males by conspecifics. But there is a further complication. After reproductively mating with a male robot, a female robot cannot reproduce for a certain number of time steps and during this non-reproductive period the female robot changes its color and this change of color is perceived by the other robots. A pregnant female knows its current state because its neural network has internal input units encoding the pregnant state of its body. The results show that while the average number of offspring is necessarily identical for females and males, males have more reproductive variability than females, that is, there are males with many offspring and males with very few or no offspring while females all have more or less the same number of offspring. Furthermore, the robots alternate appropriately between looking for food and looking for mates, with reproductive females behaving differently toward males compared to non-reproductive females, and males approaching only reproductive females.

The next step is to add an emotional circuit to the neural network of these robots and to see if in this case too the possession of the emotional circuit leads to better performance. What would make the states of this circuit social emotional states is that, unlike the emotional states of the robots described in the preceding Section, the emotional states of these robots would be activated by the

sight of a conspecific and they would allow a robot to take better motivational decisions about what to do socially.

This leads us to the second aspect that links emotions to sociality. The emotional states of the robots we have just described are social only in the sense that they are associated with social motivations but they are not social in the sense that they are expressed, that is, they are communicated to other individuals. What appears to be really important to understand the role of emotions in sociality is to construct robots that express their emotional states and in this manner cause other individuals to know their emotional states and to be influenced in their behavior by this knowledge.

How can we construct robots that express emotions they really have? As we have said, current “emotional” robots express emotions that they do not have, that is, that do not play any functional role in their behavior. To construct robots that express emotional states which they really have it is necessary to link the emotional circuit of the robots we have described in the preceding section to postures and movements of the robot's body that can be perceived by other robots. In the robots described in the preceding section, the emotional circuit influences the manner in which the robot responds to both external and internal inputs but it has no interactions with the rest of the robot's body. As we have said, this is not so in real organisms, in which the emotional circuit of the brain sends its activation to other parts of the body and these parts of the body respond by sending activation to the brain. As we have also said, some of the parts of the body activated by the emotional circuit are external and therefore their state or change of state (postures and movements) can be perceived by another individual. In this way an individual may know the emotional states of another individual by observing the postures and movements of the body of the other individual that result from the activations sent to the body by its emotional units.

Why should the emotional states of an individual be reflected (expressed) in the postures and movements of the individual's body and therefore be accessible to the sensors of another individual? Why should the other individual be able to understand the emotional states of the individual by perceiving the postures and movements of its body? Postures and movements of the body may be simply a by-product of having emotional states, with no specific adaptive value. But the richness of the expression of emotions in some animals, and especially humans, seems to indicate that expressing one's emotions has adaptive value and has evolved for this reason. This adaptive value is informing other individuals of one's emotional states. The adaptive value appears to consist in two things: predictability and manipulation. For the other individual it may have adaptive value to know the emotional states of the first individual by perceiving the postures and movements of its body. Adapting to the environment is to a large extent to be able to predict the future state of the environment given its present state in order to prepare for the future state. For social animals an important component of their adaptive environment is constituted by conspecifics in that many of the motivations of an individual can only be satisfied with the participation of other individuals. Therefore, for a social individual it is important to be able to predict the behavior of other individuals, and knowing the emotional states of other individuals is a powerful predictor of their behavior. This may explain why social animals (of some complexity) tend to attend to and to be able to understand the emotional states of other individuals as these emotional states are expressed by the postures and movements of certain parts of their body. But why should an individual care to express its emotional states through the postures and movements of its body so that another individual can know them? The explanation in this case is not in terms of predictability but in terms



of manipulation, where manipulation is behaving in such a way that the other individual will behave in some desired manner. In other words, an individual will express its emotional states because this will induce another individual which is informed of such emotional states to behave in ways which are desirable for the first individual.

Both functions of the expression of emotional states can be simulated with robots possessing an emotional circuit which interacts with the robot's body so that specific states of the circuit will cause specific postures and movements in the robot's body which can be perceived by other robots. This will allow the first robot to let its emotional state be known by other robots and therefore to influence their behavior and will allow the other robots to predict the behavior of the first robot.

Being able to predict X is an important component of (or perhaps the same thing as) understanding X. What is for individual A to understand individual B? The embodied theory of sociality says that A understands B by “simulating” B's perceived actions in its own brain (Gallese, 2010). But the behavior of B does not only consist in doing actions but also in expressing emotions, and sociality implies understanding not only B's actions but also B's expressed emotions. Hence, a more complete theory of sociality proposes that A understands B not only by observing B's actions and “simulating” these actions in its own brain but also by observing B's expressed emotions and “simulating” these emotions in its own brain–body. (Remember that, while the cognitive half of the mind is a product of the brain, the emotional half is a product of the interaction between the brain and the body.) There is an accumulating experimental literature on the role of expressed emotions in sociality. For example, Bayliss et al. (2007) have shown that how objects are affectively evaluated by an individual is influenced by the gaze and emotional expression of another individual, and Ferri et al. (2010) that the emotional expression of another individual may affect the goal-directed behavior of the observer. These are among the experimental results which our robots that have emotions should be able to replicate.

### **Emotions and Language**

An extension of the embodied view of mind to language proposes, and tries to show with experiments, that we understand nouns by internally representing in our brain the action with which we respond to the object designated by the noun and we understand verbs by “simulating” in our brain the action designated by the verb (Fischer and Zwaan, 2008). However, if the mind has two halves, the cognitive and the emotional half, an embodied conception of language should give an account of both the cognitive and emotional meanings of words. What we do in this section is describe some simple robots that (begin to) have language, and to suggest how words can evoke in these robots not only a cognitive meaning but also an emotional meaning.

Most words are sounds, or phono-articulatory movements that produce acoustic sounds, which covary with specific objects (nouns) or actions (verbs). If objects and actions evoke emotional states, words will also evoke emotional states. The “cognitive” orientation of most current theories of the mind can also be seen in the privilege accorded to the “cognitive” component of the meanings of words rather than to their “emotional” component, where the cognitive component refers to the perceptual properties of the object designed by a noun or to the action designed by a verb while the emotional component refers to the emotional states evoked by the noun or verb. But if we want to construct robots that can be said to have language, it will be necessary that their words evoke emotional states, not only “cognitive meanings,” in other robots. This requires an appropriate

robotic model of the meanings of words. We will now briefly describe such a model. (For a more detailed description of the model, see Parisi, 2010. For robotic models of nouns and verbs, see Cangelosi and Parisi, 2001.)

Our robots live in an environment with other robots and their neural network is made up of two sub-networks, the non-linguistic sub-network (NoL) and the linguistic sub-network (L). NoL is made of input units encoding non-linguistic sensory input (perceived objects and perceived actions of another robot) and output units encoding non-linguistic actions (e.g., reaching and grasping an object). L has input units encoding linguistic sounds (produced by another robot) and output units encoding phono-articulatory movements that produce linguistic sounds. Both NoL and L have a layer of internal units which connects the input units to the output units. In addition, the two internal layers also have horizontal connections linking the internal units of NoL to those of L, and vice versa. Given these horizontal connections, perceiving an object can lead to executing a non-linguistic action but also to executing a phono-articulatory action, i.e., producing the word which designates the object. And hearing a word may lead to reproducing the word by executing the appropriate movements of one's phono-articulatory organs (imitating the sound of the word) or to executing a non-linguistic action. When the robot learns language (which, in children, begins at 1 year of age), the robot learns the appropriate connection weights for these horizontal connections. What is the role of these connection weights? In the robot's social experience one specific linguistic sound tends to co-vary with one specific object or action, and vice versa. The robot incorporates these co-variations in the connection weights of its neural network so that when the robot hears a linguistic sound and activation spreads from the internal layer of L to the internal layer of NoL, the pattern of activation evoked in the internal units of NoL is more or less the same pattern of activation evoked by perceiving the object or action which co-varies with the linguistic sound. When the robot perceives an object or action, activation spreads from the internal units of NoL to the internal units of L, which causes the robot to produce the sound which in its experience co-varies with that object or action. In other words, the robot is able to both produce and understand language.

A crucial assumption of the model we have described is that there is no entity in the robot's neural network which can be called the meaning of a word. If we assume that the internal units of NoL have internal connections linking the units of the same layer, when an activation pattern is evoked in the internal units of NoL, this activation pattern will evoke a second activation pattern in the same or other units, then a third, and so on, so that the initial activation pattern is only the first step of a process which has no natural end and no fixed boundaries and is influenced by a number of factors such as the linguistic and non-linguistic context in which the word is being experienced by the robot, inter-individual differences among the robots, the frequency with which a word has been experienced by the robot, and others. Notice that the internal layer of NoL is not a specialized "semantic module" but is the entire brain (minus the L sub-network). This implies that the spreading of activation in NoL can invade the emotional circuit of the robot's neural network and in this manner it can trigger emotional states in the robot. These emotional states triggered by words constitute the emotional meaning of words. Words have emotional meaning if a word co-varies in the robot's experience with non-linguistic experiences that cause emotional states in the robot.

This of course is only a first step toward a robotic account of the emotional meaning of words. Here is a list of interesting questions that this account should be able to answer. Why, while all words have a cognitive meaning, only some words appear to have an emotional meaning? Why the

cognitive meaning of words appears to be more well-defined, articulated, and specific than their emotional meaning? How can we capture this fact with our robots? Words often appear to have emotional meaning not in isolation but in the context of other words, i.e., in sentences (Havas et al., 2007). Why this is so? In some experiments it has been shown that abstract words tend to have more emotional meaning than concrete words (Kousta et al., 2009a). Why? Other experiments indicate that emotional words are processed more rapidly than non-emotional words (Kousta et al., 2009b). Why?

## **Summary**

Theories expressed as robots have two important advantages compared to verbally expressed theories: they are unambiguous because they refer to things that can be observed and measured and they generate many detailed and non-controversial empirical predictions, which are the behaviors of the robot and what happens in the robot's "brain." We have described some robotic models that begin to address the question of how to develop a theory of mind that takes into consideration both the cognitive and the emotional halves of the mind. We have shown that if we add an emotional circuit to the neural network that controls a robot's behavior, the robot's behavior becomes more effective because the emotional circuit allows the robot to take more correct and faster motivational decisions. Unlike current "emotional" robots, robots that possess this circuit can be said to actually have emotions in that one can show that the circuit plays a clear functional role in the robot's behavior. We have then addressed the question of how emotions are related to sociality and we have distinguished two aspects of this relation. Robots can have social motivations, that is, motivations that can only be satisfied with the participation of conspecifics, and emotions can be said to be social if they cause the robot to make better motivational decisions when the choice set includes social motivations. The other link between emotions and sociality is the expression of emotions. The emotional circuit allows the brain (neural network) to interact not only with what is inside the body but also with the external body, causing postures and movements of some parts of the body (especially the face) that can be perceived by other robots. We have advanced the hypothesis that this expression of emotions has two adaptive advantages: it allows a robot that expresses its emotions to manipulate the behavior of other robots that perceive the expressed emotions and it allows the robot which perceives the emotions expressed by another robot to predict and anticipate the behavior of the other robot. Finally, we have briefly discussed which type of neural network should control the behavior of robots that have language and how words can have both cognitive and emotional meanings for a robot.

The robotic models we have described are very simple and very tentative and one still has to show that they are able to explain (reproduce) the constantly accumulating empirical facts about brain and behavior that involve the other half of the mind. (For some attempts at simulating the brain involved in emotions, see Rolls and Treves, 1998; Rolls, 1999; Avila-Garcia and Canamero, 2004.) Furthermore, we have said that current "emotional robots" do not actually have emotions because emotions do not play any clearly identifiable functional or adaptive role in their behavior, but the results obtained with our robotic models should be compared with those of alternative models, both robotic and non-robotic. An important test of the models will be their ability to reproduce not only the results of experiments but also other empirical facts such as inter-individual differences in the emotional half of the mind, i.e., differences in personality and character rather than in cognitive ability, the pathologies of the emotional half of the mind, i.e., psychiatric and psychological

disturbances rather than neurological ones (Stein and Ludik, 2008), and the expression of the emotional half of the mind in the production and appreciation of works of art.

## References

Adolphs R. (2005). "Could a robot have emotions? Theoretical perspectives from social cognitive neuroscience," in *Who Needs Emotions? The Brain Meets the Robot*, eds Fellous J.-M., Arbib M. A., editors. (Oxford: Oxford University Press; ), 9–25.

Arbib M. A., Fellous J. M. (2004). Emotions: from brain to robot. *Trends Cogn. Sci.* 8, 554–561. doi: 10.1016/j.tics.2004.10.004. [PubMed] [Cross Ref]

Avila-Garcia O., Canamero L. (2004). "Using hormonal feedback to modulate action selection in a competitive scenario," in *From Animals to Animats 8: Proceedings of the 8th International Conference on Simulation of Adaptive Behavior*, eds Schaal S., Ijspeert A. J., Billard A., Vijayakumar S., Hallam J., Meyer J. A., editors. (Cambridge, MA: Bradford Book), 243–252.

Barsalou W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–609. [PubMed]

Barsalou W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639. [PubMed] [Cross Ref]

Bayliss A. P., Frischen A., Fenske M. J., Tipper S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition* 104, 644–653. doi: 10.1016/j.cognition.2006.07.012. [PubMed] [Cross Ref]

Borghetti A. M., Di Ferdinando A., Parisi D. (2002). "The role of perception and action in object categorization," in *Connectionist Models of Cognition and Perception*, eds Bullinaria J. A., Lowe W., editors. (Singapore: World Scientific; ), 40–50.

Breazeal C. (2002). *Designing Sociable Robots*. Cambridge, MA: MIT Press.

Breazeal C., Brooks R. (2005). "Robot emotion. A functional perspective," in *Who Needs Emotions? The Brain Meets the Robot*, eds Fellous J.-M., Arbib M. A., editors. (Oxford: Oxford University Press; ), 271–310.

Brooks R. (1986). A robust layered control system for a mobile robot. *IEEE J. Robot. Autom.* 2, 14–23. doi: 10.1109/JRA.1986.1087032. [Cross Ref]

Canamero L. (2005). Emotion understanding from the perspective of autonomous robots research. *Neural Netw.* 4, 445–455. doi: 10.1016/j.neunet.2005.03.003. [PubMed] [Cross Ref]

Cangelosi A., Parisi D. (2001). "How nouns and verbs differentially affect the behaviour of artificial organisms," in *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, eds Moore J. D., Stenning K., editors. (Hillsdale, NJ: Erlbaum), 170–175.

Damasio A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Grosset/Putnam.

- Damasio A. (2004). "William James and the modern neurobiology of emotion," in *Emotion, Evolution, and Rationality*, eds Evans D., Cruse P., editors. (Oxford: Oxford University Press; ), 175–188.
- 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. *Appl. Bionics Biomech.* 6, 369–397. doi: 10.1080/11762320903123567. [Cross Ref]
- Di Ferdinando A., Parisi D. (2004). "Internal representation of sensory input reflects the motor output with which organisms respond to the input," in *Seeing and Thinking*, ed. Carsetti A., editor. (Boston, Kluwer; ), 115–141.
- Fellous J.-M., Arbib M. A. (2005). *Who Needs Emotions? The Brain Meets the Robot*. Oxford: Oxford University Press.
- Ferri F., Stoianov I. P., Gianelli C., D'Amico L., Borghi A. M., Gallese V. (2010). When action meets emotions: how facial displays of emotion influence goal-related behavior. *PLoS ONE* 10, e13126. doi: 10.1371/journal.pone.0013126. [PMC free article] [PubMed] [Cross Ref]
- Fischer M. H., Zwaan R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol.* 61, 225–250. doi: 10.1080/17470210701623605. [Cross Ref]
- Freina L., Baroni G., Borghi A. M., Nicoletti R. (2009). Emotive concept-nouns and motor responses: attraction or repulsion? *Mem. Cogn.* 37, 493–499. doi: 10.3758/MC.37.4.493. [Cross Ref]
- Gallese V. (2008). Empathy, embodied simulation and the brain. *J. Am. Psychoanal. Assoc.* 56, 769–781. doi: 10.1177/0003065108322206. [PubMed] [Cross Ref]
- Gallese V. (2010). "Embodied simulation and its role in intersubjectivity," in *The Embodied Self: Dimensions, Coherence, and Disorders*, eds Fuchs T., Sattel H. C., Henningsen P., editors. (Stuttgart: Schattauer; ), 78–92.
- Gallese V., Fadiga L., Fogassi L., Rizzolatti G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. doi: 10.1093/brain/119.2.593. [PubMed] [Cross Ref]
- Gibson J. J. (1977). "The theory of affordances," in *Perceiving, Acting, and Knowing. Toward an Ecological Psychology*, eds Shaw R., Bransford J., editors. (Hillsdale, NJ: Erlbaum; ), 67–82.
- Glenberg A. M., Webster B. J., Mouilso E., Havas D., Lindeman L. M. (2009). Gender, emotions, and language comprehension. *Emot. Rev.* 1, 151–161. doi: 10.1177/1754073908100440. [Cross Ref]
- Harnad S. (1990). The symbol grounding problem. *Physica D* 42, 335–346. doi: 10.1016/0167-2789(90)90087-6. [Cross Ref]
- Havas D. A., Glenberg A. M., Rinck M. (2007). Emotion simulation during language comprehension. *Psychon. Bull. Rev.* 14, 436–441. doi: 10.3758/BF03194085. [PubMed] [Cross Ref]
- James W., Lange C. (1922). *The Emotions*. Baltimore: Williams and Wilkins.

Kousta S., Vigliocco G., Vinson D., Andrews M. (2009a). "Happiness is... an abstract word: the role of affect in abstract knowledge representation," in Proceedings of the 31st Meeting of the Cognitive Science Society, Amsterdam.

Kousta S., Vinson D. P., Vigliocco G. (2009b). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 112, 473–481. doi: 10.1016/j.cognition.2009.06.007. [PubMed] [Cross Ref]

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. *Annu. Rev. Neurosci.* 23, 155–184. doi: 10.1146/annurev.neuro.23.1.155. [PubMed] [Cross Ref]

Maes P. (1990). "A bottom-up mechanism for behavioural selection in an artificial creature," in *From Animals to Animats*, eds Meyer J.-A., Wilson S. W., editors. (Cambridge, MA: MIT Press; ), 169–175.

Mirolli M., Parisi D. (2006). "Talking to oneself as a selective pressure for the emergence of language," in Proceedings of the 6th International Conference on the Evolution of Language, eds Cangelosi A., Smith A. D. M., Smith K., editors. (Singapore: World Scientific), 214–221. doi: 10.1142/9789812774262\_0028. [Cross Ref]

Mirolli M., Parisi D. (2009). Language as a cognitive tool. *Minds Mach.* 14, 517–528. doi: 10.1007/s11023-009-9174-2. [Cross Ref]

Mitchell M. (1998). *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.

Nolfi S., Gigliotta O. (2010). "Evorobot: a tool for running experiments on the evolution of communication," in *Evolution of Communication and Language in Embodied Agents*, eds Nolfi S., Mirolli M., editors. (Berlin: Springer; ), 297–301.

Parisi D. (2004). Internal robotics. *Connection Sci.* 16, 325–338. doi: 10.1080/09540090412331314768. [Cross Ref]

Parisi D. (2007). "Mental robotics," in *Artificial Consciousness*, eds Chella A., Mazzotti R., editors. (Thorverton: Imprinting Academic; ), 191–211.

Parisi D. (2010). Robots with language. *Front. Neurobot.* 4:10. doi: 10.3389/fnbot.2010.00010.. [Cross Ref]

Parisi D., Petrosino G. (2010). Robots that have emotions. *Adapt. Behav.* 18, 453–469. doi: 10.1177/1059712310388528. [Cross Ref]

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 *Emotions in Humans and Artifacts*, eds Trappal R., Petta P., Payr S., editors. (Cambridge, MA: MIT Press; ), 87–102.

Rizzolatti G., Craighero L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230. [PubMed] [Cross Ref]

Robbins P., Aydede M. (2009). *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press.

Robinson P., el Kaliouby R. (2009). Computations of emotions in man and machines. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3441–3447 . doi: 10.1098/rstb.2009.0198. [PMC free article] [PubMed] [Cross Ref]

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 *Advances in Cognitive Systems*, eds Gray J., Nefti-Meziani S., editors. (Herts: IET Publisher; ), 271–299.

Seth A. K. (1998). “Evolving action selection and selective attention without action, attention, or selection,” in *Proceedings of the 5th Conference on the Simulation of Adaptive Behaviour*, eds Pfeifer R., Blumberg B., Meyer J-A., Wilson S., editors. (Cambridge, MA: MIT Press).

Seth A. K. (2007). The ecology of action selection: insights from artificial life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1545–1558. doi: 10.1098/rstb.2007.2052. [PMC free article] [PubMed] [Cross Ref]

Stein D. J., Ludik J., editors. (eds). (2008). *Neural Networks and Psychopathology*. Cambridge: Cambridge University Press.

Ziemke T. (2008). On the role of emotion in biological and robotic autonomy. *Biosystems* 91, 410–408. doi: 10.1016/j.biosystems.2007.05.015.