Emotion-driven Learning for Animat Control

Sandra Clara Gadanho, John Hallam
Department of Artificial Intelligence
University of Edinburgh
5, Forrest Hill
Edinburgh EH1 2QL
sandrag@dai.ed.ac.uk
john@dai.ed.ac.uk

Abstract

Models of emotion are often suggested as a way of providing an evaluation of the current behaviour of an agent. In this work, we investigate whether emotions can actually provide suitable reinforcement signals for a Q-learning system to learn adaptive policies. For this purpose a recurrent network model of emotion consistent with the somatic somatic marker hypothesis of Damásio was developed. Experimental work was done in a realistic mobile robot simulator in a simple foraging-like task. Experiments revealed that having emotions providing a context evaluation for direct use as a reinforcement signal does not work, but using them as modifiers for learning system parameters could be fruitful.

1. Introduction

The fact that emotions are considered to be essential to human reasoning suggests that they might play an important role in artificial creatures as well. This paper reports work done in trying to use emotions in autonomous robot control. Emotions seem to have solved in mammals many cognitive problems involved in autonomous systems with multiple goals and uncertain environments (Oatley, 1987). It has been argued before in the field of Artificial Intelligence that emotions are the ultimate source of intelligence and might provide robots with the autonomy they need (Toda, 1994); doubts have even been posed on whether machines can exhibit intelligence without any emotions (Minsky, 1986).

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, by creating an illusion of life in a believable character (Bates, 1994). This kind of external demonstration of emotions is very important in social interaction and is of great adaptive advantage in both humans and animals (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 (Cañamero, 1997; Bates et al., 1994).

The work was done under an animat philosophy (Wilson, 1991), by building bottom-up a biologically inspired complete agent by synthesis. A robot was equipped with a recurrent network model of "emotions" which incorporates the important computational features of Damásio's somatic marker hypothesis (Damásio, 1994). Based on a simplified hormone system, the developed emotion model is far from the complexity of true emotions experienced by humans. Yet, in order to have a simple and concise discourse, language will be used that might, implicitly, attribute emotional feelings to the robot. Experiments were carried out on a simulated Khepera robot in an animal-like adaptation task to access the usefulness of emotions as a source of context value in a reinforcement learning task.

In the following sections, sections 2 and 3, a detailed description of the emotion model developed and of the adaptive controller used is presented. Next, in section 4, a full description of the experiments is given and the results are discussed. In the discussion, the results of using an emotion based controller are compared with a random and an human designed controller. Finally, section 5 concludes that emotions do not provide as good a reinforcement as expected, but may be used successfully in the modulation of learning.

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[1] (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.

Many emotions theorists agree that emotions are most helpful on focusing attention in the relevant features of the problem at hand (De Sousa, 1987; Tomkins, 1984; Plutchick, 1984; Scherer, 1984; Panksepp, 1982).

Inspired by the ideas that have been presented, an emotion model was developed that is described next.

2.1 Model

A large subset of theories of emotions is based on cognitive appraisal theories (Lazarus, 1982; Power and Dalgleish, 1997), although some evidence exists to suggest that emotions can be aroused without cognition (Zajonc, 1984).

Following the psychologists' main stream, most AI models of emotions are based on an analytic approach (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 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 — is 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 instance, toxic food was added to the environment, disgust would become useful to keep the robot away from it.

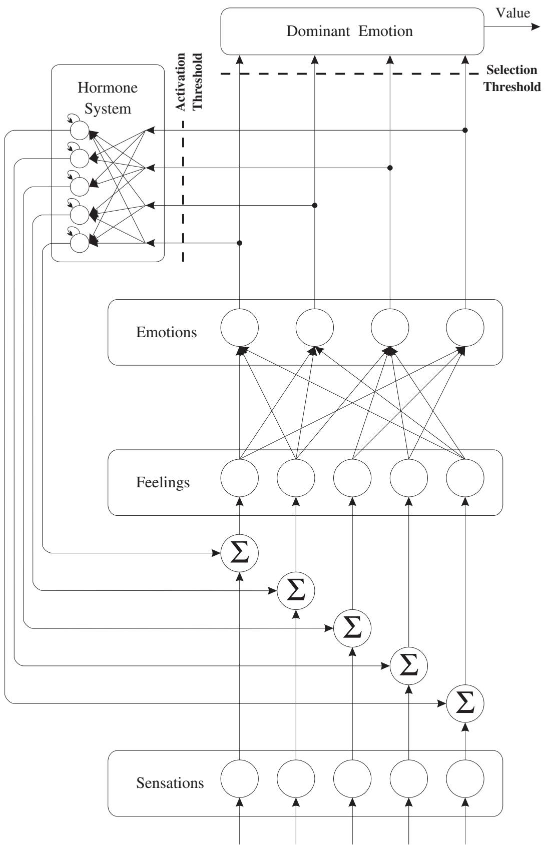


Figure 1 Emotion model.

The emotions chosen are also usually included in the definitions of primary emotions[2] (Shaver et al., 1987;

---

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

[2] 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

Power and Dalgleish, 1997), which is a good indicator of their relevance and need. Other emotions, like love and hate, which some authors like to suggest as primary emotions, were not included because they do not seem very basic [3] 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 judgement 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 constant feeling dependencies and a bias. 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.

Furthermore, based on what was suggested in (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 judgement 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 provides not only a value judgement 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 behaviour (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

It is widely accepted that robot learning is beneficial, because it is often difficult or even impossible for the designer to anticipate all possible scenarios the robot will be confronted with. In the present work an adaptive controller based on reinforcement learning was used. The controller developed — figure 2 — tries to maximise the evaluation received by selecting one of six possible discrete actions, taking into account the current sensor readings and robot feelings, and the previously received evaluations.

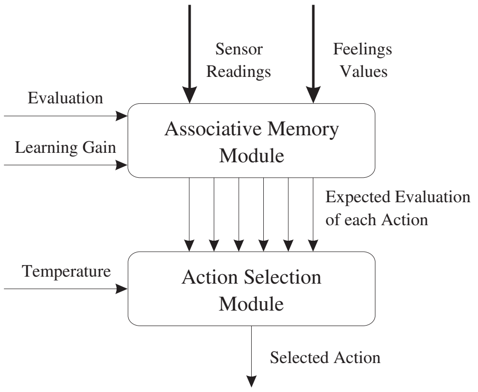


Figure 2 Adaptive action selection controller.

The controller has two separate modules:

Associative Memory Module — This plastic module associates the sensor readings and feelings with

---

environment.

[3] There are even arguments against considering them as emotions, *e.g.* (Sloman, 1987).

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 six available actions:

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

Each network has a three layer feed-forward network topology with: 22 input units, one for each distance and light sensor (values were scaled to $[0, 1]$), one for each feeling and a bias; 5 hidden units; and 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 initialised 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 is back-propagation[4].

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 (Lin, 1992) was used. The networks were used to learn utility functions that model $util(x, a)$:

$$util(x, a) = R + \gamma eval(y) \qquad (1)$$

The value R is the immediate reinforcement. The function $eval(y)$ is the expected cumulative discounted reinforcement starting from state y reached 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:

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

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:

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

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. The reason for this is to have a more coherent behaviour. Otherwise, the robot would mainly tremble, because it would be selecting different actions all the time.

4. Experiments

4.1 Setup

Experiments were carried out in a Khepera simulated robot (Michel, 1996) within the environment shown in figure 3. This is a closed environment with some walls and three lights surrounded by bricks[5].

The emotion-based robot controller was tested in the context of an animal-like creature with self-maintenance needs, making it easier to conceptually ground the emotion system. 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:

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.

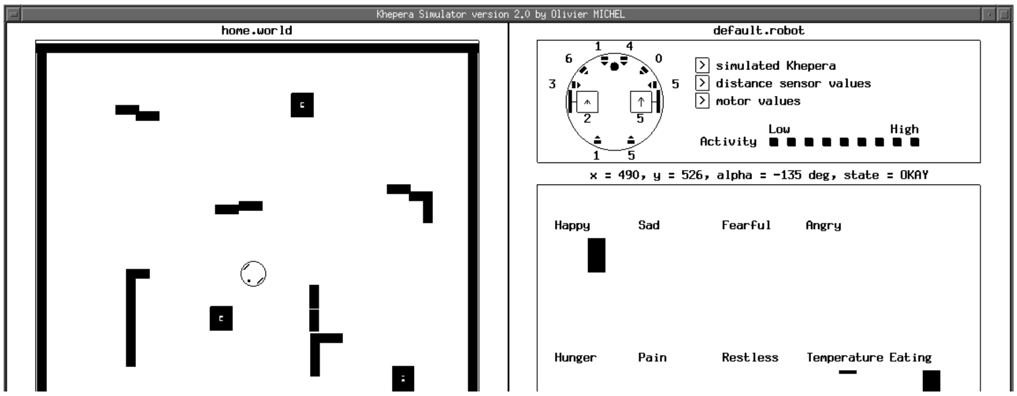
PAIN — is active if the robot is bumping into obstacles.

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] See, for example, (Hertz et al., 1991) for a full description.

[5] The lights had to be surrounded by bricks to avoid the robot becoming permanently stuck to them because of a deficiency in the simulator. The lights can still be perceived by the robot as the bricks are transparent to light.



☐ simulated Khepera
☐ distance sensor values
☐ motor values

```
x = 490, y = 526, alpha = -135 deg, state = OKAY
```