Robot Learning Driven by Emotions

Sandra Clara Gadanho and John Hallam

Department of Artificial Intelligence

University of Edinburgh

5, Forrest Hill

Edinburgh EH1 2QL

{sandrag,john}@dai.ed.ac.uk

Telephone: +44 131 6503084

Fax: +44 131 6506899

Abstract

The adaptive value of emotions in nature indicates that they might also be useful in artificial creatures. Experiments were carried out to investigate this hypothesis in a simulated learning robot. For this purpose, a non-symbolic emotion model was developed that takes the form of a recurrent artificial neural network where emotions both depend on and influence the perception of the state of the world.

This emotion model was integrated in a reinforcement-learning architecture with three different roles: influencing perception, providing reinforcement value and determining when to re-evaluate decisions. Experiments to test and compare this emotion-dependent architecture with a more conventional architecture were done in the context of a solitary learning robot performing a survival task.

This research led to the conclusion that artificial emotions are a useful construct to have in the domain of behavior-based autonomous agents with multiple goals and faced with an unstructured environment, because they provide a unifying way to tackle different issues of control, analogous to natural systems' emotions.

Key Words: robotics; emotions; reinforcement learning

1 Introduction

Recent research provides considerable support for the claim that emotions are essential to human reasoning. This suggests that something analogous to emotions might play an important role in artificial-intelligence creatures. This is the topic of the present research which employed artificial emotions in the control of a solitary agent that adapts to its environment using reinforcement learning techniques.

A view shared by many researchers in the emergent field of emotional agents is that it is the functional aspect of emotions in cognition that should be taken into account when modeling emotional agents (Frijda and Swagerman, 1987; Cañamero, 1998) and not the replication of the experience of human emotions as reported by the individuals' subjective cognitive observations. This view was adopted for the current work which consisted of the test of several functional roles of emotions under an animat philosophy (Wilson, 1991), *i.e.* by building a complete agent where emotions form an integral part of the whole. Furthermore, it was considered important not only to develop a fully functional agent that successfully performs the task that it is devised for, but also to demonstrate that the introduced mechanisms are advantageous when compared with more traditional mechanisms. This implies an engineering approach to modeling emotions (Wehrle, 1998): the primary objective pursued was to enhance the performance of the robot and not to improve knowledge about the nature of emotions themselves. However, the effective use of emotions might hopefully contribute some clues to their understanding.

The experiments reported here compared the performance of emotional and non-emotional robots in a survival task that consists of maintaining adequate energy levels in a simulated environment with obstacles and energy sources. In order to master this task, the agents used reinforcement learning techniques (*e.g.*, Sutton and Barto, 1998; Kaelbling et al., 1996, for surveys on RL) to learn the coordination of simple hardwired behaviors which is quite a difficult problem and extensively hardwired in other robot applications.

To test the utility of emotions in the adaptation of the robot to its environment, an emotion model was designed and integrated into a reinforcement learning framework for robot control. This is a simple model based on a recurrent network, where perception and emotions influence each other. Through this mutual influence some persistence of emotional state is achieved, while maintaining reactiveness to new perceptual states. The model endows the agent with emotional states that are coherent with its contextual interaction. In the design of the model, it was considered more important to have simplified emotions that could be afforded by the

robot-environment interaction than to equip the robot with human-like emotions (Cañamero, 1998). Nevertheless, to make the text more concise and easier to follow, language is used which might implicitly attribute human emotions to the robot.

Apart from the influence on perception imposed by the model used, emotions were used in the reinforcement learning framework to fulfill the following roles: reinforcement specification and detection of significant events. So, three possible forms of emotional influence were examined:

Perception

Many emotions theorists agree that emotions are most helpful for focusing attention on the relevant features of the problem at hand and, in particular, for determining the salience of the perceptual information (Cytowic, 1993). Emotions not only help discriminate the perceptual information congruent with the emotional state (Niedenthal et al., 1994) but also direct attention toward the perceptual information that might be useful to cope with the emotional state (Derryberry and Tucker, 1994).

In the experiments, the perception of state within the learning algorithm is subject to the influence of emotions. The perceived state is is changed to be more compatible with the current emotions, thus making the relevant features of the environment more salient because those are usually the ones associated with the current emotional state. In nature, emotionally charged objects are also made more salient in perception, but this was not modeled in the experiments.

Reinforcement

Emotions are usually associated with either pleasant or unpleasant feelings that can act as reinforcement. This allows emotions to motivate the agent to approach or avoid certain emotional scenarios. It is often assumed that human decision making consists of the maximization of positive emotions and minimization of negative emotions (*e.g.*, Tomkins, 1984). Mowrer (1960) proposed that during learning stimuli are primarily associated with emotions which then drive the behavior associations. Proving a source of context evaluation, or reinforcement, is the most usual role attributed to emotions in the functionality of an artificial learning agent (*e.g.*, Wright, 1996; Kitano, 1995; Albus, 1990; Bozinovski, 1982).

This role of providing reinforcement was also tested in the experiments reported here.

Control triggering

One of the most difficult problems faced when employing reinforcement learning techniques in robotics applications is to determine when a discrete state transition occurs. This is not an issue in traditional reinforcement learning problems, because these usually consist of grid worlds where state transition consists of a single discrete action which moves the agent to one of the cells in the neighborhood of the cell where the agent is located. In a continuous world the determination of this transition is not clear and it is usually implementation dependent. This transition can be triggered by some internal or external event and must be identified by the designer, because it determines when the controller needs to re-evaluate its previous decision and make a new one.

Experiments have been done to test whether emotions can successfully fulfill the role of determining state transitions by triggering the adaptive controller whenever there is a significant change in the emotional state.

Emotions are frequently pointed to as a source of interruption of behavior (Sloman and Croucher, 1981; Simon, 1967) in the domain of more traditional symbolic Artificial Intelligence architectures. In general, it is considered that behavior should be interrupted and eventually replaced whenever a strong emotion is felt. In the current work, it is considered that if the emotional intensity falls, then behavior should also be changed, because the crisis that gave rise to the emotion has probably been solved. So state transition is triggered not only by sudden rises of emotional intensity but also by abrupt drops. Implicit in this approach is the fact that the emotion model being used is continuous and so does not provide a clear cut onset or termination of emotions, requiring that abrupt changes be detected instead.

In essence, the reported research work shows how emotions can influence control in multiple ways. Although the "emotions" used are much simplified and quite different from the human counterparts, they try to capture several functional aspects of emotions and were an integrant part of a complete agent. The next section presents reasons why emotions can be advantageous in artificial creatures and describes the model of artificial emotions developed. The following section, Section 3, gives a detailed description of the experiments made to test the different emotion roles presented above. Finally, the conclusions reached are presented in Section 4.

2 Emotions in Robots

2.1 Emotions interaction with cognition

In their quest for true intelligence, people usually adopt a Cartesian approach that regards emotions as a hindrance carried over from their early evolutionary development, at odds with their aspiration to high rationality. Psychologists, too, have tended to concur with this popular view of emotions as useless or even disruptive to rationality (Toda, 1993). However, the view that emotions are an integral part of rational behavior is receiving increasing support from brain research studies (LeDoux, 1998; Damasio, 1994; Cytowic, 1993).

Studies show that human decisions are not always logical (Grossberg and Gutowski, 1987). Pure logic is not enough and shows serious faults when used to model human intelligence in Artificial Intelligence systems (Dreyfus, 1992). Furthermore, emotions have been suggested in the field of Artificial Intelligence as the ultimate source of intelligence that might provide robots with the autonomy they need (Toda, 1994). Doubts have even been posed on whether machines can exhibit intelligent behavior without emotions (Minsky, 1986; Charland, 1995).

Emotions influence cognition in terms of different elementary mechanisms like perception (Niedenthal et al., 1994; Derryberry and Tucker, 1994; Cytowic, 1993), memory (Blaney, 1986; Bower, 1981; LeDoux, 1998), attention (LeDoux, 1998; De Sousa, 1987; Ortony et al., 1988; Frijda and Swagerman, 1987; Sloman and Croucher, 1981; Simon, 1967) and reasoning (LeDoux, 1998; Bechara et al., 1997; Damasio, 1994).

Some of the properties of emotions which have an adaptive value that might be transposed to artificial creatures are:

- Control of attention. Emotions influence perception and orient reasoning by focusing the agent's attention on the most relevant features to solve its immediate problem (*e.g.*, Beck, 1983; Morignot and Hayes-Roth, 1995). In particular, they have been attributed the role of interrupting the agent from what it is doing when new problems arise that need to be attended to (Sloman et al., 1994; Beaudoin and Sloman, 1993).

- Source of reinforcement for adaptive agents (*e.g.*, Wright, 1996; Albus, 1990; McCauley and Franklin, 1998), or the role of monitoring the agent's performance so that the agent's plans or actions can be changed if necessary (*e.g.*, Michaud et al., 1996; Shibata et al., 1996).

- Memory filters which allow better recall of events that are congruent with current emotional state (mood congruence, *e.g.* Araujo, 1994) or events that were learned while the agent was on a congruent emotional state (mood dependency, *e.g.* Bower and Cohen, 1982). This preferential recall of some events over others can affect the decision-making process by making the agent more or less optimistic depending on whether it is happy or sad, respectively (Seif El-Nasr et al., 1998; Bower and Cohen, 1982).

- Assistance in reasoning. The agent's affective system quickly obtains perceptual cues that can be used to direct the access of the cognitive information relevant for the cognitive system's deliberation (Damasio, 1994; Ventura et al., 1998).

- Behavior tendencies or even stereotyped responses are usually associated with particular emotional scenarios. These built-in responses allow for appropriate behavior to be automatically triggered in emergency situations, avoiding spending unavailable time on elaborate reasoning.

- Physiological arousal of the body. A strong emotion is usually associated with a general release of energy in anticipation of demanding action response. The translation of this feature to an artificial system can consist in the modulation of system parameters (Cañamero, 1997; Bates et al., 1992a), like for example the level of behavioral activity or speed, which are directly relevant to the overall performance of the system.

- Support for social interaction (*e.g.*, Picard, 1997; Bates, 1994; Aubé, 1998; Breazeal, 1998; Klein, 1996). The expression of emotions allows the individuals to transmit to others messages that are often crucial to their survival and therefore have great adaptive value (Darwin, 1965).

2.2 Emotion Model

Like many other psychological terms (*e.g.* intelligence, consciousness), emotion is difficult to describe and existing emotion models employ mostly working definitions that tend to conflict with each other. On the one hand, although there are behavioral and physiological aspects of emotions that can be monitored, an essential part of emotions consists of a private internal experience not subject to direct observation by others than the individual experiencing them, making proper scientific analysis difficult. On the other hand, emotions are intrinsically related to other psychological processes (*e.g.* cognition) and the artificial separation from them created

by the traditional scientific approach together with the artificiality of the experimental setups often hide away the true nature of real emotional experiences (Kaiser and Wehrle, 1996).

A large subset of theories of emotions is based on elaborate cognitive appraisal theories (*e.g.*, Lazarus, 1982; Power and Dalgleish, 1997; Ortony et al., 1988) that stress the role of conscious reasoning in the generation and definition of emotions, in spite of emotions also being aroused by crude subconscious experiences involving simple information processing without the need for high level reasoning processes (Zajonc, 1984; Izard, 1993). Moreover, these theories usually rely strongly on verbal reports which can be very deceptive for several reasons (Gadanho, 1999), the most pertinent being the fact that the emotion generation mechanisms are not necessarily knowable to the conscious mind (LeDoux, 1998).

Following the psychologists' main stream, most Artificial Intelligence models of emotions are based on an analytic and symbolic approach (Sloman et al., 1994; Frijda and Swagerman, 1987; Dyer, 1987; Pfeifer, 1982; Pfeifer and Nicholas, 1985; Bates et al., 1992b) that tries to endow the model with the full complexity of human emotions as perceived from an observer's point of view.

In opposition to the traditional approach, a synthetic bottom-up approach based on the animat approach (Wilson, 1991) was preferred for the current work. For this reason, it was considered more appropriate to investigate the generation of emotions by simple associations of stimuli instead. This led to the development of a non-symbolic emotion model. Recently, models have been suggested that also follow a bottom-up approach (Velásquez, 1998; Cañamero, 1997; Foliot and Michel, 1998; Wright, 1996). The problem with reproducing most of these models is that they usually provide so little architectural specification that they allow almost total freedom of implementation. Furthermore, the evaluation of their practical implementations is often difficult, because in general they are presented as an end result, *i.e.* the adaptiveness value of the presence of emotions is not evaluated, but only presented as fact. In these conditions, unless an objective and accurate description of the end product is given, only its direct observation can make any kind of evaluation possible.

The most significant emotion features that the designed model tries to capture are:

- Emotions have valence, *i.e.*, they provide a positive or negative value.
- Emotions have some persistence in time, *i.e.* sudden unrealistic swings between different emotions should not be allowed, particularly when the emotions in question differ a lot. The occurrence of a certain emotion depends not only on direct sensory input, but also

on the agent's recent emotional history.

- Emotions color perception in that what is perceived is biased by the current emotional state.

The model that was developed — Figure 1 — is based on four emotions ($\mathcal{E}$): Happiness, Sadness, Fear and Anger. These emotions were selected because they are believed to be the most universally expressed emotions along with Disgust (Ekman, 1992) and are usually included in the definitions of basic or primary emotions (see, for example Shaver et al., 1987; Power and Dalgleish, 1997; Goleman, 1995). Furthermore, they are adequate and useful for the robot–environment interaction afforded by the experiments. Others might prove too sophisticated or out of place. For instance, there seems to be no situation where it is appropriate for the robot to feel disgust. However, if, for instance, toxic food were added to the environment, disgust would become useful to keep the robot away from it.

$$\mathcal{E} = \{\text{Happiness, Sadness, Fear, Anger}\} \quad (1)$$

The model determines the intensity of each emotion based on the robot's current internal feelings ($\mathcal{F}$). In the current work, the robot feelings are defined as the internal perception of its raw sensations, or its subjective sensations. The intensity of each emotion is calculated through linear weighted dependencies from feelings. The set of feeling used in the experiments reported here is given in Equation 2.

$$\mathcal{F} = \{\text{Hunger, Pain, Restlessness, Temperature, Eating, Smell, Warmth, Proximity}\} \quad (2)$$

Furthermore, the emotional state also influences the robot's feelings. The feelings that give rise to an emotion are also the ones aroused by the emotion. This way, each emotion tries to influence the feelings in such a way that the resulting feelings match the state that gives rise to that particular emotion. An emotion only influences the feelings if its intensity value is significantly large, *i.e.* its value is above an activation threshold.

The emotions influence the feelings through a hormone system, by producing appropriate hormones. The hormone system in the model is a very simplified one. It consists of having one hormone associated with each feeling. A feeling intensity is not a value directly obtained from the value of the sensation that gives rise to it, but from the sum of the sensation and

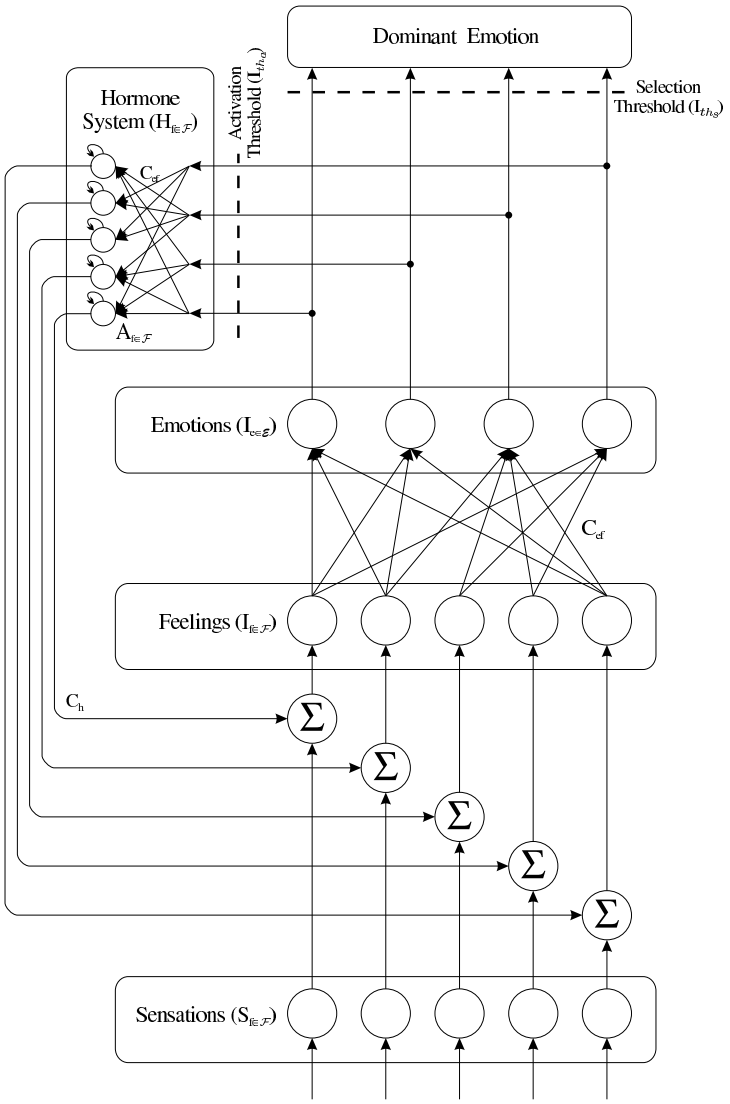

Figure 1: The emotion model. In this model, emotions do not depend directly on the agent's immediate perception of the world, *i.e.* its sensations. They depend on the feelings which are a combination of the sensations and the hormones produced recently by active emotions. This adds to the emotion state some memory of the recent past.

hormone value. The hormone quantities produced by each emotion are directly related to its intensity and its dependencies on the associated feelings. The stronger the dependency on a certain feeling, the greater quantity of the associated hormone is produced by an emotion.

On the one hand, the hormone mechanism introduces a competition between the emotions to gain control over the feelings which is ultimately what selects which emotion will be dominant. On the other hand, the robot feelings are not only dependent on its sensations but are also dependent on its emotional state, *i.e.* the intensity of its emotions.

A formal description of the model's functions is given by Equations 3 to 8. The function $\mathrm{Th}_{[b_-,b_+]}(x)$ was needed to confine values within an interval $[b_-, b_+]$.

$$\mathrm{Th}_{[b_-,b_+]}(x) = \begin{cases} b_- & \text{if } x < b_- \\ b_+ & \text{if } x > b_+ \\ x & \text{otherwise} \end{cases} \tag{3}$$

Equation 4 shows how the intensity value of emotion e at step n (I_{e_n}) is calculated from the intensity of the feelings (I_{f_n}) at that step. This calculation involves an emotion bias (B_e) and coupling coefficients (C_{ef}) between the emotion e and the feelings $f \in \mathcal{F}$.

$$\forall e \in \mathcal{E}, \forall n \in \mathbb{N}, \quad I_{e_n} = \mathrm{Th}_{[0,1]}(B_e + \sum_{f \in \mathcal{F}} (C_{ef} I_{f_n})) \tag{4}$$

The calculation of the feeling's intensity has to take into account both the influences provided by the hormone system (H_{f_n}), which are dependent on a coefficient parameter (C_h), and the value of the respective sensation (S_{f_n}). The sensations' values are directly derived from the sensory data. The hormone values are responsible for the memory of the emotion system, and depend both on their previous values and the emotion influences (A_{f_n}). Emotions only influence the hormone values if their intensity is above the activation threshold (I_{th_a}). To calculate the value of the hormones (H_{f_n}), two different system parameters are used, the attack gain (α_{up}) and the decay gain (α_{dn}). The first one is used when the emotions and their influences are increasing and the other when the emotion intensities are fading away. In general, the attack gain is much higher than the decay gain. This way the decay of emotions is slow while the emergence of new emotions is much faster. The values of the parameters used in the emotions model are given in table 1, except for the emotions' dependencies (C_{ef}) and bias (B_e) which will be given later in the context of the agent's task.

Table 1: Parameter values used in the experiments.

Parameter	Definition	Value
I_{th_a}	Emotion Activation Threshold	0.2
I_{th_s}	Emotion Selection Threshold	0.2
C_h	Hormone Coefficient	0.9
α_{up}	Hormone Attack Gain	0.98
α_{dn}	Hormone Decay Gain	0.996

$\forall f \in \mathcal{F}, \forall n \in \mathbb{N}$,

$$I_{f_n} = \mathrm{Th}_{[0,1]}(C_h H_{f_n} + S_{f_n}) \tag{5}$$

$$H_{f_n} = \begin{cases} 0 & \text{if } n = 1 \\ \alpha_n H_{f_n} + (1 - \alpha_n) A_{f_{n-1}} & \text{if } n > 1 \end{cases} \tag{6}$$

$$A_{f_n} = \sum_{e \in \mathcal{E}:\ I_{e_n} > I_{th_a}} C_{ef} I_{e_n} \tag{7}$$

$$\alpha_n = \begin{cases} \alpha_{up} & \text{if } |A_{f_n}| > |H_{f_n}| \\ \alpha_{dn} & \text{otherwise} \end{cases} \tag{8}$$

Figure 2 shows the response of an emotion e to a sensation on which it has a dependency (C_{ef}) of 0.8 weight. This dependency is actually indirect, through the respective feeling f. When the sensation value (S_f) is 1.0, the emotion intensity (I_e) is 0.8 which is the highest value possible in this example. The influence of the hormone (H_f) is only noticeable after the sensation returns to value zero. Before that, the feeling intensity (I_f) is saturated by the stimulus itself. When the stimulus disappears, the emotion intensity has a sudden drop in value because it becomes dependent solely on the total value of hormone (H_f) that accumulated while the sensation was on. The values of hormone and emotion gradually decay to zero without the presence of the sensation.

As a concrete example of the dynamics of the model in terms of robot-environment interactions consider the situation of the robot colliding with an obstacle. The collision itself produces a pain sensation that will be captured by the pain feeling. Assuming that fear has a strong dependency on pain[1], then the fear intensity will rise. If this intensity is high enough then fear

[1]Although this dependency is used in the experiments, aversive stimulation such as pain is more usually

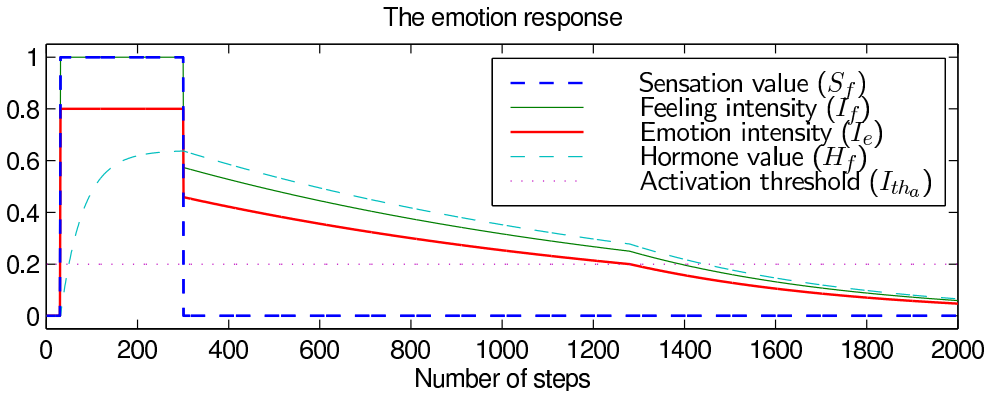

Figure 2: Emotional response to a sensation. The emotion e depends indirectly on this sensation through the respective feeling f. The emotion intensity reaches a maximum value as soon as the sensation value is set to one and it decreases drastically once the sensation value is set back to zero, but the emotion is still active long after the sensation has been set to zero.

will produce hormones. In particular, the hormone associated with pain will quickly build up during the collision. This will make the fear emotion grow stronger and possibly overtake other existing emotions. When the robot finally manages to cease the collision, it will still have pain not because the pain sensation is still there, but because the hormone associated with pain has a high value. So the fear emotion will persist while the hormone gradually decreases in value. This means that while the robot is gaining distance from the obstacle, the fear will still be there. Nevertheless, it will usually fade away as soon as a short distance is gained and the risk of further collisions has diminished.

The hormones' values can increase quite rapidly, allowing for the quick build up of a new emotional state, and decrease slowly allowing for the persistence of an emotional state even when the cause that gave rise to it is gone — another of the characteristic features of emotions. It should be noted that the time scales involved in the persistence of an emotion after the stimulus is gone, particularly when in the presence of a new stimulus that favors another emotion, are quite small. This allows for what is perceived as quick changes of emotions, in opposition to the much slower process of changes in mood. One can only talk of moods when talking of the residual hormone values that might exist in the system and are not strong enough to stimulate the existence of a dominant emotion. That would be consistent with the theory that moods are differentiated from emotions in terms of level of arousal (Panksepp, 1995). These residual

connected to anger in humans (*e.g.*, Izard, 1993).

hormone values can act as moods in the sense that they might favor the appearance of certain emotions.

Emotions were divided into two categories: positive and negative. The ones that are considered "pleasant" are positive (only Happiness, in the set of emotions used), the others are considered negative. This way a value judgment can easily be obtained from the emotion model by considering the intensity of the current dominant emotion and whether it is positive or negative. The dominant emotion is the one with the highest intensity, unless no emotion intensity exceeds a selection threshold (I_{th_s}). In this case, there will not be a dominant emotion and the value judgment is neutral.

In summary, the model of emotions described provides not only an emotional state, based on simple sensations, that is coherent with the current situation, but also influences the perception. Side issues associated with emotions as moods and temperaments were not directly built into the architecture and are only exhibited as a by-product. Different temperaments, for instance, can be achieved by having different emotion dependencies on feelings or changing other parameters of the system.

The model of emotions behaves appropriately when tested on the robot, in the sense that the robot consistently displays plausible contextual emotional states during the process of interacting with the environment. Furthermore, because its emotions are grounded in subjective "feelings", and not direct sensory input or "sensations", it manages to avoid sudden changes of emotional state, from one extreme emotion to a completely different one. The more different the two emotions are, the more difficult it is to change from one to the other. Emotion persistance could have been modelled at the level of the emotions themselves, but modelling it at the level of feelings allows for a richer interaction between emotions in their competion for dominance.

3 Experiments

This section presents a detailed account of the experiments made. First, a description of the robot, its task, its feelings and its emotions is given. Secondly, the robot learning controller used is presented and its interactions with the emotion system are fully explained. Finally, the experimental procedure is described and the results obtained are reported and discussed.

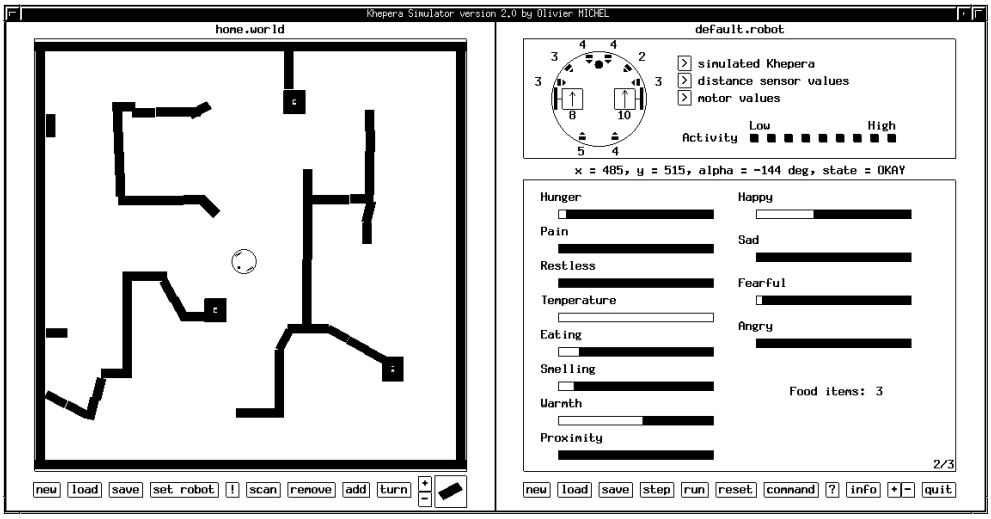

Figure 3: The simulated robot and its environment. The simulation window shows the robot's environment on the left and its internal state on the right.

3.1 The Robot and its Task

All the experiments were carried out in a realistic simulator developed by Michel (1996) of a Khepera robot — a small robot with a left and a right wheel motor, and eight infrared sensors that allow it to detect object proximity and ambient light. Six of the sensors are located at the front of the robot and two at the rear. The robot environment — Figure 3 — consisted of a closed environment with some walls and three lights surrounded by bricks.

Simulated feeding needs were added to the robot. The robot is always losing energy: the more it uses its motors the more energy is used up. It can recover its energy from lights in its environment. The main reason for having lights as food sources is to allow the robot to distinguish its food sources with its poor perception capabilities.

To gain energy from a food source, the robot has to bump into it. This will make energy available for a short period of time. At the same time an odor will be released that can be sensed by the robot. It is important that the agent is able to discriminate this state through its sensors, because the agent can get only get energy during this period. This energy is obtained by receiving high values of light in its rear light sensors, which means that the robot must quickly turn its back to the food source as soon as it senses that energy is available. To receive further energy the robot has to restart the whole process by hitting the light again so that a new time window of released energy is started.

A food source can only release energy a few times before it is exhausted. In time, the food

source will recover its ability to provide energy again, but meanwhile the robot is forced to look for other sources of energy in order to survive. The robot cannot be successful by relying on a single food source for energy, *i.e.* the time it takes for new energy to be available in a single food source is longer than the time it takes for the robot to use it. When a food source has no energy, the light associated with it is turned off and becomes a simple obstacle for the robot.

The task can be translated into multiple goals: moving around the environment in order to find different food sources and, if a food source is found, extracting energy from it. The extraction of energy was complicated in order to make the learning task harder by requiring the agent to learn sequences of behaviors. Furthermore, the robot should not keep still in the same place for long durations of time or collide with obstacles.

The sensations available to the robot were:

- **Hunger —** The robot's energy deficit;
- **Pain —** High if the robot is bumping into obstacles;
- **Restlessness —** Increases if the robot does not move and decreases otherwise, it is reset whenever a behavior is selected;
- **Temperature —** Rises with high motor usage and returns to zero with low motor usage;
- **Eating** — High when the robot is acquiring energy, *i.e.* when the hunger sensation is decreasing;
- **Smell** — Active when there is energy available — note that the robot does not have real smell sensors, odor perception is simulated for the experiments;
- **Warmth —** Directly dependent on the intensity of light perceived by the robot's light sensors;
- **Proximity** — Reflects the proximity of the nearest obstacle perceived by the distance sensors.

3.2 The Emotion System

In order to have the robot's emotional state compatible with its task, the emotions' dependencies on feelings are such that:

Table 2: The emotions' dependencies on feelings.

	Hunger	Pain	Restlessness	Temperature	Eating	Smelling	Warmth	Bias
Happiness	−0.2	−0.3	−0.2	0.2	0.4	0.3	0.0	0.1
Sadness	0.7	0.0	0.1	−0.2	−0.4	0.0	−0.2	−0.1
Fear	−0.2	0.7	−0.2	0.1	−0.2	−0.2	0.0	0.0
Anger	0.2	0.1	0.7	−0.2	−0.2	0.0	0.0	0.0

- The robot is **happy** if there is nothing wrong with the present situation. It will be particularly happy if it has been using its motors a lot or is in the process of getting new energy at the moment. Even just the smell of food can make it happy.
- If the robot has very low energy and it is not acquiring energy, then its state will be **sad.** It will be more sad if it cannot sense any light.
- If the robot bumps into obstacles then the pain will make it **fearful.** It will be less fearful if it is hungry or restless.
- If the robot stays in the same place too long it will start to get restless. This will make it **angry.** The anger will persist for as long as the robot does not move away or change its current action. A hungry robot will tend to be more angry.

Table 2 presents the actual values for each of the emotion biases and dependencies on feelings. The appropriate values were carefully chosen by hand, to reflect the qualitative dependencies described above, using a simple process of trial and error which needed few iterations. No emotion dependencies were created for the feeling of Proximity; this feeling is used only to determine state within the adaptive controller.

3.3 The Adaptive Controller

The adaptive controller selected was a well known reinforcement-learning algorithm which has given good results in the field of robotics: Q-learning (Watkins, 1989).

One of the major problems responsible for the slowness of reinforcement learning is the slow iterative process of spreading the rewards and punishments through the input space, which can be greatly minimized if the algorithm has added mechanisms to spread the reinforcements to similar input states. The system implemented profits from generalization over the input space by using neural networks to learn the utility values of each action (Lin, 1993). Apart

from accelerating the learning process, it also minimizes the memory space needed to store the policy, which is often stored in the form of a look-up table with one value for each action and sensor state combination.

For more complex tasks skill decomposition is usually advisable as it can significantly reduce the learning time, or even make the task feasible. Researchers report that a monolithic approach can fail to solve the long-term temporal credit assignment (Mahadevan and Connell, 1992). One of the reasons pointed out is the loss in accuracy of the propagation of credit assignment with long action sequences (Lin, 1992). Skill decomposition usually consists of learning some predefined behaviors in a first phase and then finding the high-level coordination of these behaviors. Although the behaviors themselves are often learned successfully (Mahadevan and Connell, 1992; Lin, 1993), behavior coordination is much more difficult and is usually hard-wired to some extent (Mahadevan and Connell, 1992; Lin, 1993; Mataric, 1994).

In the current work, the primitive behaviors were hand-designed and only the harder task of behavior coordination was learned in the hope that emotions might be helpful in solving some of problems found at this level. The developed controller tries to maximize the evaluation received by selecting between one of three possible hand-designed behaviors:

Avoid obstacles — Turn away from the nearest obstacle and move away from it. If the sensors cannot detect any obstacle nearby, then remain still.

Seek Light — Go in the direction of the nearest light. If no light can be seen, remain still.

Wall Following — If there is no wall in sight, move forwards at full speed. Once a wall is found, follow it. This behavior by itself is not very reliable in that the robot can crash, *i.e.* become immobilized against a wall. The avoid-obstacles behavior can easily help in these situations.

The controller — Figure 4 — implements a Q-learning algorithm, using neural networks, very similar to the one reported by Lin (1992). It will be defined next in terms of two separate modules: the Associative Memory Module and the Behavior Selection Module.

3.3.1 Associative Memory Module

This plastic module uses three neural networks to associate the robot feelings with the current expected utility of each of the three robot behaviors. These are three-layer feed-forward neural

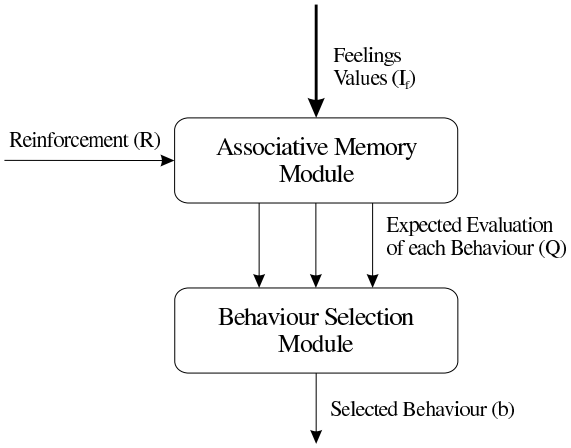

Figure 4: The Adaptive Controller. The associative memory estimates an evaluation for each behavior which is used in the stochastic behavior selection.

networks, with: 9 input units, one for each feeling and a bias; 10 hidden units; and 1 output unit that represents the expected outcome of the associated behavior.

The activation function used in both the hidden and output units was the hyperbolic tangent: $\tanh(\beta x) = \frac{1-e^{-2\beta x}}{1+e^{-2\beta x}}, \beta = 0.25$. The weights between the hidden layer and the output layer are initialized with random values, and the weights between the input layer and the hidden layer are set to zero. This way all the networks provide an initial neutral evaluation. The learning algorithm used to train the networks was back-propagation (see, for example, Hertz et al., 1991) with learning rate set to 0.3.

The input of the neural-networks consists of the perceived world state and the single output of each neural-network models the following function for one of the actions a:

$$\mathrm{Q}(s_n, a) = R_{n+1} + \gamma\ \mathrm{V}(s_{n+1}) \tag{9}$$

This function represents the expected discounted cumulative reinforcement that an agent will receive after executing action a in response to the world state s_n. The immediate reinforcement received in the next state (s_{n+1}) is R_{n+1}. The utility of the state s_{n+1}, or $\mathrm{V}(s_{n+1})$, is its expected discounted cumulative reinforcement if the optimal policy is followed by the agent, *i.e.* if it always selects the action with the highest Q function value. The value γ is the discount factor which is set to 0.9.

In each learning step of the algorithm (see Figure 5), the neural-network associated with last behavior b selected is updated for the previous state (s_{n-1}). An iteration of the back-

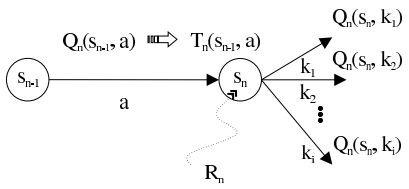

Figure 5: Learning iteration of the reinforcement-learning algorithm.

propagation algorithm is made using as target value $T_n(s_{n-1}, b)$. The calculation of this target value depends on the current estimate $Q_n(s_n, k)$ for each action k provided by the outputs of the networks when using the current state (s_n) as input.

$$T_n(s_{n-1}, b) = R_n + \gamma \max\{Q_n(s_n, k) \mid k \in \text{behaviors}\} \tag{10}$$

3.3.2 Behavior Selection Module

The utility values provided by the associative memory are used for the stochastic selection of the next behavior. The higher the value provided by the associated net, the higher the probability of a behavior to be selected. The function used to calculate the probability $P_n(s_n, b)$ of each behavior b being selected is based on the Boltzmann-Gibbs distribution:

$$P_n(s_n, b) = \frac{e^{\frac{Q_n(s_n,b)}{T}}}{\sum_{k \in \text{behaviors}} e^{\frac{Q_n(s_n,k)}{T}}} \qquad T = 0.1 \tag{11}$$

3.4 The Emotional Controller

The emotional controller — Figure 6 — was achieved by the integration of the emotion system with the learning controller. It implements emotional perception, reinforcement and control triggering.

These emotional mechanisms and the alternative mechanisms used for experimental comparison are presented next.

3.4.1 Perception

In the adaptive controller selected, the agent's perception is considered to be the neural network inputs of the associative memory. These define the input state which is perceived by the agent and is associated with rewards and punishments during the learning process. Selecting a correct

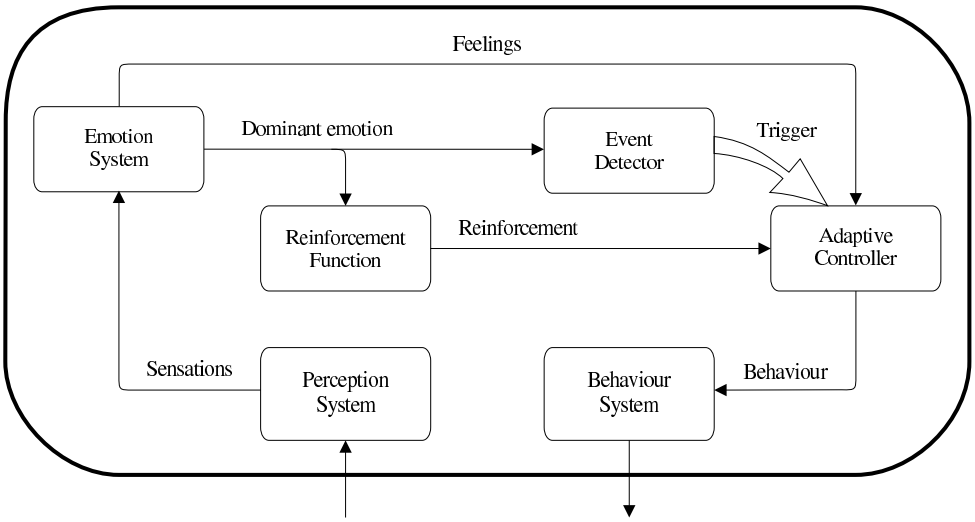

Figure 6: The emotional controller. The emotion system (Figure 1) uses the sensations returned by the perception system to calculate the current feelings and the dominant emotion. These values are used by the adaptive controller (Figure 4) to select the behavior executed by the behavior system: the feelings define the current state; and the dominant emotion determines when a state transition occurs and the reinforcement received.

input space is an important step in the design of a learning agent because it implicitly informs the agent which elements of its environment might be important for achieving its task. The input state selected was the robot's feelings. In the emotion model developed, feelings are influenced by emotions through the hormone system which rises the value of those feelings that are related to the current emotional state. So the represented robot state is emotion-dependent: the state which the robot learns to associate with rewards is being biased by emotions.

To test *whether perception being influenced by the hormone system or not has an impact on the performance of the robot* — Figure 7 — the experimental results of using feelings as network inputs was compared with those obtained by replacing them by sensations which are emotion independent.

3.4.2 Reinforcement

The reinforcement function must specify the correct behavior of the agent by giving it rewards when it is performing well and punishments when its behavior is inadequate. For this reason it implicitly specifies the agents' goals or motivations or its task. In the emotional controller developed, emotions were considered the fundamental source of motivation and were therefore

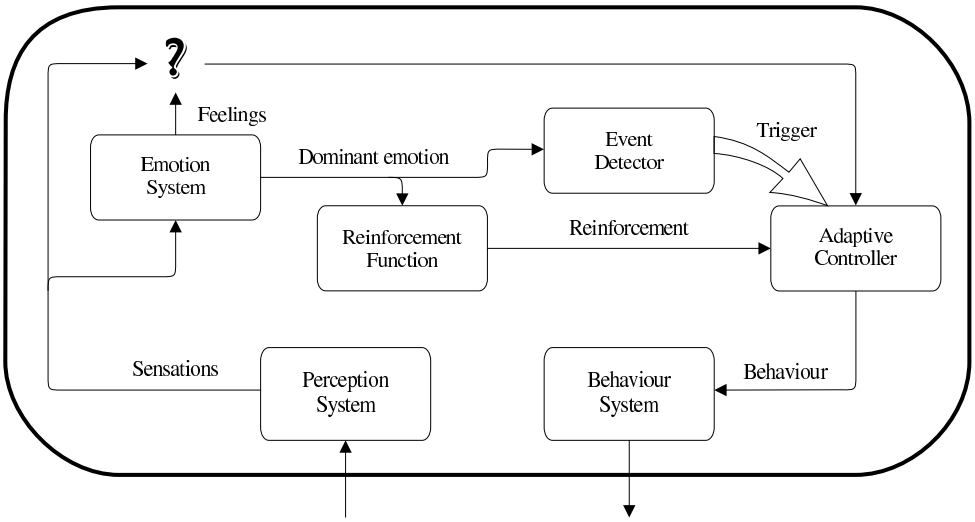

Figure 7: Emotions' influence on perception. In the emotion controller, the feelings define the agent's perception of the world state within the adaptive controller. An alternative non-emotional perception is to use sensations instead of feelings.

used as reinforcement. The learning task consisted of the maximization of positive emotions and minimization of negative emotions.

The emotional controller uses the emotion-dependent reinforcement described in Equation 13, *i.e.* $R_n = R_{e_n}$. The reinforcement magnitude was set to be the intensity of the current dominant emotion or zero if there was no dominant emotion. If the dominant emotion was negative then its positive intensity value would be negated.

$$\forall e \in \mathcal{E}, \quad \text{sign(e)} = \begin{cases} 1 & \text{if } e \text{ is positive} \\ -1 & \text{if } e \text{ is negative} \end{cases} \tag{12}$$

$$R_{e_n} = \begin{cases} 0 & \text{if } \forall e \in \mathcal{E}, I_{e_n} < I_{th_s} \\ I_{e_n} \text{sign}(e) \text{ where } e = \arg\max_{e \in \mathcal{E}}(I_{e_n}) & \text{otherwise} \end{cases} \tag{13}$$

The experiments made with the behavior-based controller to test *whether the emotion-dependent reinforcement was adequate* — Figure 8 —- compared the results obtained using emotion-dependent reinforcement with a more traditional reinforcement based on direct sensations. The sensation-dependent reinforcement was based on the emotion-dependent one but without any of the temporal or lateral side effects introduced by the hormone system. The sensations were used instead of the feelings to calculate the value of each emotion and the highest

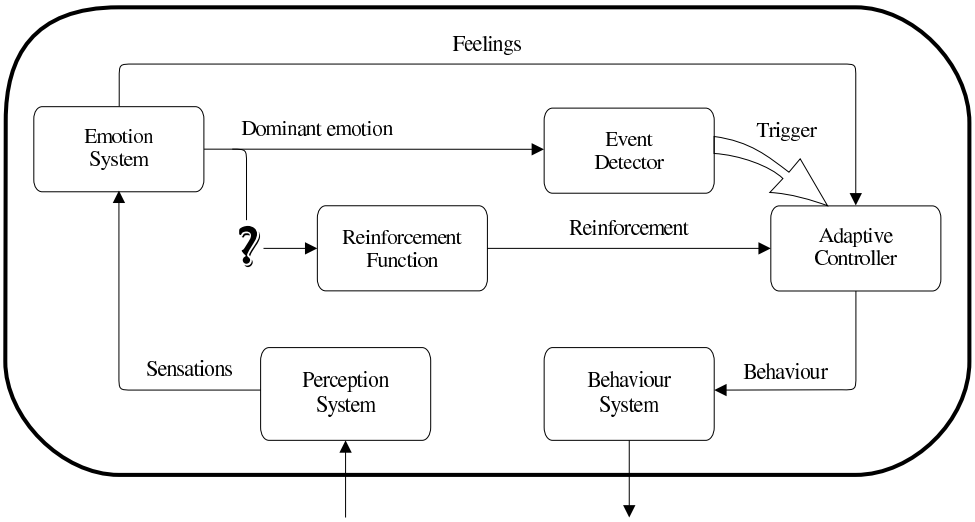

Figure 8: Emotions determining reinforcement value. The reinforcement used by the emotion controller is derived from the dominant emotion. The agent is rewarded if the emotion is positive and punished if the emotion is negative. The value of the rewards or punishments is the emotion intensity. Alternatively, a more conventional non-emotional reinforcement can be used.

of these values was selected to be the reinforcement value. As before, the values associated with negative emotions were negated. Figure 9 illustrates this modification and Equation 14 shows the resulting reinforcement function.

$$R_{s_n} = \left(B_e + \sum_{f \in \mathcal{F}} (C_{ef} S_{f_n}) \right) \text{sign}(e) \quad \text{where} \quad e = \arg\max_{e \in \mathcal{E}} \left(B_e + \sum_{f \in \mathcal{F}} (C_{ef} S_{f_n}) \right) \tag{14}$$

3.4.3 Control Triggering

One problem of behavior coordination which is quite difficult and task dependent is deciding when to change behavior. This is not a problem in traditional reinforcement learning where agents live in grid worlds and state transition is perfectly determined. In robotics, however, agent states change asynchronously in response to internal and external events and actions take variable amounts of time to execute (Mataric, 1994). As a solution to this problem, some researchers extend the duration of the current action according to some domain specific conditions of goal achievement or applicability of the action. Others will interrupt the action when there is a change in the input state (Rodriguez and Muller, 1995; Asada, 1996). Rodriguez

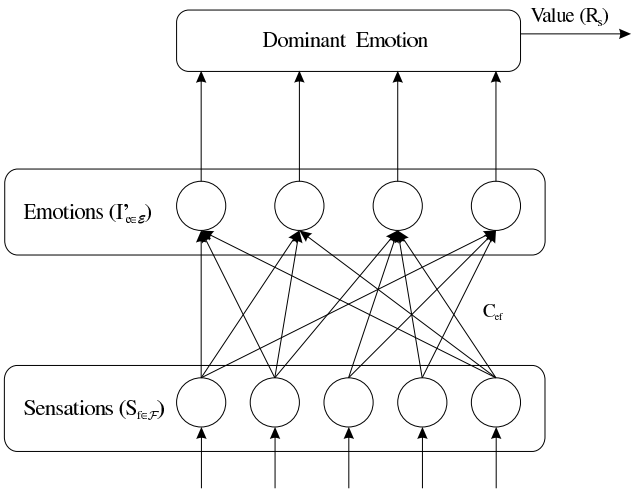

Figure 9: Truncated emotion model used to obtain the sensation-dependent reinforcement.

and Muller (1995) argue that new decisions should only be taken when there is a change in the input state, on the basis that otherwise the choice is uniquely determined by the current state of knowledge. However, this may not be a very straightforward solution when the robot is equipped with multiple continuous sensors that are vulnerable to noise.

In the specific case of the adaptive controller selected, reported experiments involving learning of behavior coordination (Lin, 1993) required the introduction of several simplifications. To start with, each behavior was associated with pre-defined conditions of activation. For example, the behavior of door-passing could only be selected if a door was nearby. Furthermore, the introduction of a persistence rule proved essential for good results. This rule ensured that the same behavior was kept until the goal of the behavior had been achieved or a previously inapplicable behavior became applicable.

In practice, the duration of behaviors must be long enough to allow them to manifest themselves, and short enough so that they do not become inappropriate (due to changing circumstances) long before being interrupted. The ideal would be to know when a significant change has occurred in the environment that makes a re-evaluation necessary. Using emotions to trigger state transition seems reasonable, because emotions can provide a global summarized vision of the environment. Any important change in the environment is liable to be captured by changes in the emotional state.

The selection of this asynchronous triggering mechanism has the added advantage of helping the learning at the level of the neural-networks. When learning on-line by back-propagation,

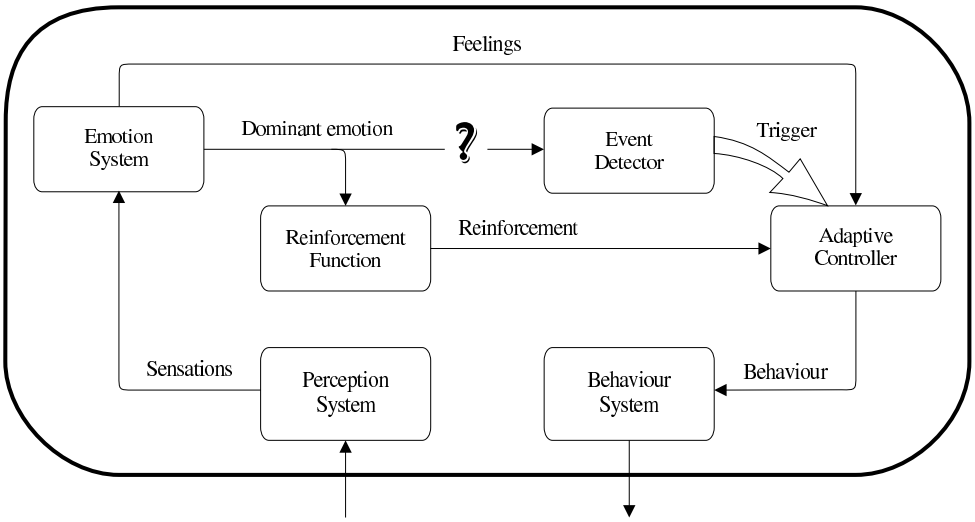

Figure 10: Emotions triggering control. The emotional controller triggers the adaptive controller whenever there are significant changes in the dominant emotion. An alternative non-emotional triggering mechanism which is traditionally used in reinforcement-learning is to trigger the adaptive controller at regular intervals.

the neural networks have a tendency to be overwhelmed by the large quantity of consecutive similar training data and forget the rare relevant experiences. Detecting and using only a few relevant examples for training can help with this problem.

In order to test *whether emotions can successfully be used to trigger state transitions*, *i.e.* determine when behavior selection should be re-evaluated — Figure 10 —, two triggering mechanisms were designed:

Event-triggered — This consisted in having state transitions triggered by the detection of significant changes in the emotional state. From the robot's point of the view, an event occurs whenever there is a significant change in emotional state, as this should reflect a relevant event in the robot-environment interaction. An event is detected whenever:

- there is a change of dominant emotion, including changes between emotional states and neutral emotional states (*i.e.* states with no dominant emotion);
- the current dominant emotion value is statistically different from the values recorded since a state transition was last made, *i.e.* if the difference between the new value and the mean of the previous values exceeds both a small tolerance threshold and ξ times the standard deviation of those previous values, where ξ is a constant set to 2

Table 3: Different mechanisms used by emotional and non-emotional controllers.

Controller	Reinforcement	Perception	Control Triggering
Emotional	Emotion-dependent	Feelings	Event-triggered
Non-emotional perception	Emotion-dependent	Sensations	Event-triggered
Non-emotional reinforcement	Sensation-dependent	Feelings	Event-triggered
Non-emotional triggering	Emotion-dependent	Feelings	Interval-triggered
Non-emotional	Sensation-dependent	Sensations	Interval-triggered

in the experiments (details in Gadanho, 1999);

- A maximum limit of 10 000 steps is reached.

Interval-triggered — A simple alternative to emotion-dependent event detection used for comparison. This consists in triggering the adaptive controller at regular intervals. The success of the learning agent depends on selecting an adequate interval duration. Experiments (Gadanho, 1999) showed that it is important to synchronize the duration of behavior execution with the dynamics of the robot-environment interaction and thus allow compatible time-scales between them. The interval of 35 steps was considered the best suited for the task at hand, given the outcome of pilot experimentation.

3.5 The Experimental Procedure

Five different experiments were carried out to assess the competence of the emotional system as an adaptive controller: one experiment for the emotional controller; one experiment for each of the emotional mechanisms, which consisted in replacing the mechanism by its non-emotional counterpart; and one for a non-emotional controller where all the emotional mechanisms were replaced by the non-emotional counterparts. Table 3 presents the differences between the experiments.

Each experiment consisted in having thirty different robot trials of three million learning steps. In each trial, a new fully recharged robot with all state values reset was placed at a randomly selected starting position. For evaluation purposes, the trial period was divided into sixty smaller periods of fifty thousand steps. For each of these periods the following statistics were taken:

Emotion — mean of the emotion-dependent reinforcement value, which is a measure of how

positive the robot's emotional state is and therefore a good measure of its overall performance;

Energy — mean energy level of the robot;

Distance — mean value of the Euclidean distance d, taken at one hundred steps intervals, between the opposing points of the rectangular extent that contains all the points the robot visited during the last interval,

$$d = \frac{1}{100}\sqrt{(x_{max} - x_{min})^2 + (y_{max} - y_{min})^2}$$

it is a measure of how much distance was covered by the robot;

Collisions — number of collisions;

Events — number of times the adaptive controller was triggered.

The results presented are an average of the different statistics over the several trials with errors representing the 95% confidence interval. The last two statistics were presented as a percentage over the total number of steps in the evaluation period in order to have the experimental results independent of the duration of this period.

3.6 The Results

A summary of the experimental results obtained for all controllers tested is presented in Table 4. The means of the values and their 95% confidence interval obtained in the last half of the trials are presented. Other results showed clearly that all controllers' learning algorithm had converged before the table values were taken.

Experiments showed a small non-significant advantage in performance of the emotional perception, both in terms of emotion and collisions, which means that the robot does well with a biased view of reality but does not demonstrate that emotions can be useful in this domain.

There was also no significant difference between the performance obtained with the two different reinforcements. Again the mean values of emotion and collisions are slightly improved, but results do not clearly show that emotion-reinforcement is better. The results do show that the distance and energy values are higher for the sensation-dependent reinforcement. Nevertheless, this is not considered an improvement in performance over the emotional reinforcement since the robot is only required to maintain these values high and not to maximize them.

Table 4: Summary of the comparison between emotional and non-emotional controllers.

Controller	Emotion	Events (%)	Energy	Collisions (%)	Distance
Emotional	0.24 ± 0.01	0.5 ± 0.0	0.63 ± 0.01	0.6 ± 0.3	1.0 ± 0.2
Non-emotional perception	0.22 ± 0.03	0.5 ± 0.0	0.61 ± 0.03	1.2 ± 0.7	1.0 ± 0.2
Non-emotional reinforcement	0.22 ± 0.02	0.5 ± 0.0	0.70 ± 0.02	1.6 ± 1.1	1.4 ± 0.1
Non-emotional triggering	0.21 ± 0.02	2.9 ± 0.0	0.62 ± 0.04	1.7 ± 0.1	0.9 ± 0.0
Non-emotional	0.21 ± 0.02	2.9 ± 0.0	0.64 ± 0.04	1.4 ± 0.1	0.9 ± 0.0

Experiments showed that the event-triggered mechanism required many fewer control iterations to successfully carry out the task. Furthermore, the controller with non-emotional triggering performed worse in terms of obstacle avoidance.

The results in Figure 11 compare the learning performance of the emotional controller and that of the non-emotional controller. Both controllers are successful in learning the task, managing to keep a high level of energy while minimizing the number of bumps. The most important differences between the emotional and non-emotional controllers are the number of events and the number of collisions. Results indicate that the main factor responsible for the advantage of the emotional controller over the non-emotional controller is the control triggering mechanism.

The emotional controller also has a slight advantage in terms of obstacle avoidance when compared with all the other controllers, suggesting that the temporal synchrony between the different mechanisms might be a beneficial factor. Although the emotional controller suffers from intrinsic delays with respect to the robot-environment interaction due to the emotions' persistence, its perception, reinforcement and triggering mechanism are in perfect synchronism.

3.7 Discussion

Experiments showed that emotions can be used as an attention mechanism at different levels of a reinforcement learning task:

- making more evident the particular aspects of the environment which are relevant for the current emotional state;
- providing a straightforward reinforcement function which attributes value to the different

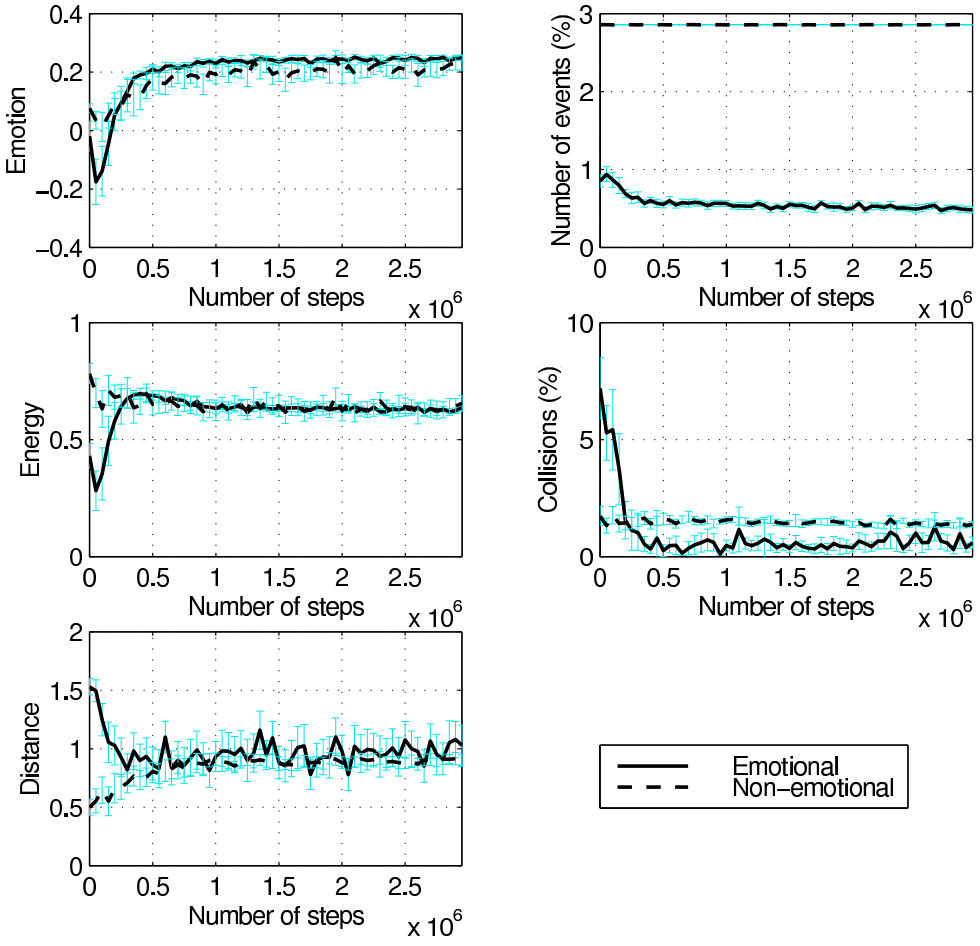

Figure 11: Comparing the emotional with the non-emotional controller. The non-emotional controller does not use emotions in the agent control, but emotions are calculated while the robot is running as a measure of its global performance. Other measures of comparison included the number of times the controller was triggered, its average energy value, the number of collisions with obstacles and the average distance it covered. The error bars represent the 95% confidence interval.

environmental situations, and therefore specifies their relative importance;

- determining the occurrence of the significant changes in the environment that ought to trigger state transition, by looking at sudden changes in the emotional system state.

These were three different mechanisms that worked well experimentally. Each one of them had different levels of performance when compared with alternative methods.

No significant differences were found in using emotion-dependent perception, *i.e.* making the emotionally relevant aspects of the environment more salient, or not. This result might be task dependent but is certainly controller dependent because the adaptive controller used can easily ignore the differences in magnitude of the input values introduced by emotions by compensating for them with changes in the neural-network's weights. A proper assessment of this emotion role would benefit from the employment of a controller equipped with proper mechanisms of attention for input processing, *i.e.* a controller where different weights could be given to the analysis of the different inputs.

The emotion-dependent reinforcement has been tested previously in a similar controller that chooses between primitive actions instead of behaviors (Gadanho and Hallam, 1998). The empirical results obtained for this action-based controller showed that emotion-dependent reinforcement was inadequate and performed much worse than the more traditional sensation-dependent reinforcement. The reason found for this result was that the sensation-dependent reinforcement has the advantage of providing a reinforcement value that reflects the situation as it stands at the moment and not the mixed evaluation of present and recent past situations that the emotion-dependent reinforcement provides. However, experiments made with the behavior-based controller showed no difference between the performance obtained with the two different reinforcements.

This led to the conclusion that the previous results (Gadanho and Hallam, 1998) obtained were controller dependent and that emotions are more appropriate for behavior-level control than to action-based control. Behavior-based control provides more appropriate time scales than action-based control for the use of emotion-dependent reinforcement. If the robot has to select and evaluate a primitive action at each time step the emotions' persistence in time becomes a severe hindrance for their successful use as reinforcement. Moreover, the difference between behavior-based and action-based control is not only restricted to temporal duration. The behaviors themselves are by nature distinct from simple actions. They are not defined by constant motor values, but rather by a simple reactive "goal" that determines the motor values

at each step as a function of the agent's current perception. This allows them enough versatility to run for longer durations of time which is their main advantage for use with emotions. The persistence of emotions over time in natural systems also suggests that they should be related to a higher level of decision-making which does not rely on simple primitive actions but on complex action patterns more suitably expressed at a behavioral level.

The emotion-dependent reinforcement has the characteristic of only depending on one emotion at a time, if any. The reinforcement information that might be provided by emotions other than the dominant emotion is ignored. For example, if the robot is sad and bumps into an obstacle then fear will overcome sadness and only fear will be taken into consideration for reinforcement. This means that reinforcement information will mostly ignore the hunger feeling and will be dominated by the pain feeling. To test whether this is an advantage for the adaptive controller or not, experiments were made that used the reinforcement values provided by all the emotions in the calculation of the reinforcement function at each iteration. The poor performance of this reinforcement function (Gadanho, 1999) supports the view that the non-linearities of the emotion-dependent reinforcement have an active role in its success.

The emotion-dependent event detector allowed drastic cuts in the frequency of triggering of the adaptive controller while maintaining overall performance. This can be particularly advantageous in the case of very time-consuming learning controllers, where each triggering of the controller can result in a significant loss of precious real time. A analysis of the distribution of the emotional states of the emotional controller in general and in the particular situations where events were detected (Gadanho, 1999), showed that the robot control is triggered more often in adverse situations. This has adaptive value, because it focuses the agent's attention on the need to change behavior when the current behavior becomes inappropriate. Furthermore, behaviors that have some immediate goal like avoid-obstacles and seek-light tend to have shorter durations than the wall-following behavior, because they are terminated by the event-triggering mechanism the moment their goal is reached.

Furthermore, a more comprehensive comparison between the two (Gadanho and Hallam, 2001) reveals that an event-triggered controller is also a better learner than the interval-triggered controller: it has the advantage of being both a more time-efficient learner and able to master more difficult tasks.

Adding the right amount of behavioral persistence is a difficult problem that has been found before in the design of artificial creatures (Mahadevan and Connell, 1992; Lin, 1993; Mataric, 1994; Blumberg, 1996). In fact, when using a learning architecture very similar to the one used

by the non-emotional controller presented here, Lin (1992) had to recur to severe simplifications of the behavior-coordination learning task. These simplifications consisted on having behaviors associated with pre-defined conditions of activation and only interrupting a behavior once it had reached its goal or an inapplicable behavior had become applicable. The emotion-controller proposed here offers a less task-dependent solution to solve the behavior persistence problem.

4 Conclusion

The emotional controller proved to be more successful than an equivalent non-emotional controller where the emotional mechanisms were replaced by their non-emotional counterparts. The triggering mechanism was strongly responsible for the difference in performance between the two, but the joint use of all three mechanisms also seems to provide an advantage.

The use of emotions as an abstraction has the advantage of allowing different components of reinforcement learning to be brought together under the same construct. This was found helpful for two reasons:

- the synchrony and coherence between the different components achieved by this unified solution represented a slight enhancement of the agent's performance;
- the design of the different components was simplified to the design of a single construct, the robot's emotions;

Furthermore, the use of emotions provided a new perspective over these different task-dependent components of a reinforcement-learning framework. This resulted in the introduction of innovative mechanisms that were tested in the robot experiments. The most important innovations are in terms of the reinforcement function and the specification of state transition:

- a multi-dimensional reinforcement function that takes into consideration the different problems faced by the robot with variable degrees of attention dependent on the robot's current priorities;
- a simplified definition of state transition based on detection of significant events captured by variation in the reinforcement function value.

The emotional system selects between different reinforcement functions according to the context of the world, *i.e.* it might choose to ignore other problems that exist when faced with

a more important one. For example, the reinforcement function might not punish the robot for its collision with an obstacle and instead reinforce it for successfully extracting energy from a light. The attribution by the reinforcement function of variable degrees of attention to each of the different problems might be taken to be a source of confusion for the learning process. However, experiments showed that instead of confusing the learning process, this was actually advantageous and that the learning algorithm was exploiting the non-linearities of the reinforcement function.

The presented event detection mechanism also profits from the novel structure of the reinforcement function. Apart from providing an absolute reinforcement value that varies with the robot's situation, the developed reinforcement function based on emotion also differentiates and prioritizes the different problems faced by the robot. This added information allows the detection of events when there is a difference in type of dominant problem and not just in problem degree.

The work reported here was a first study in the interaction of emotions and control in a robot agent which could benefit from several extentions. The emotion model could be refined to have more complex emotions and the capability to learn new emotion associations so that the learning agent could be able to master more difficult tasks. Furthermore, the interaction between emotions and control could be extended to capture other functionalities found in natural emotions like influences on memory, reasoning, behavioral tendencies and modulation of behavioral responses. The proper evaluation of these influences and a better understanding of the influences already studied would probably benefit from a more elaborate control architecture.

Acknowledgments

The first author was supported by a scholarship from the Portuguese program PRAXIS XXI for doing the work reported. We thank Gillian Hayes, Rosalind Picard and the reviewers for helpful comments on earlier versions of this material. Facilities for this work were provided by the University of Edinburgh.

References

Albus, J. S. (1990). The role of world modeling and value judgment in perception. In Meystel, A., Herath, J., and Gray, S., editors, *Proceedings of the* 5^{th} *IEEE International Symposium on Intelligent Control*. Los Alamitos, CA: IEEE Computer Society Press.

Araujo, A. F. R. (1994). *Memory, emotions & neural networks*. PhD thesis, Sussex University.

Asada, M. (1996). An agent and an environment: A view on body scheme. In Tani, J. and Asada, M., editors, *Proceedings of the 1996 IROS Workshop on Towards real autonomy*, pages 19–24, Senri Life Science Center, Osaka, Japan.

Aubé, M. (1998). A commitment theory of emotions. In *AAAI Fall Symposium on Emotional and Intelligent: The tangled knot of cognition*, Technical Report FS-98-03, pages 13–18. Menlo Park, CA: AAAI Press.

Bates, J. (1994). The role of emotions in believable agents. Technical Report CMU-CS-94-136, Carnegie Mellon University, School of Computer Science.

Bates, J., Loyall, A. B., and Reilly, W. S. (1992a). An architecture for action, emotion, and social behavior. In Castelfranchi, C. and Wemer, E., editors, *Artificial social systems: Fourth European workshop on modeling autonomous agents in a multi-agent world*, pages 55–68. Heidelberg, Germany: Springer-Verlag. Also available as technical report CMU-CS-92-144, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Bates, J., Loyall, A. B., and Reilly, W. S. (1992b). Integrating reactivity, goals, and emotion in a broad agent. Technical report CMU-CS-92-142, Carnegie Mellon University, School of Computer Science, Pittsburgh, PA. Also appeared in the Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, Bloomington, Indiana, July 1992.

Beaudoin, L. and Sloman, A. (1993). A study of motive processing and attention. In A.Sloman, D.Hogg, G.Humphreys, Ramsay, A., and Partridge, D., editors, *Proceedings of AISB93*, pages 229–238. IOS Press, Oxford.

Bechara, A., Damasio, H., Tranel, D., and Damásio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275:1293–1295.

Beck, R. C. (1983). *Motivation: Theories and principles*. Prentice Hall Inc., Englewood Cliffs, New Jersey, second edition.

Blaney, P. (1986). Affect and memory. *Psychological bulletin*, 99(2):229–246.

Blumberg, B. (1996). *Old Tricks, New Dogs: Ethology and interactive creatures*. PhD thesis, MIT.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36(2):129–148.

Bower, G. H. and Cohen, P. R. (1982). Emotional influences in memory and thinking: Data and theory. In Clark, M. S. and Fiske, S. T., editors, *Affect and cognition — The seventeenth annual Carnegie symposium on cognition*, pages 291–331. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bozinovski, S. (1982). A self-learning system using secondary reinforcement. In Trappl, R., editor, *Cybernetics and Systems*, pages 397–402. Elsevier Science Publishers, North Holland.

Breazeal, C. (1998). Early experiments using motivations to regulate human-robot interaction. In Canamero, D., editor, *AAAI Fall Symposium on Emotional and Intelligent: The tangled knot of cognition*, Technical Report FS-98-03, pages 31–36. Menlo Park, CA: AAAI Press.

Cañamero, D. (1997). Modeling motivations and emotions as a basis for intelligent behavior. In *Proceedings of the First International Symposium on Autonomous Agents, AA'97*, Marina del Rey, CA. The ACM Press.

Cañamero, D. (1998). Issues in the design of emotional agents. In Canamero, D., editor, *AAAI Fall Symposium on Emotional and Intelligent: The tangled knot of cognition*, Technical Report FS-98-03, pages 49–54. Menlo Park, CA: AAAI Press.

Charland, L. C. (1995). Emotion as a natural kind: towards a computational foundation for emotion theory. *Philosophical psychology*, 8(1):59–84.

Cytowic, R. E. (1993). *The man who tasted shapes*. Abacus, London.

Damasio, A. R. (1994). *Descartes' error — Emotion, reason and human brain*. Picador, London.

Darwin, C. (1965). *The expression of the emotions in man and animals*. The University of Chicago Press. Original work published in 1872.

De Sousa, R. (1987). *The rationality of emotion*. The MIT Press.

Derryberry, D. and Tucker, D. M. (1994). Motivating the focus of attention. In Niedenthal, P. M. and Kitayama, S., editors, *The heart's eye*, pages 167–196. Academic Press, New York.

Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. The MIT Press.

Dyer, M. G. (1987). Emotions and their computations: Three computer models. *Cognition and emotion*, 1(3):323–347.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3/4):169–200.

Foliot, G. and Michel, O. (1998). Learning object significance with an emotion based process. In *SAB'98 Workshop on Grounding Emotions in Adaptive Systems*, pages 25–30, Zurich, Switzerland.

Frijda, N. H. and Swagerman, J. (1987). Can computers feel? theory and design of an emotional system. *Cognition and Emotion*, 1(3):235–257.

Gadanho, S. C. (1999). *Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions*. PhD thesis, University of Edinburgh.

Gadanho, S. C. and Hallam, J. (1998). Emotion-driven learning for animat control. In Pfeifer, R., Blumberg, B., Meyer, J.-A., and Wilson, S. W., editors, *From animals to animats 5 — Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*, pages 354–359. The MIT Press.

Gadanho, S. C. and Hallam, J. (2001). Emotion-triggered learning in autonomous robot control. *Cybernetics and Systems — Special Issue: Grounding emotions in adaptive systems*, 32(5):531–559.

Goleman, D. (1995). *Emotional Intelligence*. Bloomsbury Publishing, London.

Grossberg, S. and Gutowski, W. (1987). Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review*, 94(3):300–318.

Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Addison-Wesley Publishing Company, Reading, MA.

Izard, C. E. (1993). Four systems for emotion activation: Cognitive and noncognitive processes. *Psychological Review*, 100(1):68–90.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285.

Kaiser, S. and Wehrle, T. (1996). Situated emotional problem solving in interactive computergames. In Frijda, N. H., editor, *Proceedings of the* 8^{th} *Conference of the International Society for Research on Emotions, ISRE'96*, pages 276–280, Storrs, CT. ISRE Publications.

Kitano, H. (1995). A model for hormonal modulation of learning. In Mellish, C. S., editor, *IJCAI-95 — Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, volume 1, pages 532–538, Montreal, Canada. Morgan Kaufmann.

Klein, J. T. (1996). Computer response to frustration. Msc in Media Arts and Sciences, The MIT School of Architecture and Planning.

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, 37:1019–1024.

LeDoux, J. E. (1998). *The Emotional Brain*. Phoenix, London.

Lin, L.-J. (1992). Self-improving reactive agents based on reinforcement learning planning and teaching. *Machine Learning*, 8:293–321.

Lin, L.-J. (1993). *Reinforcement learning for robots using neural networks*. PhD thesis, Carnegie Mellon University. Technical report CMU-CS-93-103.

Mahadevan, S. and Connell, J. (1992). Automatic programming of behavior-based robots using reinforcement learning. *Artificial intelligence*, 55:311–365.

Mataric, M. J. (1994). Reward functions for accelerated learning. In Cohen, W. W. and Hirsh, H., editors, *Machine Learning: Proceedings of the Eleventh International Conference*, pages 181–189, San Francisco, CA. Morgan Kaufmann Publishers.

McCauley, L. and Franklin, S. (1998). An architecture for emotion. In Canamero, D., editor, *AAAI Fall Symposium on Emotional and Intelligent: The tangled knot of cognition*, Technical Report FS-98-03, pages 122–127. Menlo Park, CA: AAAI Press.

Michaud, F., Lachiver, G., and Dinh, C. T. L. (1996). A new control architecture combining reactivity, planning, deliberation and motivation for a situated autonomous agent. In Maes, P., Mataric, M. J., Meyer, J.-A., Pollack, J., and Wilson, S. W., editors, *From animals to animats 4 - Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pages 245–254. The MIT Press.

Michel, O. (1996). *Khepera Simulator* package version 2.0: Freeware mobile robot simulator written at the University of Nice Sophia–Antipolis. Downloadable from the World Wide Web at `http://wwwi3s.unice.fr/~om/khep-sim.html`.

Minsky, M. (1986). *The Society of Mind*. Simon and Schuster, New York.

Morignot, P. and Hayes-Roth, B. (1995). Why does an agent act? In Cox, M. and Freed, M., editors, *AAAI Spring Symposium on Representing Mental States and Mechanisms*, pages 97–101, Menlo Park, CA. AAAI Press. Also available as report KSL 94-76 of Stanford University.

Mowrer, O. H. (1960). *Learning theory and behavior*. John Wiley & Sons, Inc., New York.

Niedenthal, P. M., Setterlund, M. B., and Jones, D. E. (1994). Emotional organization of perceptual memory. In Niedenthal, P. M. and Kitayama, S., editors, *The heart's eye*, pages 87–113. Academic Press, New York.

Ortony, A., Clore, G. L., and Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press, Cambridge, England.

Panksepp, J. (1995). The emotional brain and biological psychiatry. *Advances in Biological Psychiatry*, 1:263–288.

Pfeifer, R. (1982). Cognition and emotion: An information process approach. CIP working paper 436, Department of Psychology, Carnegie-Mellon University.

Pfeifer, R. and Nicholas, D. W. (1985). Toward computational models of emotion. In Steels, L. and Campbell, J. A., editors, *Progress in Artificial Intelligence*, pages 184–192. Ellis Horwood, Chichester, U.K.

Picard, R. (1997). *Affective Computing*. The MIT Press.

Power, M. and Dalgleish, T. (1997). *Cognition and Emotion*. Psychology Press, Philadelphia.

Rodriguez, M. and Muller, J.-P. (1995). Towards autonomous cognitive animats. In Morán, F., Moreno, A., Merelo, J., and Chacón, P., editors, *Advances in artificial life — Proceedings of the Third European Conference on Artificial Life*, Lecture Notes in Artificial Intelligence Volume 929, Berlin, Germany. Springer-Verlag.

Seif El-Nasr, M., Ioerger, T. R., and Yen, J. (1998). Learning and emotional intelligence in agents. In Canamero, D., editor, *AAAI Fall Symposium on Emotional and Intelligent: The*

tangled knot of cognition, Technical Report FS-98-03, pages 150–155. Menlo Park, CA: AAAI Press.

Shaver, P., Schwartz, J., Kirson, D., and O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of personality and social psychology*, 52(6):1061–1086.

Shibata, T., Ohkawa, K., and Tanie, K. (1996). Spontaneous behavior of robots for cooperation — Emotionally intelligent robot system. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 2426–2431, Tokyo, Japan.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, 74:29–39.

Sloman, A., Beaudoin, L., and Wright, I. (1994). Computational modeling of motive-management processes. In Frijda, N., editor, *Proceedings of the Conference of the International Society for Research in Emotions*, pages 344–348, Cambridge. ISRE Publications. Poster.

Sloman, A. and Croucher, M. (1981). Why robots will have emotions. In *IJCAI'81 — Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, pages 2369–71. Cognitive Science Research Paper 176, Sussex University.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning*. The MIT Press.

Toda, M. (1993). The urge theory of emotion and cognition: Chapter 1 Emotions and urges. SCCS technical report 93-1-01, Chuyko University. English version of the book "Kanjo (Emotion)".

Toda, M. (1994). Emotion, society and the versatile architecture. SCCS technical report 94-1-02, Chuyko University.

Tomkins, S. S. (1984). Affect theory. In Scherer, K. R. and Ekman, P., editors, *Approaches to Emotion*. Lawrence Erlbaum, London.

Velásquez, J. D. (1998). A computational framework for emotion-based control. In *SAB'98 Workshop on Grounding Emotions in Adaptive Systems*, pages 62–67, Zurich, Switzerland.

Ventura, R., Custódio, L., and Pinto-Ferreira, C. (1998). Emotions — the missing link? In Canamero, D., editor, *AAAI Fall Symposium on Emotional and Intelligent: The tangled knot of cognition*, Technical Report FS-98-03, pages 170–175, Menlo Park, California. Menlo Park, CA: AAAI Press.

Watkins, C. (1989). *Learning from delayed rewards*. PhD thesis, King's College, Cambridge.

Wehrle, T. (1998). Motivations behind modeling emotions agents: Whose emotions does your robot have? In *SAB'98 Workshop on Grounding Emotions in Adaptive Systems*, pages 71–76, Zurich, Switzerland.

Wilson, S. W. (1991). The animat path to AI. In Meyer, J.-A. and Wilson, S. W., editors, *From animals to animats — Proceedings of the First Conference on Simulation of Adaptive Behavior*, pages 15–21. The MIT Press.

Wright, I. (1996). Reinforcement learning and animat emotions. In Maes, P., Mataric, M. J., Meyer, J.-A., Pollack, J., and Wilson, S. W., editors, *From animals to animats 4 — Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pages 273–281. The MIT Press.

Zajonc, R. B. (1984). On primacy of affect. In Scherer, K. R. and Ekman, P., editors, *Approaches to Emotion*. Lawrence Erlbaum, London.