Exploring the Role of Emotions in Autonomous Robot Learning

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

Abstract

Autonomy is a very important property for a robot to have, yet implementing it in a robot is far from trivial, particularly when one requires the meaning of autonomy to include self-motivation, instead of mere automaticity.

The fact that emotions are considered to be essential to human reasoning and human motivation in particular, suggests that they might play an important role in robot autonomy. The purpose of the work reported here is to know if and how emotions can help a robot in achieving autonomy.

Experimental work was done in a simulated robot that adapts to its environment through the use of reinforcement learning. Results suggest that emotions can be useful in dividing the task in smaller manageable problems by focusing attention on the relevant features of the task at any one time.

Introduction

In the field of robotics, the criteria used to define whether a robot is autonomous or not are not well established. In general, simply requiring that, once the robot is finished, it does its task without human intervention is enough. The word's root meaning suggests an alternative definition of autonomy that has stronger requirements. Namely, a truly autonomous robot should also develop the laws that govern its behaviour. To accomplish this, a robot should have an adaptive controller that improves its performance by unsupervised learning when interacting with its environment. Although true autonomy is still a very open research subject, it is advantageous to any agent that has to deal with unexpected situations. Arguments in favour of true autonomy have been put forward in several diverse fields, for example robotics, animal robotics[1] (McFarland 1994), agents theory (Ferguson 1992) and interactive virtual environments (Blumberg 1995).

In robotics, emotions are often used to modulate activity in a fixed controller (Cañamero 1997; Bates, Loyall, & Reilly 1992a). The social role of emotions has been particularly explored. The external demonstration of emotions has been used as a sort of communication mechanism that allows the robot to report to others its internal state (e.g. its level of task achievement (Shibata, Ohkawa, & Tanie 1996)) or makes the robot capable of generating empathy emotions in people, by creating an illusion of life in a believable character (Bates 1994).

The present research focuses on how to use emotions in the control of an autonomous simulated robot that adapts to its environment using reinforcement learning. 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). Yet, the developed model is based on a simplified hormone system and is far from the complexity of true emotions experienced by humans. Experiments were carried out on a simulated Khepera robot in an animal-like adaptation task.

In the next section, a detailed description of the emotion model developed is presented. The experiments done with this model are reported in the following section.

Emotions

People like to have a Cartesian approach to life in thinking that all their reasoning is purely logical and emotions are purely disruptive. In fact, individuals do not always make rational choices (Grossberg & 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 cur-

[1] Modelling of animal behaviour using robots.

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

rent 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.

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 for focusing attention on 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.

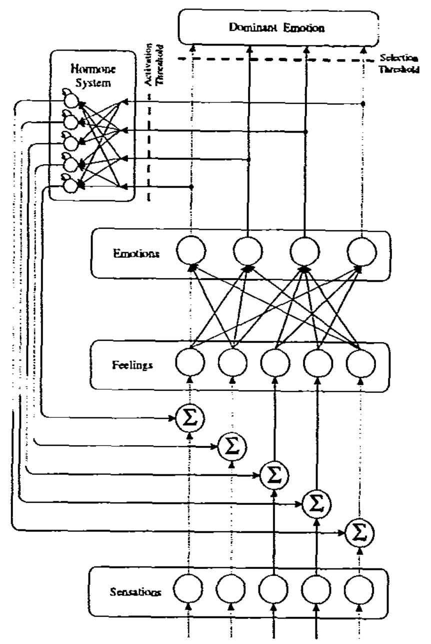

Figure 1: Emotions model.

Model

A large subset of theories of emotions is based on cognitive appraisal theories (Lazarus 1982; Power & 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, Beaudoin, & Wright 1994; Pfeifer 1982; Pfeifer & Nicholas 1985; Bates, Loyall, & Reilly 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 this kind of approach, a bottom-up approach was followed here.

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. The emotions chosen are also usually included in the definitions of primary emotions (Shaver *et al.* 1987; Power & Dalgleish 1997), which is a good indicator of their relevance and need.

The model determines the intensity of each emotion based on the robot's current internal feelings. These feelings are: Hunger, Pain, Restlessness, Temperature, Eating, Smell, Warmth and Proximity. 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 resulting 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 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.

In summary, the model of emotions described provides not only an emotional state coherent with the current situation, but also influences the body perception. In order to evaluate the role of emotions in reasoning, this state should be used for the actual control of the robot, determining its behaviour (Albus 1990; Wright 1996; Moffat, Frijda, & Phaf. 1993). The next section describes the experiments done in this direction.

Experiments

The Simulated Robot's Task

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 is easier to think of emotions in the context of an animal-like creature with self-maintainance needs.

The robot's task consists in collecting energy from food sources scattered throughout the environment. These food sources are actually lights so that the robot is able to distinguish them with its poor perception capabilities. The robot needs this energy to use during its functioning. It will use up energy faster if the velocity it demands from its motors is higher.

To gain energy from a food source, the robot has first to bump into it. If the food source still has energy left, it will make some of it available to the robot for a short period of time. At the same time an odour will be released that can be sensed by the robot. During this short period, the robot can acquire energy by receiving high values of light in its rear light sensors. This means that the robot must turn its back to the food source. To receive more energy the robot has to restart the whole process again by hitting the light again so that a new time window of released energy is started.

The robot can only extract a limited amount of energy from each food source. In time, the food source will

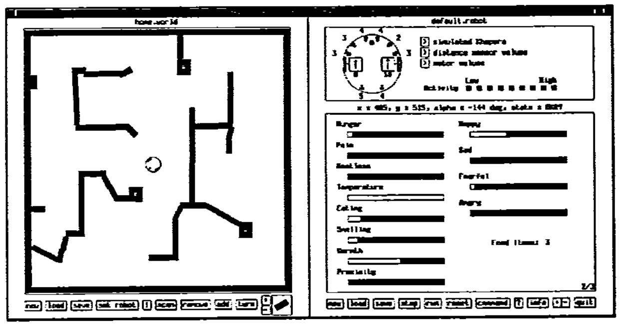

Figure 2: The robot and its environment.

recover its ability to provide energy again, but meanwhile the robot will be 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 the food source has no energy, the light associated with it is turned off.

The robot's 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. Furthermore, the robot should not keep still in the same place for long durations of time or collide with obstacles.

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

- 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 getting new energy at the moment.
- If the robot has very low energy and it is not acquiring new energy, then its state will be **sad**.
- If the robot bumps into the walls then the pain will make it **fearful**.
- 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. If the action is changed the value of restlessness is reset by the system.

All the experiments were carried out in a Khepera simulated robot (Michel 1996) within a closed environment divided by several walls and containing a few lights (see figure 2).

The Adaptive Controller

Reinforcement learning techniques have already been successfully used in the field of robotics and were therefore selected for the learning algorithm. The main problem with reinforcement learning is that learning can be very slow, particularly if the task is very complex. Yet, behaviour decomposition of the task can reduce significantly the learning time or even make the task feasible. Behavioural decomposition usually consists in learning some predefined behaviours in a first phase and then finding the high-level coordination of these behaviours. Although the behaviours themselves are often learned successfully (Mahadevan & Connell 1992; Lin 1993), behaviour coordination is much more difficult and is usually hard-wired to some extent (Mahadevan & Connell 1992; Lin 1993; Mataric 1994). One problem in particular which is quite difficult and task dependent is determining when to change behaviour. This is not a problem in traditional reinforcement leaning where agents live in grid worlds and state transition is perfectly determined. Yet, in robotics, agent states change asynchronously in response to internal and external events and actions take variable amounts of time to execute (Mataric 1994). The design of a reward function can also be a problem if there are multiple goals and immediate reinforcement is not always available. In this case, it is often impossible to have a direct translation to a traditional monolithic reward function (Mataric 1994). Both these previous problems can be found in the task at hand. Furthermore, unlike traditional reinforcement learning tasks, the task of an autonomous robot is mainly one of continuously executing a task for as long as necessary in opposition to successfully completing a task and finishing. Another major difference is that the distinction between a learning phase and a performing phase had to be eliminated, because an autonomous robot is supposed to continuously adapt to its environment.

In our work we have chosen to have the primitive behaviours hand-designed and learn only the behaviour coordination in the hope that emotions might be useful in solving some of the problems discussed previously. Three primitive behaviours were hand-designed:

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 behaviour by itself is not very reliable in that the robot can crash. Yet, the avoid obstacles behaviour can easily help in these situations.

The developed controller has two separate modules:

Associative Memory Module — This plastic module uses feed-forward networks to associate the robot feelings with the current expected value of each one of the three robot behaviours. Q-learning (Watkins 1989) was used in an implementation very similar to the one reported by Lin (Lin 1992). Neural networks learned, by back-propagation, utility functions that model $util(x,a) = R + \gamma eval(y)$, *i.e.* the immediate reinforcement (R) plus the discount factor (γ, constant set to 0.9) times the expected cumulative discounted reinforcement ($eval(y)$) starting from state y reached by executing behaviour a in state x.

Behaviour Selection Module — Taking into account the value attributed to each behaviour by the previous module, this module makes a stochastic selection based on the Boltzmann-Gibbs distribution of the behaviour to execute next.

The Role of Emotions

From what was discussed previously in this document, both in terms of emotions and of the reinforcement learning paradigm, here are some possible roles for emotions:

Reinforcement value — The design of a reward function is one of the most critical aspects for the success of a reinforcement learning task. In robotics, the role of providing an evaluation of the state of the world is often attributed to emotions, *e.g.*, (Albus 1990; Wright 1996). It is often assumed that human decision making consists in the maximization of positive emotions and minimisation of negative emotions, *e.g.*(Tomkins 1984). It should be expected that a reward function directly obtained from the emotional state would work well. This can easily be accomplished with our model. At any moment in time, a value judgement can be obtained from the emotion system by considering the intensity of the current dominant emotion and whether it is positive or negative. Experiments were done that consisted in using as reinforcement the intensity of the current dominant emotion or zero if there was no dominant emotion. If the dominant emotion was a negative one then its (positive intensity) value would be negated. This reinforcement function proved to be successful in the experiments described in this paper. Yet, in previous experiments (Gadanho & Hallam 1998a) this reinforcement function was not successful at all. The main difference between the experiments was that the first were done with an action-based adaptive controller that selected a new action in each time step while the present experiments use a behaviour-based controller. We believe that the time-scales involved in the execution of a behaviour are more adequate for emotion-dependent reinforcement and that the small time-scale associated with the execution of a single action is to blame for the previous failure. The experiments also showed that a more traditional monolithic reward function directly derived from the robot's sensations can also be used, if the state transition is triggered by emotions (see below).

Determining state — Another mechanism that was tested was having emotions influencing the robot perception. When the robot learns associations between states and rewards through its neural networks, is actually using feelings to determine state by using them as network inputs. In the emotion model developed, feelings are influenced by emotions through the hormone system. So the robot state is emotion dependent. What the robot learns to associate with rewards is actually being distorted by emotions. It is being changed to be more compatible with the active emotions, thus making the relevant features of the environment more evident. For the moment no significant differences were found in using this kind of perception or using an alternative perception that is not influenced by hormones. This shows that the robot can cope with a distorted view of reality, but does not show that emotions can be useful in this domain. Yet, the fact that no differences were found might be purely task dependent.

Triggering state transition — In a robotic environment, a new state can be found for virtually every step. The perception of the world will always be at least slightly different from step to step due to noise. Nevertheless, making a re-evaluation of the system every step by selecting a new behaviour and performing an evaluation on the previous behaviour is not wise. It is both a computational waste and an hindrance to successfully learning the advantages of each one of the behaviours. On the other hand, if the behaviours are left running for too long, events may occur that will make them inadequate for the new situation. 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 summarised vision of the environment. Any important change in the environment is liable to be captured by changes in the emotional state. Experiments showed that emotions could fulfill this role successfully (Gadanho & Hallam 1998b).

Conclusions

The model of emotions behaved appropriately when tested on the animat, 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 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, because it is not considered cognitively interesting, yet without it emotions lack their characteristic inertia (Moffat, Frijda, & Phaf. 1993).

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

- providing a straightforward reinforcement function which works like a powerful attention mechanism in a reinforcement learning task by attributing value to the different environmental situations;
- making more evident the relevant aspects of the environment, *i.e.* those directly related with the current emotional state, by influencing the robot current state through the hormones;
- determining the occurrence of the significant changes in the environment that should trigger state transition, by looking at sudden changes in the emotional system state.

These three different mechanisms all worked well experimentally.

Our ultimate intention is to scale up our architecture to develop more complex emotions on the top of the ones described, but that is not the primary focus of this work.

Acknowledgements

The first author is supported by a scholarship from the Portuguese programme PRAXIS XXI. Facilities for this work were provided by the University of Edinburgh. Thanks to Nuno Chagas for helpful comments and suggestions on an earlier version of this paper.

References

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

Bates, J.; Loyall, A. B.; and Reilly, W. S. 1992a. An architecture for action, emotion, and social behavior. In *Artificial social systems: Fourth European workshop on modeling autonomous agents in a multi-agent world.* Elsevier.

Bates, J.; Loyall, A. B.; and Reilly, W. S. 1992b. Integrating reactivity, goals, and emotion in a broad agent. CMU-CS 92-142, Carnegie Mellon University, School of Computer Science.

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

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

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.* The ACM Press.

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

De Sousa, R. 1987. *The rationality of emotion.* The MIT 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.* Ph.D. Dissertation, University of Cambridge.

Gadanho, S. C., and Hallam, J. 1998a. Emotion-driven learning for animat control. In *From animals to animats 5. Proceedings of the fifth international conference of the society for adaptive behavior 98.* The MIT Press.

Gadanho, S. C., and Hallam, J. 1998b. Emotion-triggered learning for autonomous robots. DAI Research Paper 916, University of Edinburgh. Submitted to the *SAB'98 Workshop on Grounding Emotions in Adaptive Systems.*

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.

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

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.* Ph.D. Dissertation, Carnige Melon University. Tecnical 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 *Proc. of Conference on Machine Learning - 1994*, 181–189.

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

Michel, O. 1996. Khepera simulator version 2.0. 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., ed., *Prospects for Artificial Intelligence.* IOS Press. 219–228.

Panksepp, J. 1982. Toward a general psychobiological theory of emotions. *The behavioural and brain sciences.*

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

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

Plutchick, R. 1984. Emotions: A general psychoevolutionary theory. In Scherer, K. R., and Ekman, P., eds., *Approaches to emotion.* London: Lawrence Erlbaum.

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

Scherer, K. R. 1984. On the nature and function of emotion: A component process approach. In Scherer, K. R., and Ekman, P., eds., *Approaches to emotion.* London: Lawrence Erlbaum.

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.; Beaudoin, L.; and Wright, I. 1994. Computational modeling of motive-management processes. In Frijda, N., ed., *Proceedings of the Conference of the International Society for Research in Emotions*, 344–348. Cambridge: ISRE Publications. Poster.

Tomkins, S. S. 1984. Affect theory. In Scherer, K. R., and Ekman, P., eds., *Approaches to emotion.* London: Lawrence Erlbaum.

Watkins, C. 1989. *Learning from delayed rewards.* Ph.D. Dissertation, 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 *From animals to animats 4.* The MIT Press.

Zajonc, R. B. 1984. On primacy of affect. In Scherer, K. R., and Ekman, P., eds., *Approaches to emotion.* London: Lawrence Erlbaum.