Using Emotions for Behaviour-Selection Learning

Maria Malfaz [1] and **Miguel A. Salichs** [2]

Abstract. Emotions play a very important role in human behaviour and social interaction. In this paper we present a control architecture which uses emotions in the behaviour selection process of autonomous and social agents. The state of the agent is determined by its internal state, defined by its dominant motivation, and its relation with the external objects including other agents. The behaviour selection is learned by the agent using standard and multiagent Q-learning algorithms. The considered emotions are fear, happiness and sadness. The role of these emotions in this architecture is different, while the learning algorithms use happiness/sadness of the agent as positive/negative reinforcement signals, the emotion fear is used to prevent the agent of choosing dangerous actions as well as a motivation.

1 INTRODUCTION

In previous works [7] an emotion-based architecture was proposed for an autonomous and social robot. The use of emotions in this architecture is oriented to the behavior selection process rather than the human-robot interaction issue. Several authors such as Canamero [3] and Gadanho [5] have used emotions to improve the adaptation of the robot to the environment and therefore the autonomy of the robot. Many others such as Breazeal [2], Fujita [4], Shibata et al [11] have implemented emotional models on robots to enhance the human-robot interaction. Their robots, Kismet and Leonardo, AIBO and Necoro, include the possibility of showing emotions, by facial and sometimes body expressions. We intend to make use of emotions in robots trying to imitate their purpose in nature, which includes, but is not limited to, interaction.

We think that the role that each emotion plays, and how the mechanisms associated to each one work are very specific. That means that each emotion must be incorporated to the robot in a particular way. In a recent paper [10] we have shown how some emotions such as happiness, sadness and fear are used to learn the right behaviour policy for each situation.

2 CONTROL ARCHITECTURE

The control architecture proposed here, see Fig. 1, has been designed for an autonomous robot. An autonomous agent/robot should be capable of determining its goals, and it must be capable of selecting the most suitable behaviour in order to reach them. Similarly to other authors [3], [2], [12], our agents autonomy relies on a motivational model.

[1] Carlos III University of Madrid, Spain, email: mmalfaz@ing.uc3m.es
[2] Carlos III University of Madrid, Spain, email: salichs@ing.uc3m.es

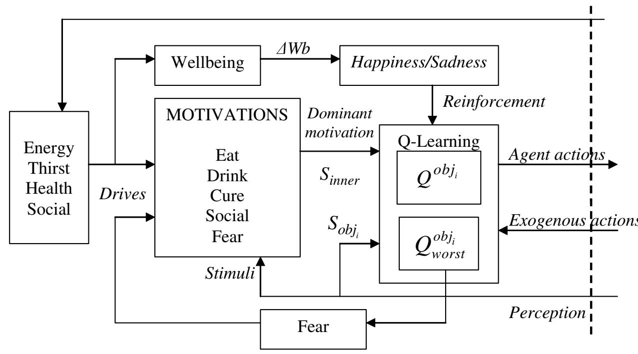

Figure 1. Control Architecture

2.1 Drives and Motivations

Motivations can be seen as homeostatic processes, which maintain a controlled physiological variable within a certain range. Homeostasis means maintaining a stable internal state [1]. The drives constitute urges to action based on bodily needs related to self-sufficiency and survival.

In order to model motivations, the hydraulic model of motivation described in [6] has been used as an inspiration. In this model, internal drive strength interacts with external stimulus strength. The intensity of motivations is a combination of the intensity of the related drive and the presence of the related external stimuli. The external stimuli are the different objects and agents that the player can find in the virtual world during the game. According to our model a motivation can obtain a high intensity due to two reasons: 1) the correspondent drive is the highest or 2) The correct stimulus is present. The dominant motivation will be the one with the highest intensity and the possibility of no dominant motivation exists.

2.2 Wellbeing

As it is shown in Fig.1 the wellbeing of the agent is a function of the values of the drives. In [8] the wellbeing is defined in such a way that it is maximum when all the needs (drives) of the agent are satisfied. There exist some personality factors that weight the importance of the values of the drives on the wellbeing of the agent.

On the other hand, based on the definitions of the emotions given in [7], the emotions of happiness and sadness depends on the variation of the wellbeing of the agent. When the wellbeing of the agent is increased a certain amount, happiness arises, on the contrary, sadness is produced when the wellbeing decreases a certain amount.

2.3 Behaviour Selection

In this architecture the agent learns, using different reinforcement learning algorithms (RL), the best behaviour at each step using the emotions of happiness/sadness as the reward. This use of emotions as reinforcement has been proposed previously by Rolls [9]. Therefore, in this architecture behaviours are not selected to satisfy the goals determined by the dominant motivation but to optimize the wellbeing of the agent. This implies that the final goal of the agent is to reach Happiness and avoid Sadness.

In the proposed architecture the agent will use the Q-learning algorithm as the RL algorithm when the agent is not interacting with other players. In the case of social interaction, the agent uses a Q-multiagent RL algorithm.

The state of the agent is the aggregation of his inner state, the dominant motivation S_{inner}, and the states S_{obj} related to each of the external objects, including external agents, which can interact with him. For the RL algorithms the objects are considered as independent. This means that the state of the agent in relation with each object is $S \epsilon S_{inner} \times S_{obj_1}$.

2.4 Fear

Fear is produced when the agent knows that something bad may happen. In this architecture there are two types of fear, one related to actions executed by the agent and the other related to exogenous actions carried out by other elements of the environment such as other agents.

2.4.1 To be afraid of executing risky actions

In order to avoid risky actions, the worst result experimented by the agent for each pair action-state is stored in a variable called $Q^{obj_i}_{worst}$. The effect of being afraid can be considered by following a policy that considers not only the average of the result obtained but also the worst one. We define a daring factor that measures the daring degree of the agent. By varying this parameter the agent gives more or less importance to the worst result obtained making the agent fearful or fearless respectively.

2.4.2 To be afraid of risky exogenous actions

When the agent may suffer some negative effects in a state as a consequence of exogenous events, feels fear. ”Fear” is expressed as a drive D_{fear}. The fear drive is equally treated as the rest of drives, and its related motivation could be the dominant one. In this case, the agent will learn by itself what to do when it is afraid.

3 EXPERIMENTS

This control architecture has been tested on virtual MUD players (a text-based multi user game), who ”live” in a virtual world. This game gave us the possibly of creating different 2-D environments to play in. This environment is fully described in [8].

The results obtained in [8] showed that using happiness and sadness as the reinforcement signal of the RL algorithms the agent learns the correct policy in order to survive in his virtual world. Moreover, the quality of life of the agent is quite good following the learned policy.

In recent experiments it has been proved that following a near fearless policy that consider the optimal policy as well as the worst expected results in some degree, improves the performance of the agent.

4 CONCLUSIONS AND FUTURE WORKS

This control architecture implements emotions on the learning behaviour process of an artificial agent, obtaining quite good results in relation with the overall performance. The learning algorithms use happiness/sadness of the agent as positive/negative reinforcement signals. Fear is used to prevent the agent choosing dangerous actions or being in dangerous states where non-controlled exogenous events, produced by external objects or other agents, could danger him.

Using this control architecture the agent learns behavioural patterns, this means that the agent learns the sequence of behaviours to be executed in order to satisfy the drive related to the dominant motivation. The agent learns, for example, that he has to get food when he is hungry, even when this behaviour is was not previously linked with the ”Eat” motivation.

The final goal of the project is to implement this architecture on a real personal robot. Meanwhile in the next future, another emotion is going to be implemented: Anger. This emotion will be produced when another active object, such as other agent, could reduce the expectancies of the wellbeing of the agent.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funds provided by the Spanish Government through the projects named ”Personal Robotic Assistant” (PRA) and ”Peer to Peer Robot-Human Interaction” (R2H), of MEC (Ministry of Science and Education).

REFERENCES

[1] Kent C. Berridge., ‘Motivation concepts in behav-ioural neuroscience’, *Physiology and Behaviour*, (81), 179–209, (2004).
[2] C. Breazeal, *Designing Sociable Robots*, The MIT Press, 2002.
[3] L. Caamero, ‘Modeling motivations and emotions as a basis for intelligent behavior’, in *First International Symposium on Autonomous Agents (Agents’97), 148-155. New York, NY: The ACM Press.*, (1997).
[4] Masahiro Fujita, ‘Aibo: Toward the era of digital creatures’, *The International Journal of Robotics Research*, (vol 20, n 10), 781–794, (October 2001).
[5] Sandra Gadanho, *Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions*, Ph.D. dissertation, UNiversity of Edinburgh, 1999.
[6] K. Lorentz and P. Leyhausen, *Motivation of human and animal behaviour; an ethological view*, volume xix, New York: Van Nostrand-Reinhold, 1973.
[7] M. Malfaz and M.A. Salichs, ‘A new architecture for autonomous robots based on emotions’, in *Fifth IFAC Symposium on Intelligent Autonomous Vehicles. Lisbon. Portugal*, (July 2004).
[8] M. Malfaz and M.A. Salichs, ‘Learning behaviour-selection algorithms for autonomous social agents living in a role-playing game’, in *AISB’06: Adaptation in Artificial and Biological Systems. University of Bristol, Bristol, England*, (April 2006).
[9] Edmund Rolls, *Emotions in Humans and Artifacts*, chapter Theory of emotion, its functions, and its adaptive value, MIT Press, 2003.
[10] M.A. Salichs and M. Malfaz, ‘Using emotions on autonomous agents. the role of happiness, sadness and fear’, in *AISB’06: Adaptation in Artificial and Biological Systems. University of Bristol, Bristol, England*, (2006).
[11] Tanie K. Shibata T., Tashima T., ‘Emergence of emotional behaviour through physical interaction between human and robot’, in *the 1999 IEEE International Conference on Robotics and Automation (ICRA’99)*, (September 1999).
[12] J. Velasquez, ‘An emotion-based approach to robotics’, in *I999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (1999).