The Role of Emotions:

Exploring Autonomy Mechanisms in Mobile Robotics

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

Revised and extended version of DAI Research Paper No. 851.

Abstract

Autonomy is considered to be an important property for a robot to have, yet this is not a well explored issue. This is particularly the case when one requires the meaning of autonomy to include self-motivation, instead of mere automaticity. The fact that emotions are seen as essential to human reasoning suggests that they might play an important role in achieving the necessary self-motivation.

The purpose of the work reported here is to know if it is advantageous to have emotional robots in terms of their final behavior's autonomy. If this is the case, what should the role of emotions be?

An emotion model was developed under the animat philosophy, by building bottom-up a biologically inspired complete agent by synthesis. Experimental work was done in a realistic mobile robot simulator in order to assess the usefulness of the emotions in the robot's adaptive control. A Q-learning algorithm implemented with neural networks was used to control a robot in a simple animal-like surviving task. Experiments revealed that having emotions providing a context evaluation to be used as a reinforcement value is not very helpful. A reinforcement based on simple raw sensations is much more successful. Extensive experiments were done to understand why. Results are reported and discussed. Finally, some alternative uses for emotions are considered and tested.

Key Words: robotics; autonomy; emotions; reinforcement learning

1 Introduction

The criteria used to define whether a robot is autonomous or not are far from well established. In general, in the literature, simply requiring that the robot does its task without human intervention is sufficient. However, the root meaning of autonomy suggests an alternative definition with stronger requirements: namely, that a truly autonomous robot should also develop the laws that rule its behavior.

> **Autonomous** — f. Gr. *αὐτόνομος* making or having one's own laws, independent (f. *αὐτο-* self, own + *νόμος* law) + -ous (The Oxford English dictionary, 1989)

Considering that such strong requirements can be made, stating that a robot is autonomous just because it performs its task without any external help is too weak a claim. Nevertheless, such an ability still undeniably contributes towards extra independence. Hence, this will be considered a weak form of autonomy — autonomy should not be defined as an all-or-nothing property (Boden, 1993). Intermediate degrees of autonomy between being fully autonomous and not autonomous at all should be allowed. For that reason, the definition used in this document will be one of degree, *i.e.* the extent to which the previous definitions are fulfilled.

Although there is little significant experimental work done on the subject, there is much argument in favor of research with truly autonomous systems and simple experiments have been made since the beginnings of AI (Grey Walter, 1950, 1951). Having autonomous agents seems to be a clear advantage in several diverse fields, like for example robotics, animal robotics[1] (McFarland, 1994), agents theory (Ferguson, 1992) and interactive virtual environments (Blumberg, 1995). Although, arguments against having fully autonomous agents are easy to find (in general, the robot is supposed to do something useful for its owner and not whatever it wants), it is generally accepted that it is beneficial to have autonomy, at least to a certain extent which is still far from being achieved in today's systems.

In the field of robotics, the way to accomplish strong autonomy is by developing an adaptive controller that will improve its performance when interacting with its environment by unsupervised learning. Such improvement of performance must be and is always grounded in some kind of value scheme defined by the designer. Whether this scheme involves reinforcement values, instincts, goals or heuristic rules the aim is always the same: to give the robot some sense of *directionality* in its learning task. It is important that the autonomous robot has some sort

[1] Modeling of animal behavior using robots.

of self-motivation that allows it to establish its learning goals without the aid of an external teacher.

The fact that emotions are considered to be essential to human reasoning suggests that they might play an important role in achieving the self-motivation necessary to having strong autonomy. Experimental work was done in order to establish if there is an advantage in having emotions as the source of self-motivation in autonomous robot. Emotions have already been considered as a fundamental source of motivation in psychology (*e.g.*, Beck, 1983) and it has been argued before in the field of Artificial Intelligence (Toda, 1994) that emotions are the ultimate source of intelligence and might provide robots with the autonomy they need.

In robotics, emotions are often used in a social role as a sort of communication mechanism that allows the robot to report its internal state to others (*e.g.* its level of task achievement in Shibata et al., 1996) or makes it capable of generating empathy emotions in people (Bates, 1994), by creating an illusion of life in a believable character. This kind of external demonstration of emotions is very important in social interaction and is of great adaptive advantage (Darwin, 1965) but is not the focus of the present research. The experiments done focus on how to use emotions in the control of the robot, and in particular in its adaptation to the environment and not as modulating activity in a fixed controller (Bates et al., 1994). In particular, the utility of different roles of emotions in terms of the autonomy of the robot's final behavior was explored.

The work was done under an animat philosophy, by building bottom-up a biologically inspired complete agent by synthesis.

As the system is not equipped with self-awareness, there is no reason to claim that the robot feels emotions in the sense humans do. The robot has a weaker form of 'emotions', functionally equivalent but without the awareness. Yet, language will be used that might, implicitly, attribute emotional feelings to the robot. This kind of language is used here because it is more practical and concise to describe the robot's final behavior in terms of intentions (Malcolm, 1993).

An emotional model based on a hormone system was developed (section 2). The emotion system was used as a source of context value in a reinforcement learning task (section 3 describes the adaptive controller). Experiments made used a simulated Khepera robot in an animal-like adaptation task (section 4 gives a full description of the experimental setup). The results (section 5) showed that emotions do not provide as good a reinforcement as expected, but can be used successfully in the modulation of learning (section 6 summaries the conclusions reached).

2 Emotions

It is well established that emotions play an important role in reasoning, even in terms of basic mechanisms like memory and remembering (Schwartz and Reisberg, 1991), although most neurologists have usually given them a more disruptive than helpful function.

People like to have a Cartesian approach to life in thinking that all their reasoning is purely logical. In fact, individuals do not always make rational choices (Grossberg and Gutowski, 1987) and pure logical reasoning shows serious faults when used to model human reasoning in the field of Artificial Intelligence (Dreyfus, 1992). Recent neurophysiological research suggests that our thinking is not so detached and ungrounded. With the help of the emotions, the feelings provided by our body play an important role in reasoning. This is the central claim of the somatic-marker hypothesis[2] (Damásio, 1994).

Damásio makes a clear distinction between the concepts of emotion and feeling that will be used in the current work. Feeling designates the process of monitoring the body. Feelings offer us the cognition of our visceral and musculoskeletal state. Emotion is a combination of a mental evaluative process with dispositional responses to that process, mostly toward the body proper but also toward the brain itself. Emotions all generate feelings, but only some feelings generate emotions. If feelings are associated with emotions then the body signals will move from the background to the foreground of our attention.

Somatic markers are special instances of body feelings (visceral and non-visceral sensations) generated by emotions, that are acquired by experience based on internal preference systems and external events and which help to predict future outcomes of certain scenarios. They will force attention on the negative or positive outcome of certain options, that can be immediately defeated leaving fewer alternatives or can be immediately followed. This way, the somatic markers provide humans with a reasoning system that is free from many of the faults of formal logic, namely the need for much computational and memory power for having every option thoroughly evaluated.

Goleman (1995) has shown that the power of the emotional mind in everyday decision making is greater than the the power of the rational mind. In other words, he argues that it is more advantageous for humans to have a good emotional development than a high intelligence quotient.

Inspired by the ideas that have been presented about emotions, a model was was developed

[2]Marker because it marks an internal image and somatic because it is the body that does it.

that will be described on the next section.

2.1 Model

Most AI models are based on an analytic approach (*e.g.*, Sloman et al., 1994; Pfeifer, 1982; Pfeifer and Nicholas, 1985; Bates et al., 1992) 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 this kind of approach, a bottom-up approach (based on the animat approach, see Wilson, 1991) was followed here. In both ontological development and evolution the full richness of emotions is only achieved at a final stage. In the first stages of these processes only certain basic emotions are present and, later, other more complex emotions develop on top of these.

The model that was developed — figure 1 — was based on four basic emotions: Happiness, Sadness, Fear and Anger. These emotions were selected because they are the most universal emotions along with Disgust (Damásio, 1994) and 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. Yet, if, for example, toxic food was added to the environment, disgust would become useful to keep the robot away from it.

*** Figure 1 goes here ***

The emotions chosen are also usually included in the definitions of primary emotions[3] (*e.g.*, Shaver et al., 1987; Goleman, 1995), which is a good indicator of their relevance and need. There are other emotions, like love and hate, which some authors like to suggest as primary emotions, that were not included because they do not seem very basic (there are even arguments against considering them as emotions, *e.g.* Sloman, 1987) and the present work does not have, for the moment, any social aims.

It is the purpose of this model to have an emotion system that provides the decision making process of the robot with a value judgment based on what the robot feels. This model determines the intensity of each emotion based on the robot's current internal feelings. These feelings are: Hunger, Pain, Temperature, Restlessness and Eating. Each emotion is defined by a set of

[3]It should be noticed that the paradigm of primary emotions is not undisputed, yet most of the arguments against it are marginal to the present usage. These arguments refer to the plausibility of translating all emotions in terms of graduations of primary emotions. The point here is that these emotions are more universal and fundamental and therefore more adequate to low reasoning animats in a simplified environment.

constant feeling dependencies and a bias (see table 1). The values of the dependencies were carefully chosen to provide adequate emotions for the possible body states. For example, the sadness intensity will be high if hunger and restlessness are high and the robot is not eating.

*** Table 1 goes here ***

Furthermore, based on what was suggested by Damásio (1994), the emotion state should also influence the way the robot feels. In general, the body reactions aroused by emotions also give rise to the emotions that create them. In this model each emotion tries to influence the body state in such a way that the result body state matches the state that gives rise to that particular emotion.

When an emotion is active, *i.e.* its intensity value is significantly large, then it will influence the body through a hormone system, by producing appropriate hormones.

The hormone system in the model is a very simplified one. It consists in having one hormone associated with each feeling. A feeling intensity is not a value directly obtained from the value of the body sensation that gives rise to it, but from the sum of the sensation and hormone value. The hormone values can be (positively or negatively) high enough to totally hide the real sensations from the robot's perception of its body.

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 provides a sort of fight between the emotions to gain control over the body which is ultimately what selects which emotion will be dominant. On the other hand, what the robot feels — its body image — is not only dependent on its sensations but is also dependent on its emotional state.

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 has given rise to it is gone, another of the characteristic features of emotions.

The dominant emotion is the emotion with the highest intensity, unless no emotion intensity exceeds a selection threshold. In this case, there will not be a dominant emotion and emotional state will be neutral. Emotions were divided into two categories: positive and negative. The ones that are considered "good" 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 system by considering the intensity of the current dominant emotion and whether it is positive or negative.

In summary, the model of emotions described (see appendix A.1 for details) provides not only a value judgment about the current situation, but also influences the body perception. In order to evaluate the role of emotions in reasoning, this value should be used for the actual control of the robot, determining its behavior (Albus, 1990; Wright, 1996; Moffat et al., 1993). The next section describes an adaptive controller that will make use of such a value.

3 Adaptive Action Selection Controller

The role of the adaptive controller (see figure 2) is to produce actions so that it maximizes the evaluation received, a value that is expected to vary between -1 and 1. To achieve this purpose the developed controller can select one of six possible discrete actions, taking into account the current sensor readings and robot feelings, and the previously received evaluations.

*** Figure 2 goes here ***

It might seem that having such a small and non-continuous space of possible actions is inconsistent with the main objective, *i.e.* achieving autonomy, and in fact the solution found is not totally satisfactory in this respect. Yet, more importantly than allowing the robot more freedom of movement, this solution allows the robot freedom of choice. The action the robot takes at each point is not predefined by the designer, but is selected by the robot according to what it has learned so far. Furthermore, with the set of actions provided, the robot can still have a large repertoire of different behaviors by sequentially selecting the appropriate action at each time step.

The controller developed has two separate modules:

Associative Memory Module — This plastic module associates the sensor readings and feelings with the current expected value of each one of the actions that the robot can take.

Action Selection Module — Based on the information provided by the previous module, this module makes a stochastic selection of the action to take at each step.

3.1 Associative Memory Module

The associative memory consists of six neural networks that try to predict the outcome of selecting each one of the available actions (see appendix B.1 for details):

- slow move forward;
- turn left;
- turn right;
- fast move forward;
- stop;
- move slowly backwards with a slight twist to the right.

Each network is a three layer feed-forward network with:

- 22 input units: one for each distance and light sensor (IR values were rescaled to $[0, 1]$), one for each feeling and a bias;
- 5 hidden units;
- 1 output unit that represents the expected outcome if the action associated with this net is selected in the present situation.

The activation functions used were the hyperbolic tangent in the hidden layer and the identity function in the output layer. The weights between the hidden layer and the output layer are initialized with random values and weights between the input layer and the hidden layer are set to zero. This way all the networks will provide an initial neutral evaluation. The learning algorithm used to train the networks was back-propagation (see, for example, Hertz et al., 1991, for a full description).

First attempts that used the networks to associate the received evaluation with the network inputs were a failure. Learning from delayed rewards with Q-learning (Watkins, 1989) proved to be much more successful. An implementation very similar to the one reported by Lin (1992) was used. The networks were used to learn utility functions that model $util(x, a)$:

$$util(x, a) = R + \gamma eval(y) \tag{1}$$

The value R is the immediate reinforcement. The function $eval(y)$ is the expected cumulative discounted reinforcement starting from state y matched by doing action a in state x. The discount factor (γ) was set to 0.9. For each iteration, the target value u will be given to the network whose action was used in the previous iteration:

$$u = R + \gamma Max\{util(y,k) \mid k \in actions\} \tag{2}$$

After an action a has been evaluated its network state (*i.e.* the value of each one of its units) for situation x is saved. The new utility value of each action k for the new situation y is calculated. The maximum is obtained and used in the previous formula to update the weights of the network associated with action a and the previous situation. Just before learning by back-propagation takes place, the network's saved state for situation x is restored. After the learning has taken place, the utility value for action a is recalculated.

3.2 Action Selection Module

The utility values provided by the associative memory are used for the stochastic selection of the next action to take. The higher the value provided by the associated net, the higher the probability of an action to be selected.

The function used to calculate the probability of each action is based on the Boltzmann-Gibbs distribution. For a temperature T, the probability of selecting action a is:

$$P(a) = \frac{e^{\frac{v_a}{T}}}{\sum_{i=1}^{6} e^{\frac{v_i}{T}}} \qquad v_i = \text{Value of action } i \tag{3}$$

The selection of a new action is not made every cycle; there is a certain inertia of the current action that is directly correlated with its probability (see appendix B.2 for details). The reason for this is to have a more coherent behavior. Otherwise, the robot would mainly tremble, because it would be selecting different actions all the time.

4 Experimental setup

4.1 The robot

All the experiments were carried out in a Khepera simulated robot (Michel, 1996) within the same environment (see figure 3). This is a closed environment with some walls and three lights surrounded by bricks[4].

*** Figure 3 goes here ***

The ultimate goal of the research reported in this document is to develop a fully autonomous real robot. This was one reason why self-sufficiency was considered a useful property to include in the system. Another reason for this choice is that it is easier to think of emotions in the context of an animal-like creature with self-maintenance needs.

Simulated feeding needs were therefore 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 light. More exactly, the amount of energy that the robot acquires at each step depends on whether enough light is being received by the two front sensors and on how much light is being received by those sensors.

The robot feelings are based on the following sensations from its body (see A.2 for details):

Eating — is based on the energy the robot is acquiring at the moment.

Hunger — is directly related to its energy deficit.

Temperature — depends upon the usage of the motors. The real robot's velocity does not matter; as long as high velocity is being demanded from the motors, the temperature will rise.

Restlessness — increases if the robot does not move.

Pain — is active if the robot is bumping into obstacles.

In summary, apart from feeding itself, the robot is supposed to wander in order to avoid getting restless and avoid walls because bumping into the walls causes it pain.

4.2 The experimental procedure

All experiments consisted in having the robot learn for two thousand steps followed by an evaluation of its performance for another two thousand steps, with learning turned off. In total, the robot takes one hundred and twenty thousand learning steps and sixty one evaluation tests, one test after each learning period plus one extra test before any learning takes place.

[4]The lights had to be surrounded by bricks to avoid the robot becoming permanently stuck in their concavities. The lights can still be perceived by the robot as the bricks are transparent to light.

Tests were made transparent to the experiment. When continuing with its learning the robot's state is restored to the state just after its last learning step and previous to the test.

The robot's evaluation was based on the reinforcement values it received in its testing period. There were two evaluations, each one based on a different reinforcement function. One was the mean of the emotion-dependent reinforcements and the other was the mean of the sensation-dependent reinforcements obtained during its test period (sections 5.2 and 5.3 give a detailed description of the reinforcement functions). The higher the value, the better is the evaluation. For each experiment, this procedure was performed fifty times so that an average of the evaluations over several trials could be obtained. Each trial had a new robot with all state values reset placed in a randomly selected starting position. There are twenty possible starting positions (shown in figure 4), that were chosen to maximize the differences in starting conditions, but were otherwise arbitrary.

*** Figure 4 goes here ***

The experimental data that are shown in the graphs are the means of the reinforcement values obtained during the sixty-one testing phases in each of the fifty runs. The error bars show the 95% confidence interval.

As any adaptive robot should, the robot was designed to learn continuously and therefore it might seem strange to have a distinction between a learning phase and a performance phase. The idea behind having a testing phase with no learning is that each step in a test represents a snapshot situation that the controller has to deal with. This way each test provides two thousand "individual" evaluations of the same controller.

5 Experimental results

5.1 No Learning

The experimental setup provides temporal constraints that add complexity to the problem. Figure 5 shows how the robot would perform in time without any adaptation, just with a random action selection controller. The left graph of figure 5 shows the values of a reinforcement function based on direct sensations and the right graph shows a reinforcement function based on emotions. The immediate sensations provide a steady and gradual decrease of reinforcement over time, that reflects the decrease in energy level, while emotions suffer a much more significant drop right at the start due to the recurrent nature of the emotions model.

*** Figure 5 goes here ***

The robot's performance in its adaptation task is highly constrained by the time it takes. If it takes too long, the reinforcement will lose meaning before any adaptation can be achieved. The reason for this is that if the robot does not learn to feed itself, it will get increasingly hungry. It will eventually arrive to a state where it is totally hopeless, because it will keep getting low reinforcement on account of its hunger, independently of what it does.

5.2 Emotion-dependent reinforcement values

The most usual role attributed to emotions in the functionality of a complete working agent is that of providing an evaluation of the state of the world[5] (*e.g.*, Albus, 1990).

This experiment consisted in giving the controller an emotion-dependent reinforcement value. The reinforcement 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.

Results are shown in figure 6. The analysis of the results suggested the existence of two quite differentiated populations, one that manages to learn the task and another that does not. For this reason, in the presentation of the results two populations were distinguished, based on the emotion-dependent reinforcement obtained at the last ten evaluation points. The trials that had quite negative reinforcement for these testing points formed one population which amounted to 38% of the total number of trials. The remaining trials were included in the "successful" population.

*** Figure 6 goes here ***

An alternative partition of populations was made based on the robot's final behavior. This new partition is consistent with the previous in that good behavior is usually associated with good reinforcement and bad behavior with low reinforcement. A total of 58% of the robots managed to converge to a suitable behavior, namely circling near a light or just wandering near a light in such a way that they receive plenty of light and never get hungry. Two of these robots made use of the fast-forward action, with different levels of success, to have a higher

[5] There are also some researchers (*e.g.* Michaud et al., 1996; Shibata et al., 1996) who give emotions the somehow more sophisticated role of monitoring the robot's performance so that the robot's plans or actions can be changed if necessary.

reinforcement through the increase in temperature. Yet, the remaining 42% end up behaving in a totally inappropriate way (*e.g.* bumping into walls or lights). If the robot ends circling in an open space getting increasingly hungry because it was not near any light, its behavior was also considered inappropriate. This would happen frequently enough to suspect that the circling behavior near a light was not a robust behavior, but a sort of *accidental* behavior. In fact, if the robot were moved away from the light to a open space, it would just remain with its circling behavior as if nothing had happened. In time, after hunger begins to be noticeable, it will learn to behave differently. Yet, this change of behavior will be mostly due to new learning and not to any previous learning.

At this point, the role of emotions was re-evaluated. The reason for the poor results obtained appears to be that the controller is not receiving the kind of reinforcement it needs. The controller needs a good evaluation of the situation as it stands at the moment and not the mixed evaluation of present and recent past situations that the emotions provide.

The reinforcement provided by emotions is thus quite misleading. Even when a good action selection is made, the robot may still receive negative reinforcement (and vice versa). An example will make this problem clear. Imagine that the robot bumps into a wall. It will feel pain and therefore become fearful. During the time that it is close to the wall it will be fearful, but even if it finally manages to go away, by taking a move-backwards action for example, it will still receive negative reinforcement, because the fear emotion will persist for a while even when the wall is out of reach. So although the fear intensity will get smaller, it will still be there providing undeserved negative reinforcement.

5.3 Sensation-dependent reinforcement values

To test the hypothesis suggested by the previous results, a control test was made that consisted in giving the controller a putatively more adequate reinforcement signal. This new signal was based on the previous one but without any temporal or lateral side effects. The sensations were used instead of the feelings to calculate the value of each emotion and the highest of these values was selected to be the reinforcement value (see figure 7). As in the previous experiment, the value of "bad" emotions was negated. This procedure provides a more traditional reinforcement that directly reflects the immediate situation.

*** Figures 7 and 8 go here ***

The results obtained in this experiment are shown in figure 8. They are a considerable improvement on the results obtained with the previous experiment which used emotion-dependent reinforcement.

Most of the final behaviors of the robot were quite successful. Many would converge to the circling behavior near a light or wandering near a light. These behaviors tend to be much more robust than the ones of the previous experiment, in that in general there was not so much preference for just one action. Instead, the final behaviors would use a small subset of actions involving both forward and backward turning movements much more often.

Another group, about 40%, were wandering about using the fast-forward action a lot. This kind of behavior gets them very good reinforcement, because this action raises the temperature, which is considered beneficial. Yet, because these robots do not always keep within reach of one light, their reinforcement is unstable. The reinforcement may suffer substantial drops, if the robot begins to become significantly hungry because of being away from a light source for a while. Although this kind of behavior has better reinforcement in general, it was only selected once in the emotion-dependent reinforcement experiment. There was another case in that experiment where the same sort of behavior was found, but the robot would bump into lights all the time. This led to the suspicion that this kind of behavior under emotion-dependent reinforcement might degenerate into some sort of crashing behavior. To test this hypothesis, a small experiment was made that consisted in starting the emotion-dependent reinforcement learning experiment with robots that had converged to this wandering behavior. Two out of the ten robots tested converged to crashing behavior. When the same experiment was done with sensations-dependent reinforcement all robots maintained their wandering behavior.

In the evaluation of these behaviors, the restrictive short range of the robot's sensors and the fact that the inputs of the networks provide only a view of the current situation should be taken into consideration. So one cannot expect the robot to act based on previous actions or sensings or some sort of global map of the environment.

5.4 Human designed behavior

A behavior was designed to take full advantage of its environment and achieve maximum reinforcement, by carefully selecting the initial weights of the networks. There were two main reasons for the design of this behavior:

- to determine how much reinforcement a reasonably successful behavior might receive;

- to check if a successful behavior is stable under the learning algorithm.

To simplify the process of design of this behavior, the set of actions from which the controller makes its selection was slightly modified. In the new set of actions, the backwards movement is done in a straight line.

The behavior consisted in having the robot, oriented towards a light, selecting between fast forward movement and backward movement, depending on how far from the light it was at each point. This would give rise to a kind of interleaved attraction and repulsion to light behavior. The robot's reinforcement would be optimal because its temperature would reach its maximum value, the robot would eat a lot and would not feel any hunger, pain or restlessness. The networks' initial weights were pre-defined in such a way that that the result behavior would be the one just described. Some minor settings of the weights were made in order to give it a little of avoiding behavior, although this was not very successful.

Three new experiments were made with the robot starting off with this human designed behavior, each one of them corresponding to each one of the experiments reported previously: no learning, emotion-dependent learning and sensation-dependent learning (see figures 9, 10 and 11).

*** Figures 9, 10 and 11 go here ***

Without learning this behavior would end up receiving 0.537 of mean emotion-dependent reinforcement and 0.297 of mean sensation-dependent reinforcement. When the robot was allowed to learn either with sensation-dependent or emotion-dependent reinforcement, it maintained the initial behavior and kept similar reinforcement values, apart from a small increase in variance due to the exploration characteristic of learning.

In the experiments reported previously, the robots learning from emotion-dependent and sensation-dependent reinforcement would sometimes reach reinforcement levels similar to these. Yet, the algorithm does not always manage to converge to such good solutions and in the case of some previous emotion-dependent reinforcement trials it converged to receiving very bad reinforcements. It should be noticed that the exact behavior that was designed cannot be achieved with the set of actions normally used. Nevertheless, slightly more sophisticated behaviors were learned in the previous experiments that achieved the same kind of reinforcement for long periods of time. Yet, these learned behaviors were less stable, because it is much more difficult for the controller to keep track of a light with an action set with non-invertible actions.

If the first three experiments are repeated with the robot equipped with the new set of actions, its performance in the learning task is actually worse. The reason for this lower performance is probably the fact that it is easier to avoid further encounters with an obstacle that appears in front of the vehicle if the backward movement is not done in a straight line. The results with this new set of actions agree, however, with those previously obtained in that emotions provide an inadequate reinforcement signal for this task.

5.5 Further experiments for analysis of results

5.5.1 Emotions selection threshold

The main unexplored difference between sensation-dependent and emotion-dependent reinforcement is the threshold used in emotions. If the intensity of the emotions is too small, then there will be no dominant emotion selected and the emotion-dependent reinforcement will be zero. The selection threshold used in the previous experiments was 0.2; an experiment was made that applies this same threshold to sensation-dependent reinforcement. No significant differences arise from the use of the threshold in the sensation-dependent reinforcement — the results graphs (see figure 12) are similar to those previously obtained. Apparently, the Q-learning mechanism seems to solve the threshold-added difficulty easily.

*** Figure 12 goes here ***

The results of this experiment show that the selection threshold by itself is not responsible for the low performance of emotion-dependent reinforcement.

5.5.2 Reinforcement dependent upon change of rate

It was noticed previously that one of the problems of emotions might be to continue giving a negative (or positive) reinforcement even when the situation is improving (or deteriorating). An attempt to minimize this problem was to have the reinforcement value be the difference between the previous and the current emotional value. In other words, the reinforcement value would reflect the improvement or decline of the robot's situation.

The value of this kind of reinforcement is in general smaller; for this reason, the temperature used in the selection of a new action by the Action Selection Module was decreased to 0.02.

Figures 13 and 13 show the results for using change of rate for both the emotion-dependent and sensation-dependent reinforcements. Once again, the robot performs much better with the sensation-dependent reinforcement. The use of change of rate proved to be unsuccessful in improving learning with emotion-dependent reinforcement.

*** Figures 13 and 14 go here ***

5.5.3 Body image

The value of the feelings given as input to the neural networks are not the robot's raw sensations, but are influenced by the emotions through the hormone system. The robot has a false image of its body sensations. Does this influence the learning task? A brief analysis of the networks' weights showed that the feelings had an active role in influencing the controllers preferences in the selection of actions.

Figure 15 shows the results for an experiment that feeds the controller network inputs with sensations instead of feelings. There were no significant changes in the reinforcement rewards.

*** Figure 15 goes here ***

If the networks' feelings inputs are totally removed (*i.e.* if the feelings' input values are replaced by a constant number) the results are still similar to those previously reported: performance for both sensation-dependent and emotion-dependent reinforcement does not suffer any significant change.

5.5.4 Learning all the time

The robot's controller is designed to learn continuously. So its performance should be better if it is allowed to learn during the evaluation. If not allowed to learn, a bad controller might just keep pushing the robot to bad situations, not really allowing for independent values to be obtained during evaluation and providing a much worse evaluation than the one deserved.

To test this hypothesis, a experiment was made that evaluated the robot's performance while the robot was learning (see figure 16). The robot seems to perform a bit better, although not significantly better, when learning all the time. It is also interesting to notice that the variance of the sensation-dependent reinforcements is much smaller, which can be easily explained by the evaluations not favoring extremes as much as before.

*** Figure 16 goes here ***

5.5.5 Simple Action Selection Module

Modifications were made to the "vanilla" Q-learning algorithm in terms of the action selection mechanism. This algorithm is usually associated with a stochastic action selection at every step. In the experiments reported in this document, this is not the case. In order to have the current action changed less often, some mechanisms were added that prevent an action selection in every step by maintaining the current action.

A new set of experiments was made with traditional action selection at every step. The main differences found in the new results when compared with the previous ones are that the task becomes easier to learn, but the final behavior is less impressive. As expected, the robot's final behavior is much more *hesitant* and in general, the area covered by the robot is smaller. However, the robot seems more successful at avoiding collisions which probably makes the learning task easier.

Once again, the experiment with emotions-dependent reinforcement performed considerably worse than the experiment with sensations-dependent reinforcement. Nevertheless, it received much higher reinforcements than previously.

5.5.6 Different networks

The networks used might be a bit over-complex for the task in hand, because using only one neuron in the hidden layer gave rise to similar results (see figure 17).

Not using random initializations for all the network weights might seem strange, but results show that this did not negatively influence the learning performance (see figure 18).

*** Figures 17 and 18 go here ***

5.6 Exploration of other uses of emotions

5.6.1 Emotions modulating learning rate

Human learning abilities are strongly dependent on their emotional state. In the longer term, strong emotions often give raise to vivid memories, while emotionally neutral events are easily forgotten (Schwartz and Reisberg, 1991).

This experiment explores the use of emotions as a sort of learning gain. This learning gain ($\mathcal{G}$) is used to influence the learning rate (η) in the following way:

$$\eta = 2\,\mathcal{G}(\text{emotion})\,\eta_{default}$$

To obtain the normal results with fixed learning rate, we should have a constant learning gain ($\mathcal{G}$) of 0.5.

All previous experiments used a learning rate value of 0.1. New experiments using sensation-dependent reinforcement were made with the following learning rates: 0.05, 0.20 and 0.30. The new results obtained were similar to the previous ones (see figure 8 for reference).

Next some experiments were made using a variable learning rate dependent on the robot's emotional state. The robot would have a higher learning rate if under a dominant emotion, be this positive or negative, than if there were no dominant emotion. In these experiments (results in figures 19 and 20) the learning gain was set to the intensity of the present emotion, or zero if there were no dominant emotion. This means that if a dominant emotion is not present the robot will not learn at all. It is quite surprising how the robot still manages to maintain its high rate of success in the learning task (although the learning is a bit slower) if one takes into consideration that neutral states are not being learned at all by the neural networks. If the weights of all network layers are initialized with random values (see figure 20), some (64%) of the robots present severe difficulties in learning (although they do manage to learn in the end). This is the result of the initial preferences of the system not being neutral. If these initial preferences happen to favor the wrong kind of actions, the robot will first have to unlearn these. Because the robot is not learning all the time, this can result in a high performance cost.

*** Figures 19 and 20 go here ***

It is possible to conclude that the robot does not need to waste computing time at learning if there is no emotion present. This will not compromise its performance, unless the robot's initial preferences are very misleading. Even in this case, the drop in performance is only temporary.

5.6.2 Variable emotion-dependent action-selection temperature

Previous experimental results might be enhanced if the fixed exploration versus exploitation ratio is upgraded to a more sophisticated selection algorithm that takes into consideration the deadlock situations in which the robot gets trapped. In these situations, there is an option that is by far better ranked than the others and therefore always gets selected although its practical utility turns out to be very low and it is not able to change the situation at all. This suggests that being in a bad emotional state for a while should trigger more exploration than usual.

To begin with, some experiments were made that showed that the robot learning with sensation-dependent reinforcement would perform as well as previously (see figure 8) if its action selection temperature of 0.1 was changed to 0.2 or 0.03.

Next, emotions were used to modulate the modulate the exploration versus exploitation ratio, by directly influencing the temperature parameter of the action selection module. Results in figures 21 and 22 use the following emotion-dependent functions to set the temperature:

$$\mathcal{F}_1(E) = \begin{cases} \mathrm{E}_{Intensity}/4.0 & \text{if } \mathrm{E}_{Intensity} \neq 0 \\ 0.05 & \text{otherwise} \end{cases} \tag{4}$$

$$\mathcal{F}_2(E) = \begin{cases} \mathrm{E}_{Intensity}/4.0 & \text{if E is negative} \\ 0.05 & \text{otherwise} \end{cases} \tag{5}$$

Apparently, the results for function $\mathcal{F}_2$ (figure 22) are an improvement over the results previously achieved. The use of this function allows the robot to explore new solutions when it is in a bad situation. The function $\mathcal{F}_1$, increases the temperature independently of whether the robot's emotional state is good or bad. Disrupting behavior that is being successful does not seem to be such a good idea, although no significant changes can be found in using $\mathcal{F}_1$ (figure 21) when compared with the standard experiment (figure 8).

*** Figures 21 and 22 go here ***

6 Conclusion

The model of emotions developed proved to be quite good. The robot has the right emotions at the right moments. Furthermore, because its emotions are grounded in body feelings, and not direct sensory input, it manages to avoid sudden changes of emotional state, from one extreme emotion to a completely different one. The more different the emotions are, the more difficult it is to change from one to the other. The physiological arousal caused by emotions has been repeatedly left out of cognitive theories of emotions (the FEELER by Pfeifer, 1982, can be considered an exception), because it is not considered cognitively interesting, yet without it emotions lack their characteristic inertia (Moffat et al., 1993).

Unfortunately, the experiments done failed to show that emotions can be used to increase the robot's autonomy, yet they do not show that such increase cannot be achieved. More importantly, what they do show is that the role of emotions is more intricate than that often assumed and that a simple approach to the use of emotions as context judgment values is not very successful when a more than usually realistic emotions model is used. More attention should be given to the role attributed to emotions in adaptation.

The emotions do not really provide a good evaluation of what is going on at any one moment, but are a sort of mixed evaluation the robot acquired from its past experiences. This is good to modulate its behavior, but not to be taken at face value when trying to predict the outcome of each one of its actions.

Preliminary results suggest that emotions can be successfully used in modulating the learning rate and the exploration versus exploitation ratio. Yet, no solid conclusions can be drawn with the experiments done so far. It is felt that the sensation-dependent reinforcement controller is quite successful and that further improvements in performance are hard to notice when performance has reached such a high level. Therefore it is considered that for a thorough examination of these mechanisms the complexity of the robot's task must be increased.

This document reports on-going work of an attempt to have a truly embodied emotions model. Although some preliminary results have been achieved there is still much more to be explored.

Acknowledgments

The first author is supported by a scholarship from the program PRAXIS XXI. Facilities for this work were provided by the University of Edinburgh. Thanks to Chris Malcolm and Nuno Chagas for helpful comments and discussion.

References

Albus, J. S. (1990). The role of world modeling and value judgement in perception. In Meystel, A., Herath, J., and Gray, S., editors, *Proceedings of the* 5^{th} *IEEE International Symposium on Intelligent Control*, Los Alamitos, California. IEEE Computer Society 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. (1992). Integrating reactivity, goals, and emotion in a broad agent. CMU-CS 92-142, Carnegie Mellon University, School of Computer Science.

Bates, J., Loyall, A. B., and Reilly, W. S. (1994). An architecture for action, emotion, and social behavior. In *Artificial social systems: Fourth European workshop on modeling autonomous agents in a multi-agent world*, Berlin. Springer-Verlag.

Beck, R. C. (1983). *Motivation: Theories and principles.* Prentice-Hall, Inc., second edition.

Blumberg, B. (1995). Multi-level direction of autonomous creatures for real-time virtual environments. In *Computer Graphics Proceedings, Siggraph'95*, Los Angeles, California.

Boden, M. A. (1993). Autonomy and artificiality. CSRP 307, University of Sussex.

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

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

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

Ferguson, I. A. (1992). *Touring Machines: An architecture for dynamic, rational, mobile agents.* PhD thesis, University of Cambridge.

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

Grey Walter (1950). An imitation of life. *Scientific American*, 182(5):42–45.

Grey Walter (1951). A machine that learns. *Scientific American*, 185(2):60–63.

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.

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

Malcolm, C. (1993). Behaviour, purpose & meaning. Communication from author.

McFarland, D. (1994). Animal robotics — from self-sufficiency to autonomy. In Gaussier, P. and Nicoud, J.-D., editors, *Proceedings of "From Perception to Action" conference*, pages 47–54.

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`.

Moffat, D., Frijda, N., and Phaf., R. (1993). Analysis of a model of emotions. In et al., A. S., editor, *Prospects for Artificial Intelligence*, pages 219–228. IOS Press.

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.

Schwartz, B. and Reisberg, D. (1991). *Learning and memory.* W.W. Norton and Company, Inc.

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 1996 IEEE International conference on Robotics and Automation*.

Sloman, A. (1987). Motives mechanisms and emotions. *Emotion and Cognition*, 1(3):217–234. Reprinted in M. A. Boden, editor, *The Philosophy of Artificial Intelligence*, "Oxford Readings in Philosophy" Series, Oxford University Press, pages 231-247, 1990.

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.

The Oxford English dictionary (1989). *The Oxford English dictionary.* Oxford University Press, second edition prepared by J. A. Simpson and E .S. C. Weiner.

Toda, M. (1994). The urge theory of emotion and cognition: Chapter 2 basic structure of the urge operations. SCCS technical report 94-1-01, Chuyko University.

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

Wilson, S. W. (1991). The animat path to ai. In *From animals to animats. Proceedings of the first conference on simulation of adaptive behaviour.* 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.

A Emotions model

A.1 Functions

First, let us define the set of emotions ($\mathcal{E}$) and the set of feelings ($\mathcal{F}$),

$$\mathcal{F} = \{\text{Hunger, Pain, Restlessness, Temperature, Eating}\} \tag{6}$$

$$\mathcal{E} = \{\text{Happiness, Sadness, Fear, Anger}\} \tag{7}$$

and define the function $\text{Th}_{[b_-,b_+]}(x)$ that allows to bound the value of a x value to the interval $[b_-, b_+]$.

$$\text{Th}_{[b_-,b_+]}(x) = \begin{cases} b_- & \text{if } x < b_- \\ b_+ & \text{if } x > b_+ \\ x & \text{otherwise} \end{cases} \tag{8}$$

Taking this into consideration, the intensity value of emotion e at step n (I_{e_n}), depends on the intensity of the feelings in the following way:

$$\forall e \in \mathcal{E}, \forall n \in \mathbb{N}, \qquad I_{e_n} = \text{Th}_{[0,1]}(B_e + \sum_{f \in \mathcal{F}} C_{ef} I_{f_n}) \tag{9}$$

To calculate the intensity of the feeling f at step n (I_{f_n}) the following functions are used. The feeling's intensity has to take into account both the influences provided by the hormone system (H_{f_n}) and the value of the sensations (S_{f_n}).

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

$$I_{f_n} = \text{Th}_{[0,1]}(C_h H_{f_n} + S_{f_n}) \tag{10}$$

$$H_{f_1} = 0, \qquad H_{f_{n+1}} = \alpha H_{f_n} + (1-\alpha) A_{f_n} \tag{11}$$

$$A_{f_n} = \sum_{e \in \mathcal{E}} C_{ef} I_{e_n} \tag{12}$$

$$\alpha = \begin{cases} \alpha_{up} & \text{if } |A_{f_{n-1}}| > |H_{f_{n-1}}| \\ \alpha_{dn} & \text{otherwise} \end{cases} \tag{13}$$

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}). 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 system parameters that have been used in the emotions model's functions are the following:

$$C_{ef} = \text{Value of the emotion-feeling dependency between emotion } e \text{ and feeling } f$$

$$C_h = \texttt{HormoneCoef}$$

$$\alpha_{up} = \texttt{HormAlphaUp}$$

$$\alpha_{dn} = \texttt{HormAlphaDn}$$

See appendix C for parameter values.

A.2 Sensation definitions

Energy level and sensations intensities are values bounded between zero and one.

The robot is initialized with maximum energy level. The robot takes `EnergyAutoStopSteps` steps to lose all its energy if it is stopped, and `EnergyAutoRunSteps` if it is moving at full power. The decrease in energy level is proportional to the total motor activity (*i.e.* the sum of the absolute values of both left and right motor values). `EnergyRechargSteps` is the number of steps necessary to recover all its energy if the robot receives maximum light on its sensors. The energy will increase only if the sum of the values of the robot's front light sensors is high enough (*i.e.* 60% of its maximum value). If the previous condition is met, then the increase in energy is directly proportional to the light received by these sensors. The previous descriptions of the processes of increase and decrease of energy level assume independent processes, *i.e.* each description considers that the energy value is only modified by the process being described. In reality, the effects of both processes in the energy level are calculated separately and subsequently added.

Hunger is one minus the energy level.

The Eating sensation is non-zero and directly proportional to the light perceived if the light received by the robot's front sensors is considered high enough to increase its energy level. If the energy level is very high (> 0.95) the Eating sensation is in addiction multiplied by one minus the energy level, *i.e.* the Hunger sensation.

If the robot is bumping then the pain is proportional to the number of distance sensors with high values (over 1020), otherwise it is zero. The pain value starts at 1/3 and increases by 1/6 for each high-valued distance sensor.

If the robot travels a good distance (Manhattan distance higher than 1), its restlessness will decrease; otherwise it increases. The parameter `BoredomSteps` is the number of steps the robot has to be totally stopped to reach its maximum restlessness value. The change in restlessness intensity is directly proportional to one minus the Manhattan distance (*i.e.* the sum of the absolute values of the distances covered in x and in y, which lies between -4 and 1).

If the total motor activity is above a certain threshold (14), then the temperature will rise. If this value is low (below 10), then the temperature will decrease. Basically, the robot will lose temperature with all actions in its action set except for the fast-forward action. When motors are at full power (sum = 20), the robot takes `TempRaiseSteps` steps to go from no temperature to maximum temperature. It will take `TempLowerSteps` to lower its temperature back to zero with the motors off. The temperature increase is directly proportional to the total motor activity and the decrease is directly proportional to one minus the rescaled total motor activity (rescaled to lie between 0 and 1).

B Controller

B.1 Discrete Actions

The set of actions used in the experiments is defined by the motor values shown in table 2. The second set of action used in the experiments reported in section 5.4 is defined by the same motor values, except for the backwards movement.

*** Table 2 goes here ***

B.2 Action Selection Algorithm

The present work does not use a simple Q-learning algorithm that evaluates the previous action and stochastically selects a new one in every step.

An action is only evaluated, and eventually a new one selected, every other step, unless there is a significant change in the emotional state, *i.e.*, a change from one of the following states to another:

- a positive emotion is dominant;

- a negative emotion is dominant;
- no emotion is dominant.

Even if an evaluation takes place, the probability of not choosing an action based on the above criteria (equation 3) if a is the currently selected action is:

$$P(NoSelection) = \sqrt[10]{P(a)} \tag{14}$$

This way the probability of an action being selected in a given step is extended to the probability of being consecutively selected in the following next ten steps.

Both the mechanisms described attempt to give a more coherent behavior to the robot, yet care was taken not to do this at the expense of:

- Preventing the controller from taking notice of sudden changes in the emotional state;
- Giving preference to the previous action independently of how well rated that action is;
- Giving preference to the previous action even when the conditions have changed significantly, making it inappropriate to do so.

C Parameters

*** Table 3 goes here ***

D Program Environment and Performance

The program developed, *sim*, is an extension of the X-windows simulator by Michel (1996). The original code was in C, but the extensions were done in C++. Some modifications were made to the original code in order to allow it to run in the background with no graphical output.

The *sim* program takes about five hours to run an experiment with one hundred and twenty thousand learning steps and sixty-one two thousand steps evaluation tests (*e.g.* the experiment reported in figure 8). This performance time was obtained for a Sun SparcStation 4 at 125 MHz with *sim* running in the background with low priority and other programs running on the same machine.

	Hunger	Pain	Restlessness	Temperature	Eating	Bias
Happiness	-0.2	-0.3	-0.2	0.2	0.7	0.1
Sadness	0.7	0.0	0.5	0.0	-0.4	0.0
Fear	-0.4	0.8	-0.2	0.15	0.0	0.0
Anger	0.2	0.2	0.3	-0.2	0.0	0.0

Table 1: The emotions' dependencies from feelings.

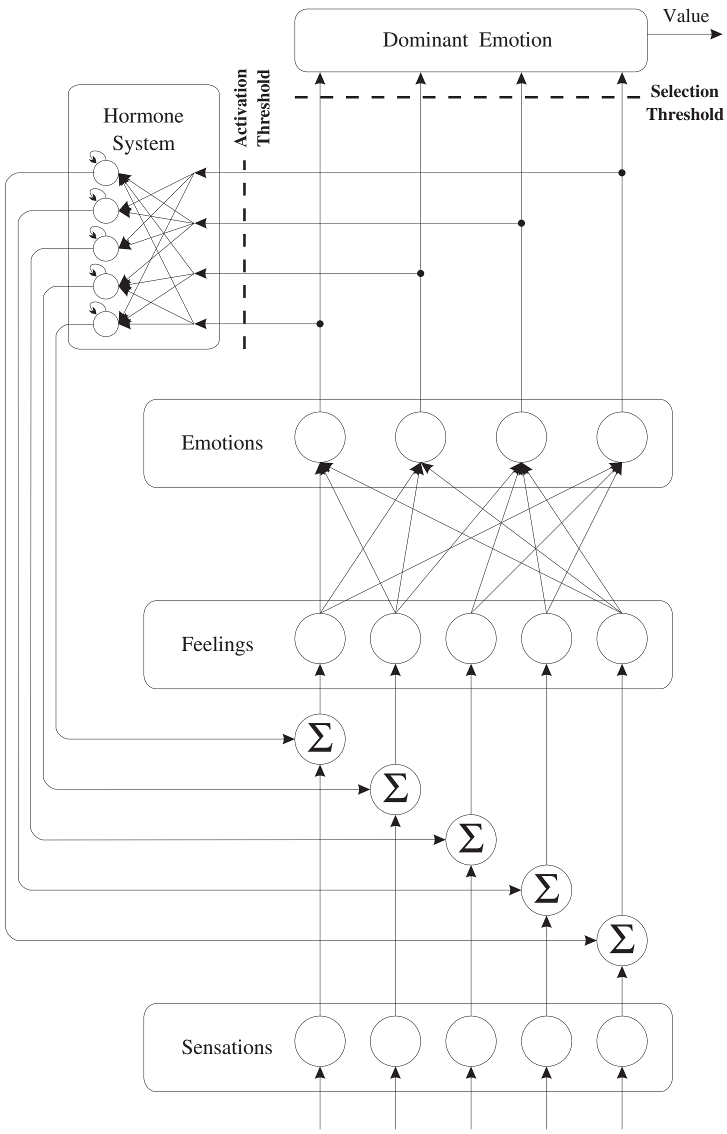

Figure 1: Emotions model.

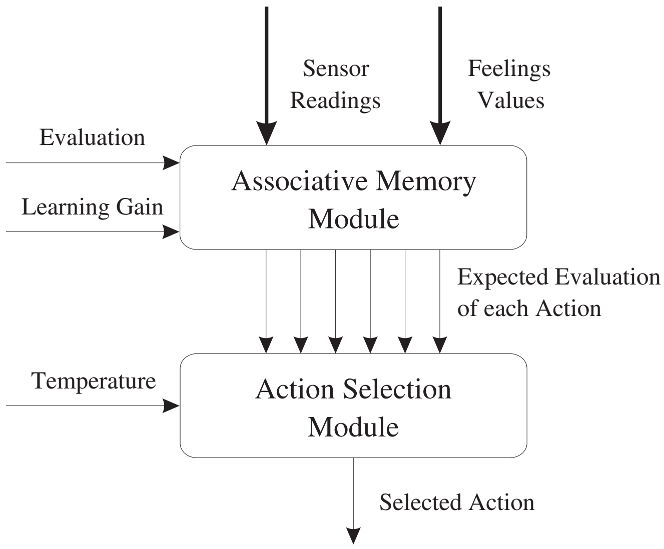

Figure 2: Adaptive action selection controller.

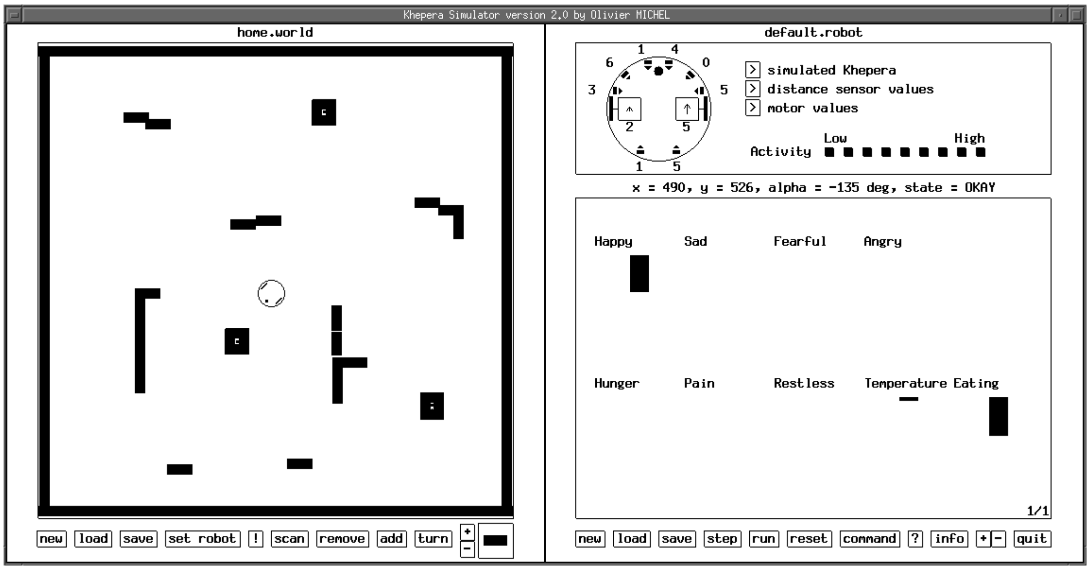

Figure 3: The robot and its environment.

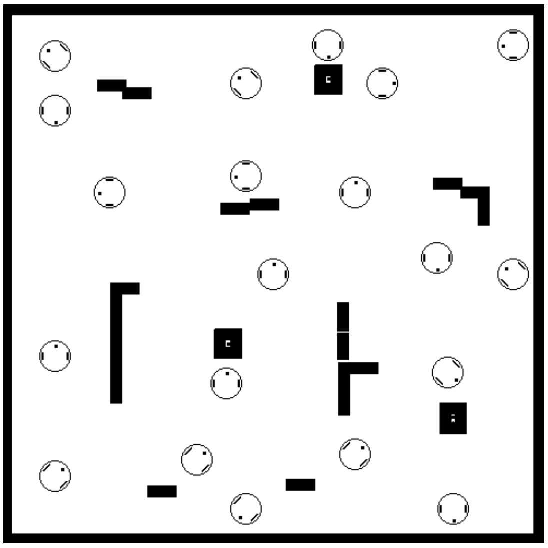

Figure 4: Starting positions.

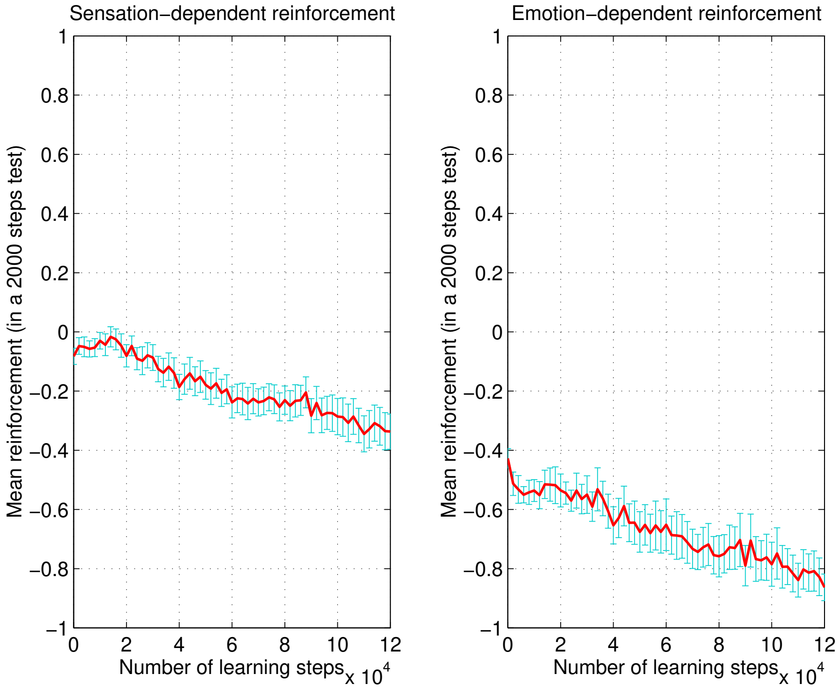

Figure 5: Controller with no learning.

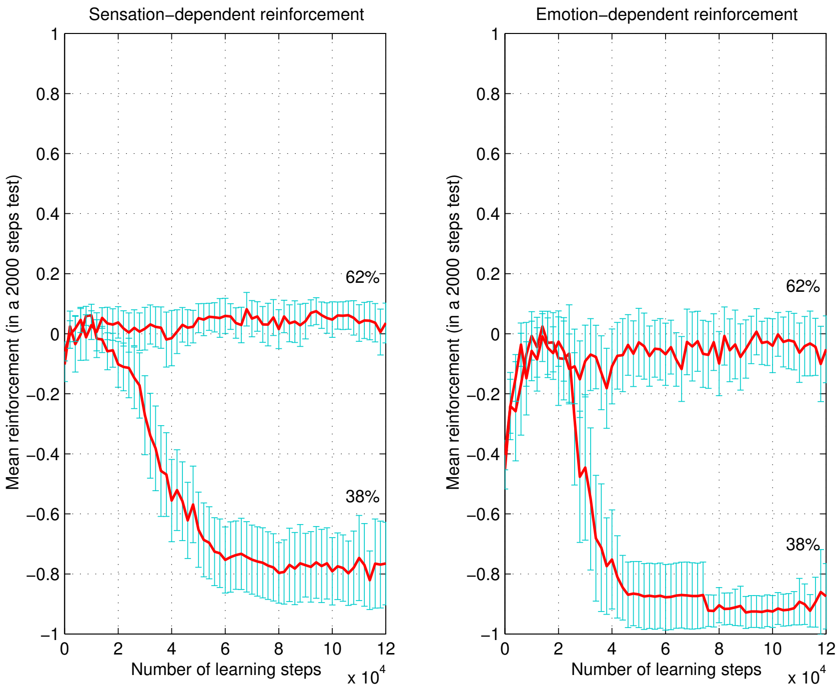

Figure 6: Learning controller with emotion-dependent reinforcement.

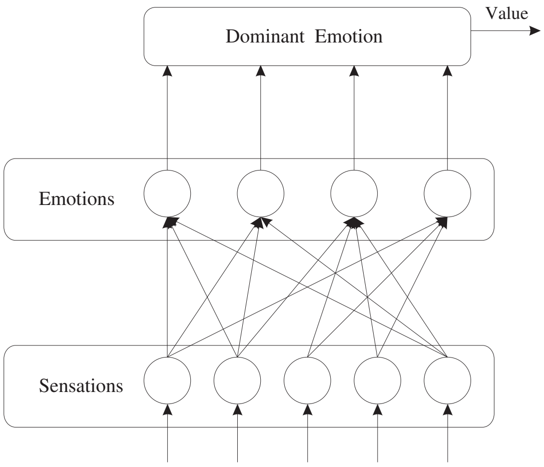

Figure 7: Truncated emotion model used to obtain the sensation-dependent reinforcement.

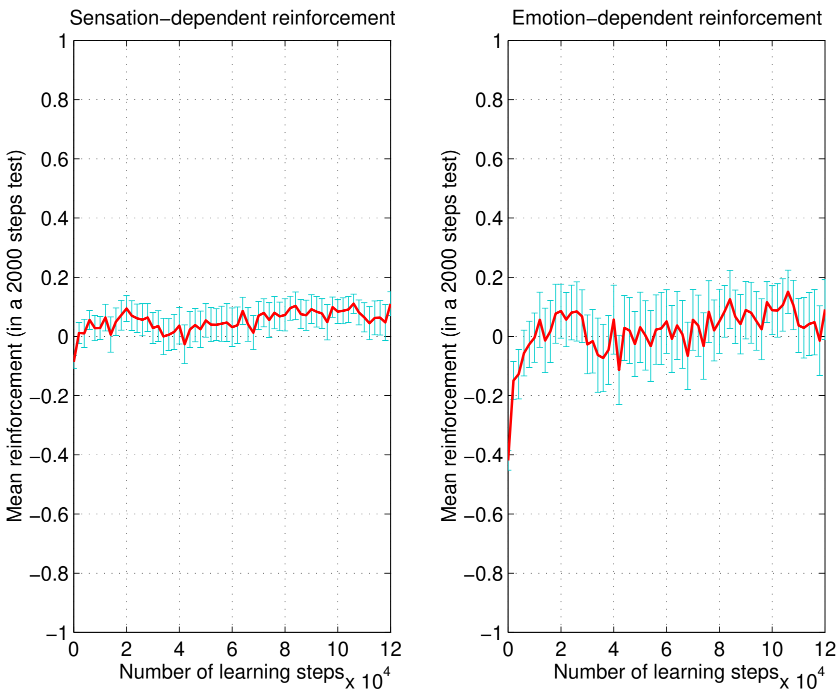

Figure 8: Learning controller with sensation-dependent reinforcement.

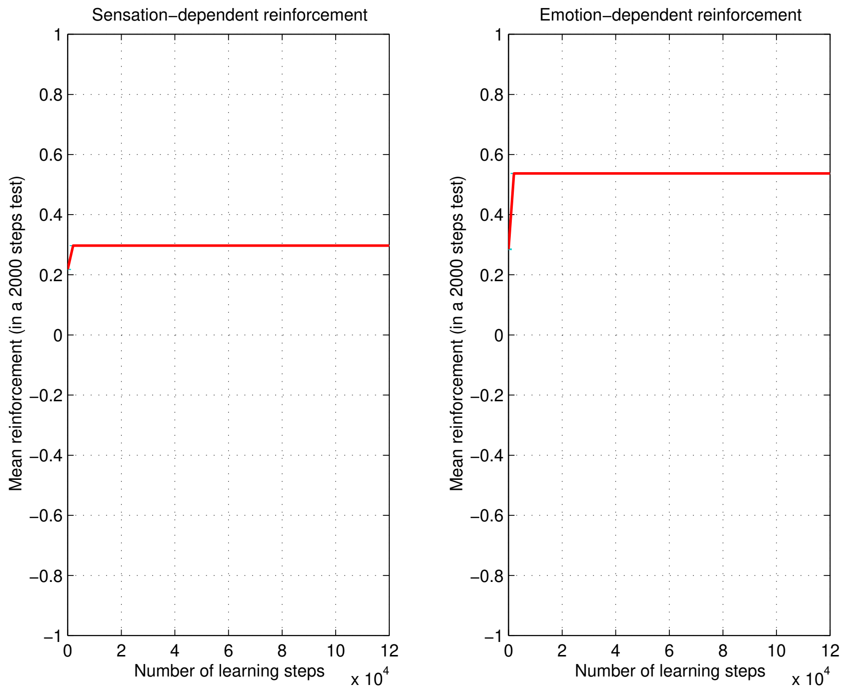

Figure 9: Controller with no learning. Human designed behavior.

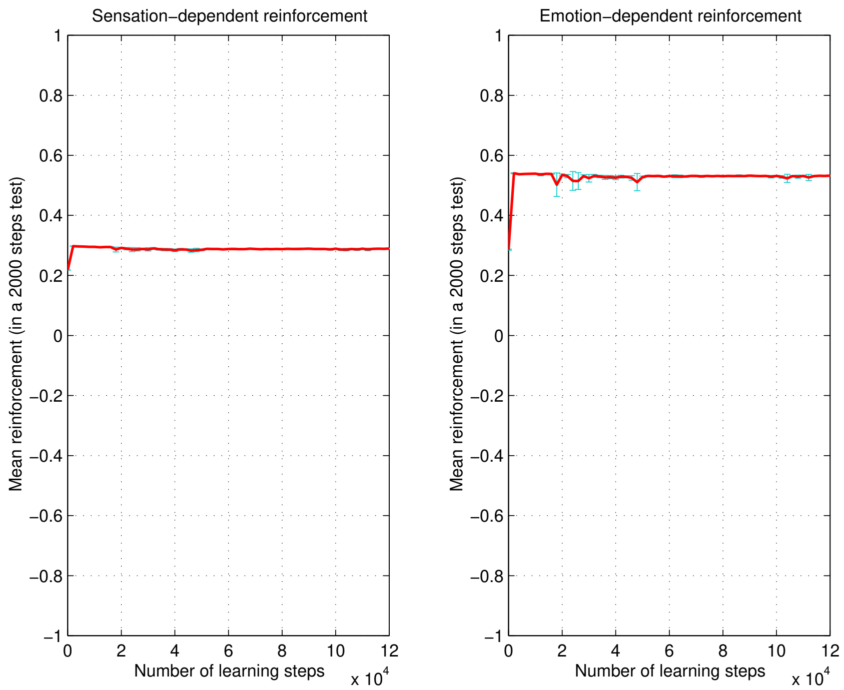

Figure 10: Learning controller with emotion-dependent reinforcement. Human designed behavior.

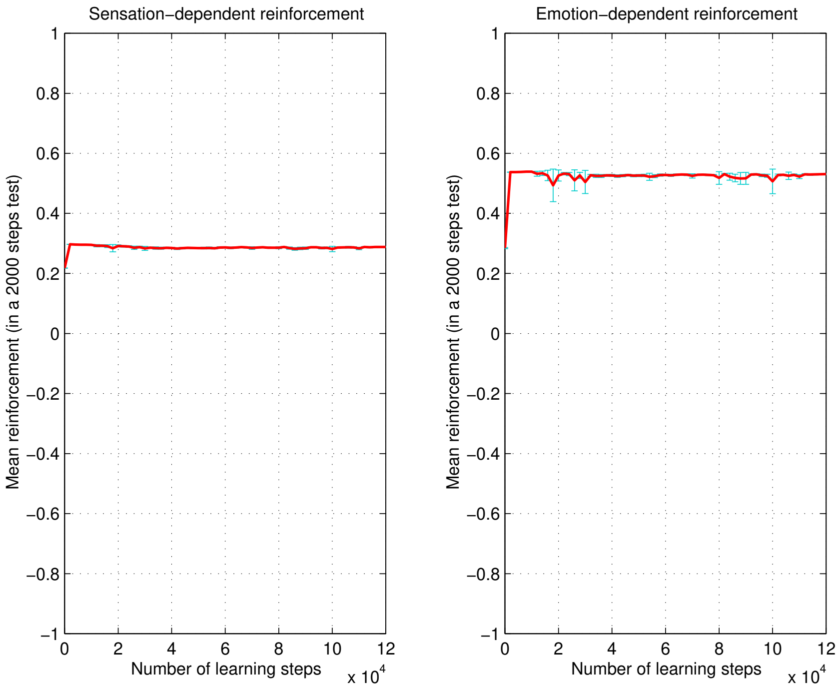

Figure 11: Learning controller with sensation-dependent reinforcement. Human designed behavior.

Figure 12: Learning controller with sensation-dependent reinforcement subject to thresholding.

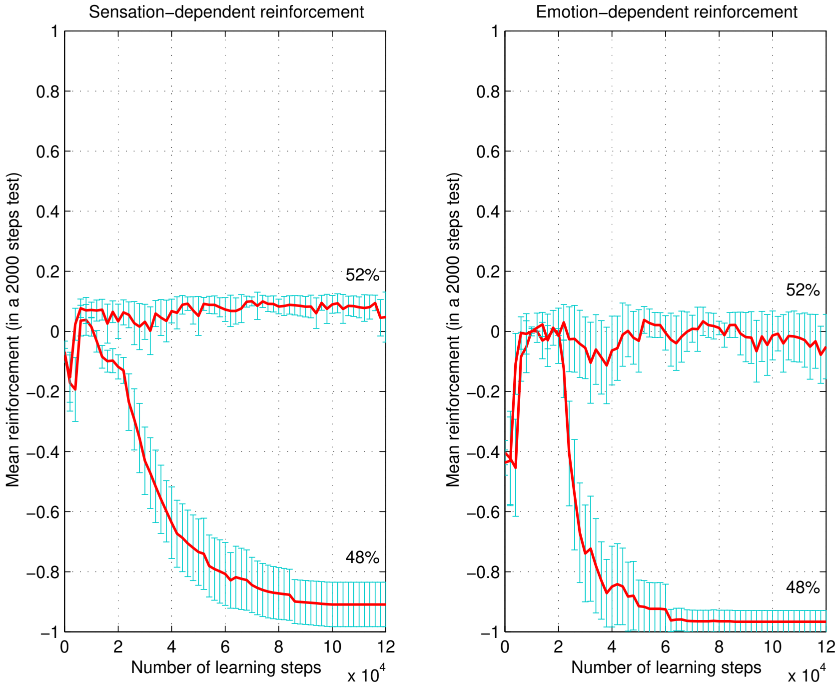

Figure 13: Learning controller with emotion change of rate dependent reinforcement.

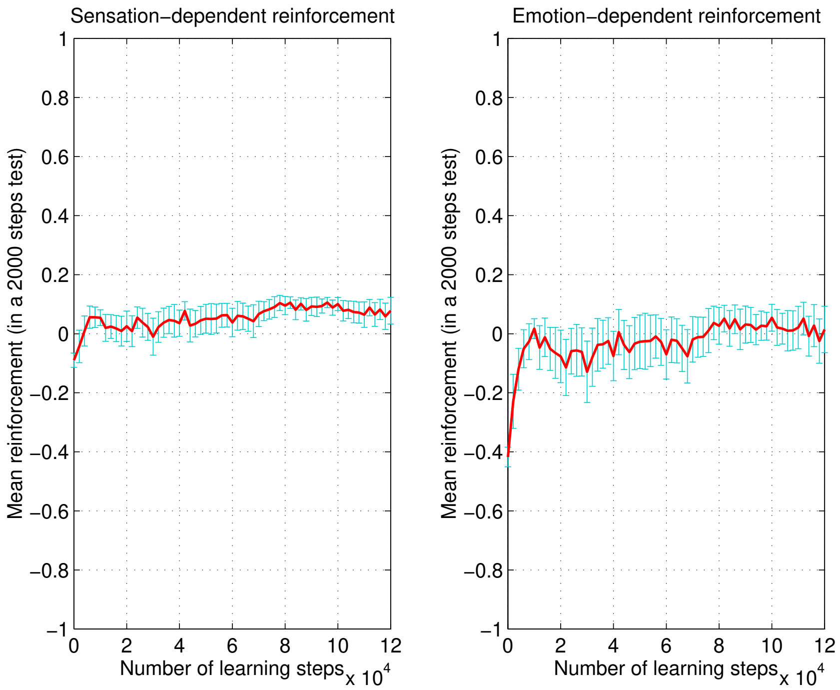

Figure 14: Learning controller with sensation change of rate dependent reinforcement.

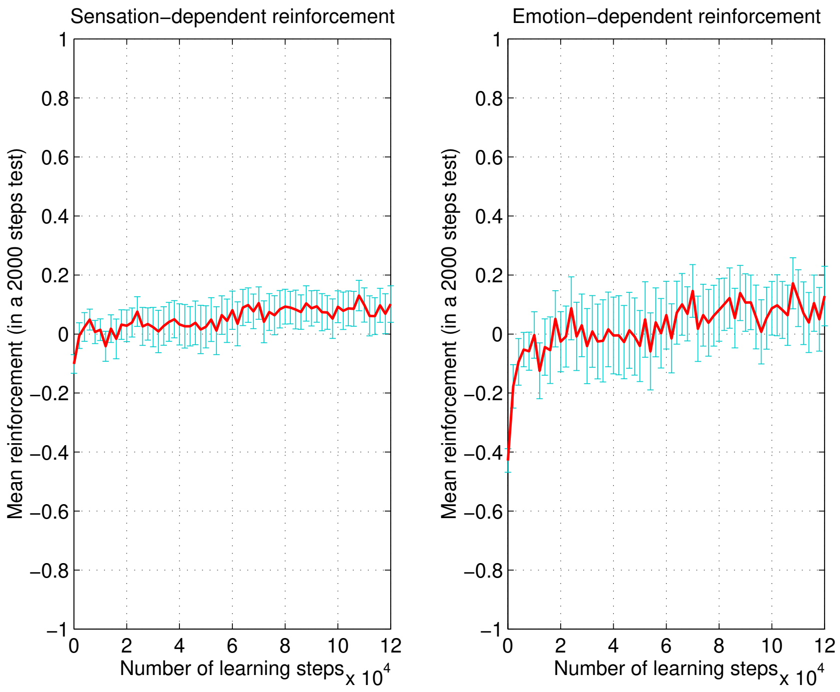

Figure 15: Learning controller with sensation-dependent reinforcement. Faithful body image.

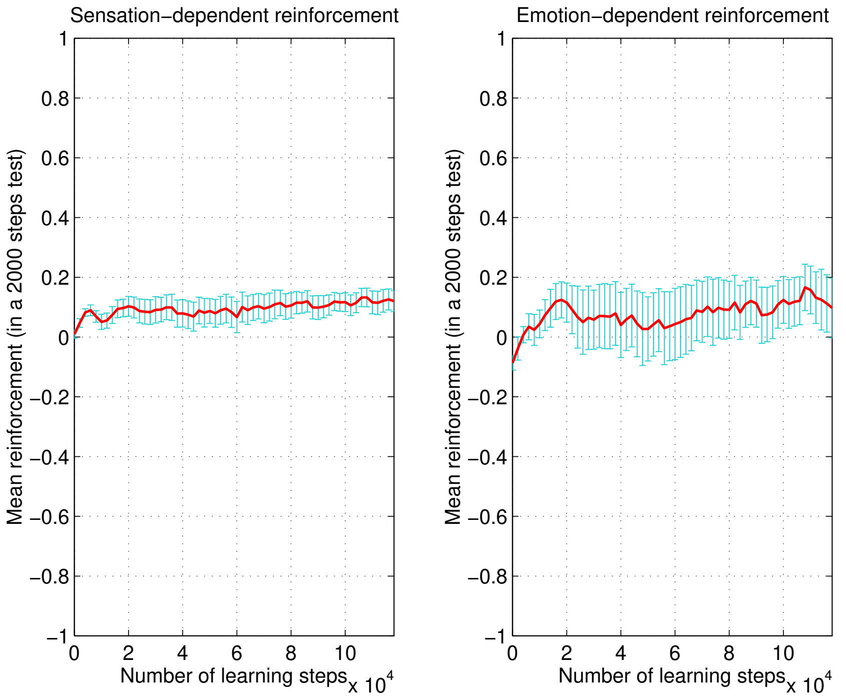

Figure 16: Learning controller with sensation-dependent reinforcement. Learning during tests.

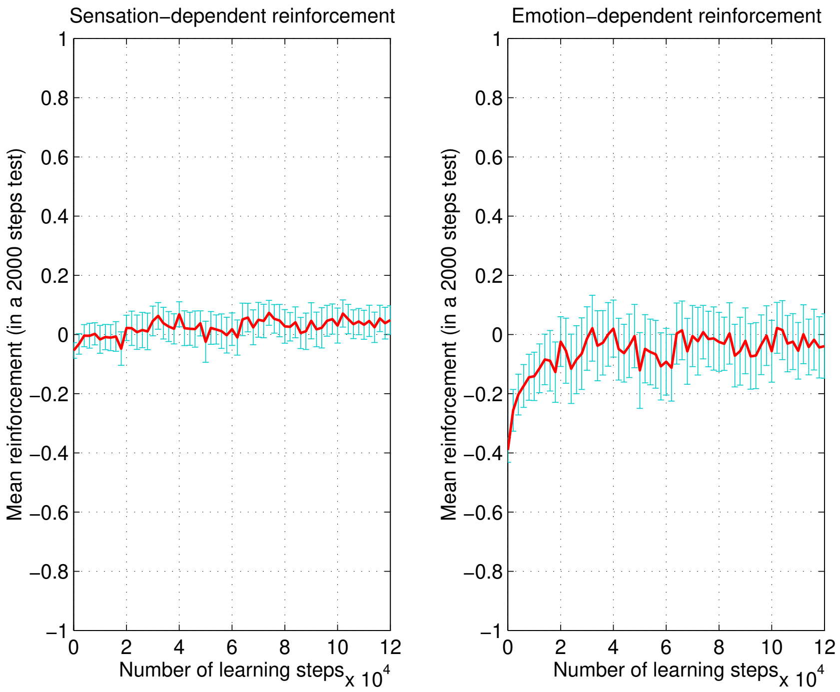

Figure 17: Learning controller with sensation-dependent reinforcement. Networks with only one neuron in the hidden layer.

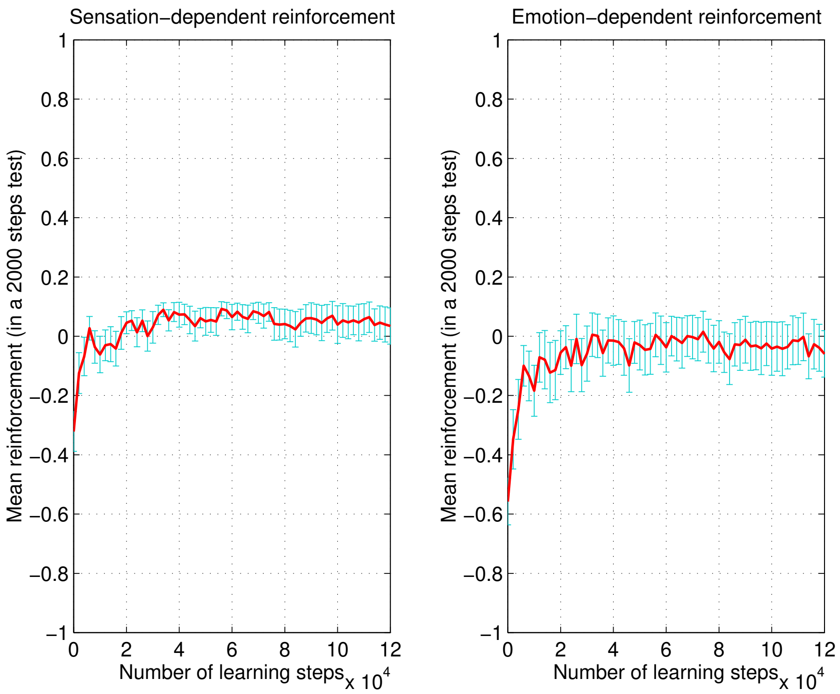

Figure 18: Learning controller with sensation-dependent reinforcement. Initial random weights in all nets.

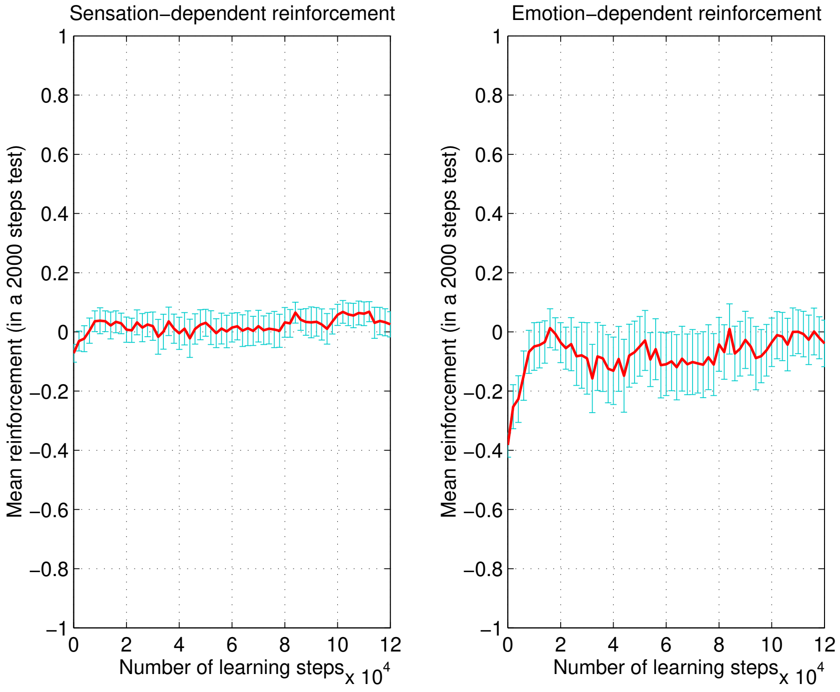

Figure 19: Learning rate dependent on current emotion.

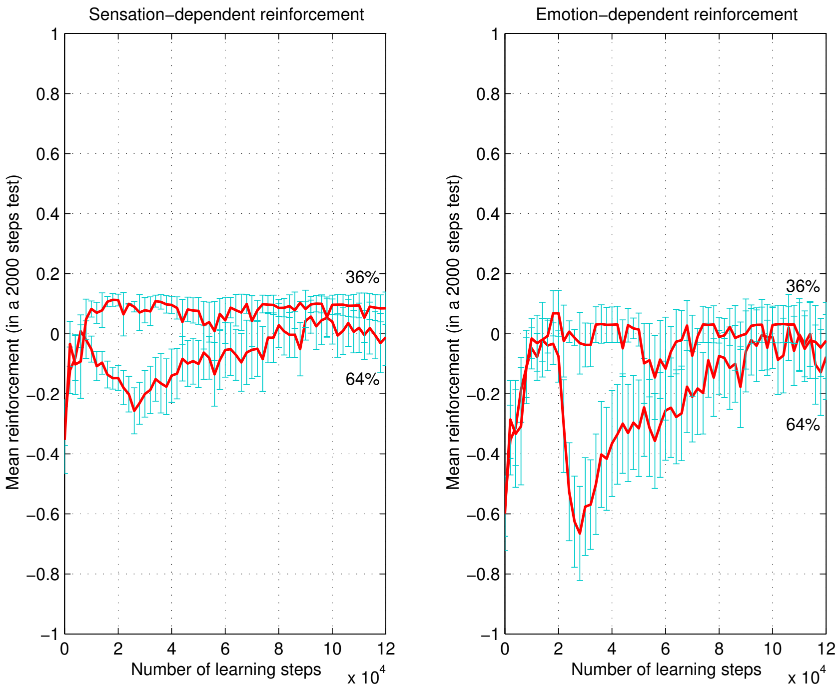

Figure 20: Learning rate dependent on current emotion. Random initial weights.

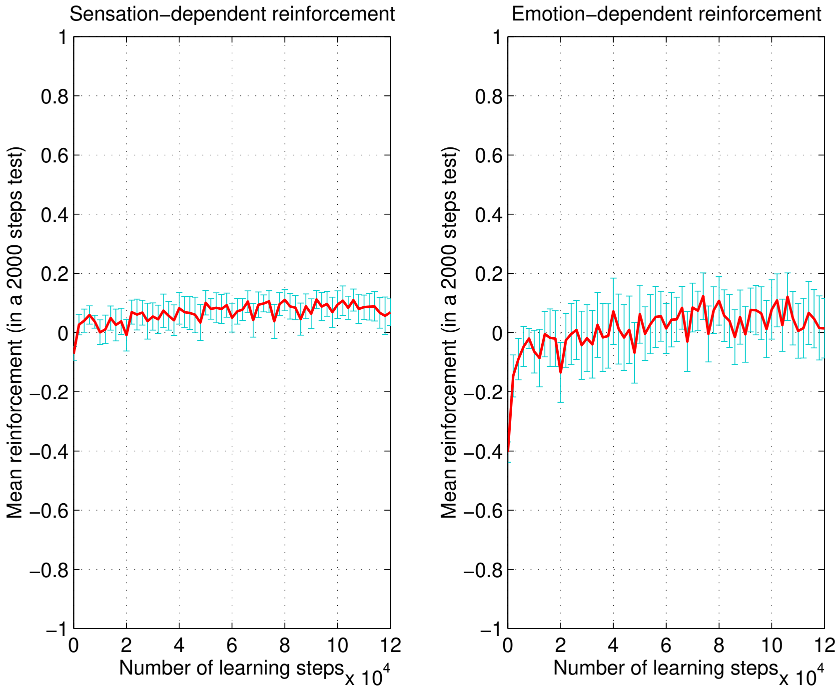

Figure 21: Temperature influenced by emotions: $\mathcal{F}_1$.

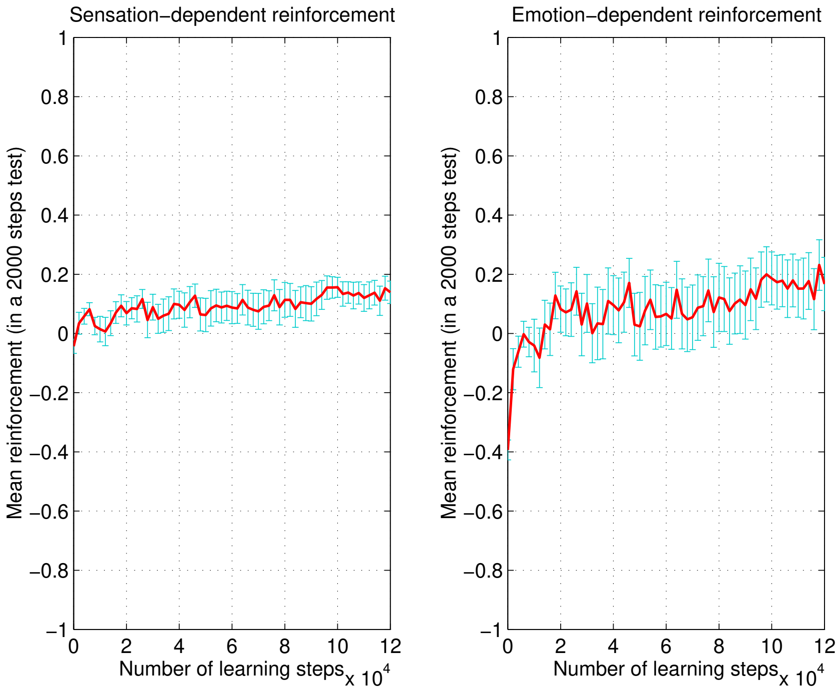

Figure 22: Temperature influenced by emotions: $\mathcal{F}_2$.

Motor power	Slow forward	Turn left	Turn right	Fast forward	Stop	Backwards
Right motor	3	5	2	8	0	-1 (-3)
Left motor	3	2	5	8	0	-5 (-3)

Table 2: Values of the robot's motors for each action. The values in brackets are the modified values for the second set of actions.

Emotions Parameters	
EmotionActiveTH	0.2
EmotionSelectTH	0.2
HormCoef	0.9
HormAlphaUp	0.98
HormAlphaDn	0.996

Controller Parameters	
BackLearnRate	0.1
ActionSelectConst	10.0

Time steps constants	
EnergyAutoStopSteps	1000000
EnergyAutoRunSteps	50000
EnergyRechargSteps	2000
TempRaiseSteps	200
TempLowerSteps	1000
BoredomSteps	600

Table 3: Parameter values used in the experiments.