

DESIGN OF AN ARCHITECTURE BASED ON EMOTIONS FOR AN AUTONOMOUS ROBOT

María Malfaz, Miguel A. Salichs

Carlos III University of Madrid
28911 Leganés, Madrid, Spain
mmalfaz@ing.uc3m.es, salichs@ing.uc3m.es

Abstract

With the emergence of personal robots recently the need to create an easier form of human-robot interaction has become apparent. Emotions will play an important role in this interaction. In the proposed architecture emotions are generated from the evaluation of the wellbeing of the robot. Moreover, behavior selection mainly depends on a specific drive and emotions have a supervisory role.

Proposed Architecture

The goal of our research is to construct autonomous personal robots. These robots are being built on mobile platforms (iRobot Magellan Pro), on which various sensors (laser, sonar and camera) and devices are installed. These sensors and devices will make possible the interaction of the robot with its environment and users.

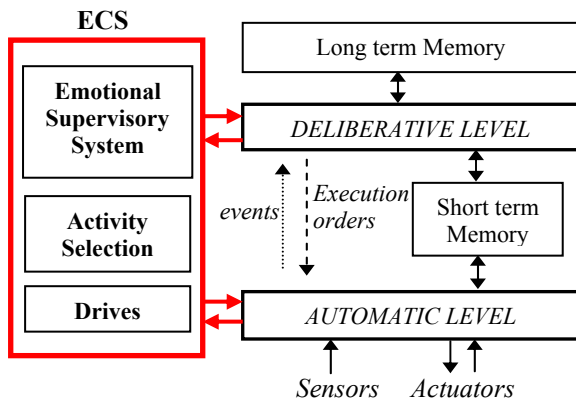


Fig. 1. AD Architecture with an emotional control system.

Figure 1 illustrates how an emotional control system (ECS) has been added to the AD architecture (Barber and Salichs 2001). Inside the ECS we can observe three different modules: Drives, Activity Selection and Emotional Supervisory System (ESS). The Drives module is the one that controls the basic drives of the robot. The Activity Selection module on the other hand, determines goals and action tendencies of the robot. Finally, the ESS module generates the emotional state of the robot.

Drives

Several authors, such as Frijda (Frijda 1987), propose that an independent system should not have to wait for someone to maintain, succour, and help it. Therefore, an autonomous agent should be capable of determining its goals, and it must be capable of selecting the most suitable behaviour in order to reach its goals. Similarly to the Velasquez's proposal (Velasquez 1998), the robot's autonomy relies on a motivational model based on drives and emotions. Within the model developed by Velasquez, drives are the robot's internal needs. These drives are related to the feedback of internal variables such as its battery and temperature. By contrast, within our control architecture the drives are defined in a different way. They are related to the feedback of internal variables (e.g. battery level) as well as external variables (e.g. social interaction).

In this architecture the following drives are initially used:

- Energy: With the battery level as its input.
- Entertainment: Depending on the activity and human-robot interaction.
- Affect: Depending on the level of affection received through the human-robot interaction.
- Health: Related to physical damage.
- Agenda: This drive reflects the motivation of the robot to carry on with activities which have been programmed previously.

As it is shown, social motivations are included as robot's needs. The Entertainment and Affect drives are similar to the Stimulation and Social drives defined by Breazeal (Breazeal 2002).

Emotions

Most of authors working with emotion-based control architectures, such as Velasquez (Velasquez 1998), Gadanho (Gadanho 2002), Breazeal (Breazeal 2002), etc., give a basic role to emotions i.e. behaviour selection is conditioned by emotions. Nevertheless, according to several authors such as Ortony (Ortony, Clore and Collins

1988) and Frijda (Frijda 1987), not all behaviours are emotional, there exist non emotional behaviours, i.e. we can avoid a chair without being afraid of it. Following this way of thinking, in our control architecture behaviour selection mainly depends on a specific drive, the dominant drive, and emotions have a supervisory role.

When there is not any active emotion, the dominant drive will decide the goal of the robot, and therefore it will determine which behaviour will become active in order to reach the goal, i.e. if the Energy drive is low, then the goal of the robot will be to recharge itself, so the behaviour “find the power station” will become active. Only when an emotion becomes active, this emotion can influence the Activity Selection module.

The emotions used in this architecture are: Happiness, Anger, Fear and Sadness. These emotions have persistence in time. This classification is the same used by other authors (Gadanhó and Hallam 2001).

Emotion generation. In this architecture emotions are generated from the evaluation of the wellbeing of the robot. Happiness is produced because something good has happened, i.e. an increment of the wellbeing is produced. On the contrary, Sadness is produced because something bad has happened, so the wellbeing decreases. Fear appears when the possibility of something bad is about to happen. In this case, we expect that the wellbeing drops off. Finally, Anger can be produced either because we expected something good to happen (an increase in the wellbeing) and something made it impossible, or because something produced something bad to happen (a decrease in the wellbeing). These definitions of emotions are based on Ortony’s works (Ortony, Clore and Collins 1988).

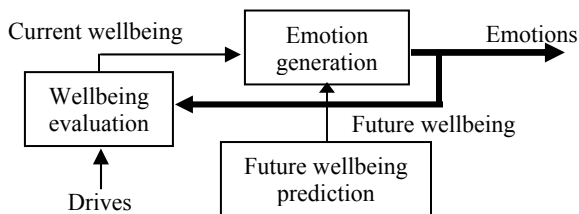


Fig. 2. Emotional Supervisory System

As it is shown in Figure 2, in order to generate these emotions we evaluate two signals: the current wellbeing and the prediction of the future wellbeing. The evaluation of the past wellbeing signal would also produce emotions such as shame, pride, gratitude, etc. Nevertheless, these kinds of emotions will not be considered by now in our architecture.

The intensity of each emotion is related to the derivate of the current or future wellbeing signal. Depending on the absolute value of the slope of the wellbeing signal, the intensity of the correspondent emotion will vary.

The current wellbeing signal is formed by two contributions. The first one is a general appraisal of the drives, and the second one is the emotion influence. The future wellbeing signal can be obtained from the analysis of the consequences of the current situation.

The functions of the ESS in our control architecture are as follows:

- To establish general goals: i.e. happiness. These global goals allow the architecture to develop behaviour selection mechanisms and a global appraisal of the situation. This global appraisal can be used also as reinforcement in learning processes.
- Supervision: When the changes of the wellbeing signal (current or future) are greater than a certain threshold, one emotion will become active. In this sense, the ESS will act as an alarm system where emotions are the alarm signals.

Acknowledgements

The authors gratefully acknowledge the funds provided by the Spanish Government through the DPI2002-00188 project, of MCyT (Ministerio de Ciencia y Tecnología).

References

- Barber, R., Salichs, M.A. 2001. A new human based architecture for intelligent autonomous robots. In *The Fourth IFAC Symposium on Intelligent Autonomous Vehicles*, p85-90. Sapporo, Japan.
- Breazeal C. 2002. *Designing Sociable Robots*. The MIT Press.
- Frijda, N. and Swagerman, J. 1987. Can computers feel? Theory and design of an emotional model. *Cognition and Emotion*. 1 (3). 235-357.
- Ortony, A., Clore, G. L., and Collins, A. 1988. *The Cognitive Structure of Emotions*. Cambridge University Press. Cambridge, UK
- Gadanhó S. and Hallam J. 2001. Emotion-triggered learning in Autonomous robot control. *Cybernetics and Systems: An International Journal* 32:531-559.
- Gadanhó S. 2002. Emotional and Cognitive Adaptation in Real Environments. In *Symposium ACE'2002 of the 16th European Meeting on Cybernetics and Systems Research*, Vienna, Austria.
- Velásquez, J 1998. When Robots Weep: Emotional Memories and Decision Making. In *Proceedings of AAAI-98*.