Using Emotions on Autonomous Agents. The Role of Happiness, Sadness and Fear.

Miguel Angel Salichs Maria Malfaz
RoboticsLab, Carlos III University of Madrid
28911 Leganés, Madrid, Spain
salichs@ing.uc3m.es mmalfaz@ing.uc3m.es

Abstract

This paper addresses the use of emotions on autonomous agents for behaviour-selection learning, focusing in the emotions fear, happiness and sadness. The control architecture is based in a motivational model, which performs homeostatic control of the internal state of the agent. The behaviour-selection is learned by the agent using a Q-learning algorithm while there is no interaction with other agents. In situations where interaction arises (e.g. interacting with other agents), agents rely on stochastic games approaches as a learning strategy. The agent is intrinsically motivated and his final goal is to maximize Happiness. 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. Preliminary tests have been carried out in a virtual world, based in a role-playing game.

1 Introduction

The goal of our project is to develop social robots with a high degree of autonomy. The social aspect of the robot will be reflected in the fact that the human interaction will not be considered only as a complement of the rest of the robot's functionalities, but as one of the basic features.

For this kind of robots, the autonomy and emotions makes them to behave as if they were "alive". This feature would help people to think about these robots not as simple machines but as real companions. Evidently, a robot that has his own "personality" is much more attractive than one that simply executes the orders that he is programmed to do.

Emotions can act as a control and learning mechanism, driving behaviour and reflecting how the robot is affected by, and adapts to, different factors over time (Fong et al, 2002). In previous works (Malfaz and Salichs, 2004), an emotion-based architecture has been proposed.

Some researchers have also used emotions in robots. Most of them have made emphasis in the external expression of emotions (Breazeal, 2002) (Fujita, 2001) (Shibata et al 1999). Their robots include the possibility of showing emotions, by facial and sometimes body expressions. In this case, the emotions can be considered just as a particular type of information that is exchanged in the human-robot interaction process. In nature emotions have different purposes and interaction is only one of them. We intend to make use of emotions in robots trying to imitate their purpose in nature, which includes, but is not limited to, interaction. The role that plays each emotion 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 this paper we will present some basic ideas on how emotions such as happiness, sadness and fear can be used in an autonomous robot.

Emotions will be 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 its 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 is produced when a decrement of the wellbeing of the robot happened due to another-initiated act.

This paper presents a control architecture for an autonomous agent based on motivations. The agent uses reinforcement learning algorithms to learn its policy while interacts with the world. The reward for these learning algorithms will be the variation of the wellbeing of the agent (happiness/sadness) due to the previous selected behaviour, calculated at each step of the process. This wellbeing is a function of the internal needs of the agent (drives). This idea of using the wellbeing of the agent as the reinforcement in the learning process for behaviour selection has been also used by Gadanho in the ALEC architecture, obtaining quite good results (Gadanho, 2003).

The remainder of the paper is organized as follows. Section 2 introduces the use of emotions in robots. Section 3 and 4 describe the proposed control architecture and the reinforcement learning algorithms respectively. Section 5 introduces the emotion fear and section 6 describes the experimental setting. Finally, conclusions and future works are summarized in section 7.

2 Emotions in robotic

One of the main objectives in robotics and artificial intelligence research is to imitate the human mind and behaviour. For this purpose the studies of psychologists on the working mind and the factors involved in the decision making are used. In fact, it has been proved that two highly cognitive actions are dependant not only on rules and laws, but on emotions: Decision making and perception (Picard, 1998). In fact, some authors affirm that emotions are generated through cognitive processes. Therefore emotions depend on ones interpretation, i.e. the same situation can produce different emotions on each agent, such as in a football match (Ortony, 1988). Moreover, emotions can be considered as part of a provision for ensuring and satisfaction of the system's major goals (Frijda, 1987).

Emotions play a very important role in human behaviour, communication and social interaction. Emotions also influence cognitive processes, particularly problem solving and decision making (Damasio, 1994). In recent years, emotion has increasingly been used in interface and robot design, primarily in recognition that people tend to treat computers as they treat other people.

There are several theories about emotions (Frijda 1987; Ortony, 1988; Sloman, 2003; Rolls, 2003), but the results of Damasio (1994) can be considered the basis, for many A.I. researchers, to justify the use of emotions in robotics and their computation.

Rosalind Picard in her book Affective Computing (1998), writes a complete dissertation about this subject based on several psychologists, including Damasio. Picard (1998) proposed a design criterion in order to create a computer that could express emotions. Moreover, she established that a computer has emotions if it has certain components that are present on the emotional systems of healthy people. Picard (2003) expounded four motives for giving certain emotional abilities to machines: The first goal is to build robots and synthetic characters that can emulate living humans and animals, such as a humanoid robot. The second is to make machines that are intelligent. A third objective is to try to understand human emotions by modelling them. Although these three goals are important, the main one is to make machines less frustrating to interact with, i.e. to facilitate the human-machine interface.

Cañamero (2003) considers that emotions, or at least a sub-group of them, are one of the mechanisms founded in biological agents to confront their environment. This creates ease of autonomy and adaptation. For this reason she considers that it could be useful to exploit this role of emotions to design mechanisms for an autonomous robot. Emotions are used as mechanisms that allow the agent (robot) to:

1. Have fast reactions.
2. Contribute to resolve the selection among multiple objectives.
3. Signal important events to others.

Bellman (2003) agrees, to some degree, with Cañamero and her reasons for considering emotions in robotics. The author states that emotions allow animals with emotions to survive better than the others without emotions. Therefore, we can presume that some type of analogy to emotional abilities is required within robots, if we want an intelligent and independent behaviour within a real environment.

Changing subject, Picard (2003) gives an advice about the implementation in machines of functions implemented by the human emotional system. Computers do not have emotions as human beings in any natural experimentation sense. Science methodology is to try to reduce complex phenomena, such as emotions, to a functional requirements list. The challenge of many computing science researchers is to try to duplicate these in computers at different levels depending on the motives of the investigation. But we must be careful when presenting this challenge to the general public, who may perceive that emotions are the frontier that separates man and machine

3 Control Architecture

An independent system should not have to wait for someone to maintain, succour, and help it (Frijda and Swagerman,1987). 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 other authors (Avila-Garcia and Cañamero, 2004), (Breazeal, 2002), (Gadanho, 2003), (Velasquez, 1998), our agent's autonomy relies on a motivational model. Figure 1 shows this proposed control architecture for behaviour selection.

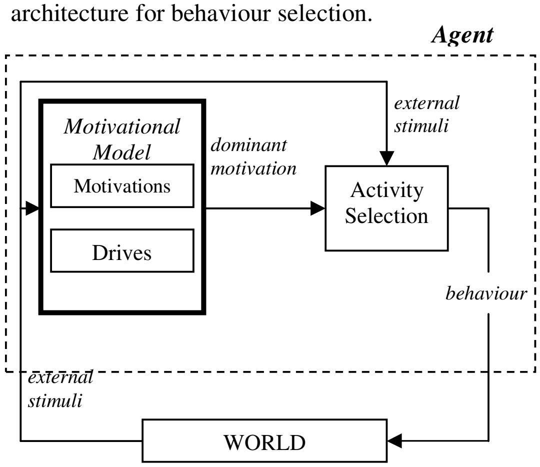

Figure 1: Control architecture for autonomous agents

3.1 Motivational Model

Motivations can be seen as homeostatic processes, which maintain a controlled physiological variable within a certain range. Homeostasis means maintaining a stable internal state (Berridge, 2004). This internal state can be parameterized by several variables, which must be around an ideal level. When the value of these variables differs from the ideal one, an error signal occurs: the drive. These drives constitute urges to action based on bodily needs related to self-sufficiency and survival. External stimuli, both innate and learned, are also able to motivate and drive behaviour (Cañamero, 1997).

In order to model motivation, the hydraulic model of motivation described by Lorentz and Leyhausen in (Lorentz and Leyhausen, 1973) has been used as an inspiration. This model is essentially a metaphor that suggests that motivational drive grows internally and operates a bit like pressure from a fluid reservoir that grows until it bursts through an outlet. Motivational stimuli from the external world act to open an outflow valve, releasing drive to be expressed in behaviour. In this model, internal drive strength interacts with external stimulus strength. If drive is low, then, a strong stimulus is needed to trigger motivated behaviour. If the drive is high, then, a mild stimulus is sufficient (Berridge, 2004). Following this idea, the intensity of motivations (M_i) is a combination of the intensity of the related drive (D_i) and the related external stimuli (w_i), as it is expressed in the following equation:

$$M_i = D_i + w_i \tag{1}$$

The ideal value for all the drives is 0. The external stimuli are the different objects that the player can find in the virtual world during the game. If the stimulus is present the value of w_i is 1, otherwise is 0.

According to (1), the intensity of a motivation is high due to two reasons: 1) the correspondent drive is high or 2) The correct stimulus is present. The dominant motivation is the one with the highest intensity.

This model can explain the fact that due to the availability of food in front of us, we sometimes eat although we are not hungry. We have also introduced activation levels (L_d) for motivations such that:

$$\begin{aligned} &\textit{if } D_i \le L_d \textit{ then } M_i = 0 \\ &\textit{if } D_i > L_d \textit{ then } (1) \text{ is applied} \end{aligned} \tag{2}$$

Therefore the possibility of no dominant motivation exists.

3.2 Wellbeing

As shown in (3), the agent's wellbeing is a function of the values of the drives (D_i) and some "personality" factors (α_i).

$$Wb = Wb_{ideal} - \sum_i \alpha_i \cdot D_i \tag{3}$$

Wb_{ideal} is the ideal value of the wellbeing of the agent, which is set to 100. The personality factors weight the importance of the values of the drives on the wellbeing of the agent. The value of the wellbeing and its variation (ΔWb) are calculated at each step. The variation of the wellbeing is calculated as the current value of the wellbeing minus the wellbeing value in the previous step.

3.3 Behaviour Selection

The action selection process consists in making decisions as to what behaviours to execute in order to satisfy internal goals and guarantee survival in a given environment and situation. For other authors (Avila and Cañamero, 2002), (Avila and Cañamero, 2004), (Cañamero, 1997) this implies that the agent can choose among some behaviors related to the dominant motivation. Therefore for each motivation there is a set of behaviours oriented to fulfill the motivational goal.

It is important to note that finally, the agent will learn that when the dominant motivation is Eat, it must select among the behaviours related to the object food, instead of those associated to water or medicine. The novelty of our approach is that these behaviours were not linked a priori with the correspondent motivations.

3.4 Happiness and Sadness

Considering the definitions of the emotions given in the introduction section:

$$\begin{aligned} &If\ \ \Delta Wb > L_h \Rightarrow Happiness \\ &If\ \ \Delta Wb < L_s \Rightarrow Sadness \end{aligned} \tag{4}$$

Where $L_h>0$ and $L_s<0$ are the minimum variations of the wellbeing of the agent that produce Happiness or Sadness respectively. Therefore these two emotions are used by the agent as the reward for the reinforcement learning algorithms.

In this architecture the agent learns, using different reinforcement learning algorithms, the best behaviour at each step using happiness/sadness as the positive/negative reward. 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 maximize Happiness.

4 Reinforcement Learning

Reinforcement learning (RL) is about learning from interaction how to behave in order to achieve a goal. The agent's objective is to maximize the amount of reward it receives over time (Sutton and Barto, 1998). Q-learning is a value learning version of RL that learns utility values (Q-values) of state and action pairs Q(s,a). It provides a simple way for agents to learn how to act optimally in controlled Markovian domains (Yang and Gu, 2004). The theory of Markov Decision Processes (MDP's), assumes that the agent's environment is stationary and as such contains no other adaptive agents (Littman, 1994). Therefore, while the agent is not interacting with the other agent, we will consider our virtual world as a MDP environment.

On the other hand, if the agent is interacting with other player, the rewards the agent receives depend not only on their own actions but also on the action of the other agent. Therefore, the individual Q-learning methods are unable to model the dynamics of simultaneous learners in the shared environment. Currently multiagent learning has focused on the theoretic framework of Stochastic Games (SGs) or Markov Games (MGs). SGs appear to be a natural and powerful extension of MDPs to multiagent domains (Yang and Gu, 2004).

Taking into account these considerations, in the proposed architecture the agent will use the standard Q-learning algorithm as the RL algorithm when the agent is not interacting with the other player. Obviously, in the case of "social" interaction, the agent must use a multiagent RL algorithm. The following subsections explain in more details these two scenarios.

In our system, the state of the agent is the aggregation of his inner state S_{inner} and the states S_{obj} related to each of the objects, including external agents, which can interact with him.

$$S = S_{inner} \times S_{obj_1} \times S_{obj_2} \cdots \tag{5}$$

For the RL algorithms the states related to the objects are considered as independent. This means that the state of the agent in relation with each object is $s \in S_{inner} \times S_{obj_i}$

4.1 Q-learning Algorithm

As mentioned previously, in MDP environments the agent will use the standard Q-Learning as a learning algorithm. As described in (Gadanho, 2002), through this algorithm the agent learns iteratively by trial and error the expected discounted cumulative reinforcement that it will receive after executing an action *a* in response to a world state *s*, the Q-values for each object is:

$$Q^{obj_i}(s,a) = (1-\alpha)\cdot Q^{obj_i}(s,a) + \alpha\cdot\left(r + \gamma \max_{a \in A_{obj_i}}\left(Q^{obj_i}(s',a)\right)\right) \tag{6}$$

where A_{obj_i} is the set of actions related to the object *i*, *s'* is the new state, *r* is the reinforcement; γ is the discount factor and α is the learning rate parameter.

The optimal policy, chooses the action that maximizes $Q^{obj_i}(s,a)$ this means

$$a^* = \arg\max_a Q^{obj_i}(s,a) \tag{7}$$

The proposed architecture differs from others in that we do not consider only the behaviours that help to satisfy the drive related with the dominant motivation but the agent must consider all the behaviours that can be performed at each step, depending on his states.

4.2 Multiagent reinforcement learning

In multiagent systems, other adapting agents make the environment no longer stationary so the Markov property is not applicable. In the learning framework of SGs, learning agents attempt to maximize their expected sum of discounted rewards. Unlike single-agent system, in multiagent systems the joint actions determine the next state and rewards of each agent. In (Littman, 1994) it is proposed a Minimax-

Q learning algorithm for zero-sum games in which the player always tries to maximize its expected value in the face of the worst-possible action choice of the opponent. The player's interests in the game are opposite. Later, Littman (Littman, 2001) proposed the Friend or Foe Q-learning algorithm, for the RL in general-sum SGs. The main idea is that each agent is identified in advance as being either "friend" or "foe". The Friend class consists of SGs in which the Q-values of the players define a game which has a coordination equilibrium. The Foe class is the one in which the Q-values define a game with an adversarial equilibrium. The Friend-Q updates similarly to regular Q-learning, and Foe-Q updates as does minimax-Q (Shoham et al, 2003).

All these algorithms extend the normal Q-function of state-action pairs $Q^{obj_i}(s,a)$ to a function of states and joint actions of all agents. Taking into account this fact and that each agent can select among n actions while they are interacting, the Q-values to be calculated are $Q^{obj_i}(s,a_1,a_2)$ where a_1 and a_2 belong to the set of n actions of each agent.

5 Fear

Fear is produced when the agent knows that something bad may happen. This means that the wellbeing of the agent might decrease. To cope with fear the action that produces the negative effect is going to be considered. We will distinguish between actions executed by the agent and exogenous actions carried out by other elements of the environment such as other agents.

5.1 To be afraid of executing risky actions

Q-learning algorithm evaluates every action carried out in a state, using the expected average value. However, since the system is non deterministic, the result of a certain action may have different values. The worst result experimented by the agent for each pair action-state is stored in a variable called $Q^{obj_i}_{worst}(s,a)$, which is updated after the execution of the action.

$$Q^{obj_i}_{worst}(s,a) = \min(Q^{obj_i}_{worst}(s,a),\ r+\gamma \max_{a\in A_{obj_i}} (Q^{obj_i}(s',a))) \tag{8}$$

where A_{obj_i} is the set of actions, *s'* is the new state, *r* is the reinforcement and γ is the discount factor.

The effect of being afraid can be considered by choosing the action that maximizes $Q^{obj_i}_{fear}$ instead of choosing the one that maximizes Q^{obj_i},

$$Q^{obj_i}_{fear}(s,a) = \beta Q^{obj_i}(s,a) + (1-\beta) Q^{obj_i}_{worst}(s,a) \tag{9}$$

Using this approach the expected result of each action is considered as well as the less favourable one. The parameter β, being $0\leq \beta \leq 1$, measures the daring degree of the agent, and its value will depend on the personality of the agent. If the agent is fearless, β will be near 1; while in a fearful agent, who tries to minimize the risk, β will be near 0. If $\beta=1$ the agent is using the optimal policy.

This means that the "fearful" policy chooses the action:

$$a^f = \arg\max_a Q^{obj_i}_{fear}(s,a) \tag{10}$$

For example, when an agent has to pass over a deep hole, he can choose between jumping over it and going around it. Jumping is easier, faster and usually safe, but very occasionally he can fail and die. On the other hand, if the agent goes around the hole he will take a lot of time and get tired but it is safer. Translating this example to our point of view, the Q-value related with jumping will be greater than the one related to going around. Using the standard Q-learning algorithm, the agent would always jump over the hole. Using the fearful policy, considering the worst thing that could happen to the agent jumping or going around, he would choose going around since it is safer than jumping.

5.2 To be afraid of malicious 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}.

Traditionally, Q-learning has been applied on Markov decision processes (MDP), which are discrete time systems. Some authors have extended the use of this algorithm to continuous time systems by considering them as semi Markov decision processes. In both cases it is commonly assumed that there are no exogenous events. In order to introduce the effects of exogenous events in continuous systems we consider the system as a discrete time system with constant period. In the limit, if the period is very small the system will tend to be a continuous time system. Moreover, we will also consider that the exogenous events can be associated to other agents or elements of the environment. These exogenous events are synchronized with the actions executed by the agent. Among these action we will include the action of "doing nothing". In this case the treatment for multiagent systems mentioned before will be applied.

The exogenous events executed by an external object or other agent can occur simultaneously to any of the actions of the agent. Therefore the negative effects of these exogenous events will be reflected in all the actions of the agent. In order to separate the effects of the actions of the agent and the effects of the exogenous events, we will focus on the study of the agent when he is "doing nothing". In that case, we suppose that all the changes suffered by the agent are a consequence of external elements.

It will be considered that a state is a "scary" state when:

$$Q_{worst}^{obj_i}(s, Nothing) < L_{fear} \tag{11}$$

being L_{fear} the minimum acceptable value of the worst result that can be expected by the agent when it is doing nothing. In this case the value of the fear drive D_{fear} will be incremented.

When

$$Q_{worst}^{obj_i}(s, Nothing) > L_{safe} \tag{12}$$

it is considered that the agent is in a "safe" state and the value of the fear drive D_{fear} will be decreased.

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.

6 Experimental Test Bed

The proposed architecture is intended to be used in a social personal robot developed by our lab and named "Maggie" (see Fig2) (Salichs et al, 2006). As a first stage of this project and due to the obvious physical difficulties of making experiments on a real robot and on a real environment, we decided to implement our architecture on virtual players, who "live" in a virtual world, a text-based multi user role game. This game gave us the possibly of creating different 2-D environments to play in, as well as a graphic interface.

Table 1 shows our agent's motivations, drives and external stimuli that the agent can find in the virtual world.

These drives have been selected tacking into account the role of the agent in the virtual world used to implement our architecture. Since our final goal is to construct an autonomous social robot, it must show social behaviours. Therefore, as it is shown, social motivations are included as robot's needs.

Table 1: Motivations, drives and related stimuli

Drive/Motivation	**External Stimuli**
Energy	Food
Thirst	Water
Health	Medicine
Sociability	Other player
Fear	

At each simulation step some of these drives, such as Energy, Thirst, Health and Sociability are incremented by a certain amount. The value of the drive Fear, as it was previously explained, increase or decrease depending on if the agent is in a "scary" state or not.

Following (3) the wellbeing of the agent is defined by:

$$Wb = Wb_{ideal} - (\alpha_1 D_{energy} + \alpha_2 D_{thirst} + \alpha_3 D_{health} + \alpha_4 D_{social} + \alpha_5 D_{fear}) \tag{13}$$

In our test bed the inner state is then:

$$S_{inner} = \{Hungry, Thirsty, Ill, Bored, Scary, OK\} \tag{14}$$

This internal state is obviously related with the dominant motivation. Therefore when the dominant motivation is for example "Eat" then the agent is "Hungry" and so on.

In relation with static objects the agent can be in the following states:

$$S_{obj} = Have\_it \times Near\_of \times Know\_where \tag{15}$$

where,

$$Have\_it = \{yes, no\} \tag{16}$$

$$Near\_of = \{yes, no\} \tag{17}$$

$$Know\_where = \{yes, no\} \tag{18}$$

In relation with other player:

$$S_{obj} = Near\_of \tag{19}$$

where,

$$Near\_of = \{yes, no\} \tag{20}$$

And the set of actions that can be executed in every state is the following:

$$A_{food} = \{Eat, Get, Go\_to, Explore\} \tag{21}$$

$$A_{water} = \{Drink, Get, Go\_to, Explore\} \tag{22}$$

$$A_{medicine} = \{Take, Get, Go\_to, Explore\} \quad (23)$$

$$A_{playmate} = \begin{cases} Explore \\ Steal\ food / water / medicine \\ Give\ food / water / medicine \\ Chat \end{cases} \quad (25)$$

Among the previously mentioned behaviours there are some of them that reduce or increase some drives, and therefore will produce a variation in the emotional state of the agent:

- Eat food: reduces to zero the Energy drive. (happiness when hungry)
- Drink water: reduces to zero the Thirst drive. (happiness when thirsty)
- Take medicine: reduces to zero the Health drive. (happiness when sick)
- Chat: reduces to zero the Social drive. (happiness when the social drive is high)
- To be taken something by other player: increases by a certain amount the Social drive. (sadness)
- To be given something from other player: reduces by a certain amount the Social drive. (happiness when the social drive is high)

Fig. 2. "Maggie" The Social Robot of the Robotic Lab.

The conducted experiments show the usefulness of the proposed architecture in facilitating the development of social autonomous agents able to learn from the experience the right behaviours to execute depending on the world state.

7 Conclusion and Future work

In this paper different reinforcement learning algorithms have been discussed and implemented for the behaviour-selection learning of non-interacting and social autonomous agents. These agents are controlled by an emotion-based architecture, which performs homeostatic control of the internal state of the agent through an embedded motivational model. This architecture has been designed for autonomous and social robots.

The agent is intrinsically motivated and his goal is his own wellbeing. 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.

In the future work, it is expected that the agent learns not only the right policy but also to identify its opponent. So far, the agent treats all its opponents as if they were all the same, and this is not true. In future scenarios, the agent will be able to behave different with the "good" opponent than with the one that tries to steal its objects every time that interacts with it.

Another emotion is going to be implemented: Anger. Anger will be produced when sadness arises due to the interaction with another 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

Avila-Garcia, O. and Cañamero, L. A Comparison of Behavior Selection Architectures Using Viability Indicators. In *Proc. InternationalWorkshop Biologically-Inspired Robotics: The Legacy of W. Grey Walter*(WGW'02). 2002.

Avila-Garcia, O. and Cañamero, L. Using Hormonal Feedback to Modulate Action Selection in a Competitive Scenario. *In Proc. 8th Intl.Conference on Simulation of Adaptive Behavior (SAB'04)*. 2004

Bellman , Kirstie L.. Emotions: Meaningful mappings between the individual and its world. In: *Emotions in Humans and Artifacts*. (Robert Trappl, Paolo Petta and Sabine Payr), pp 149-188. The MIT Press. Cambridge, Massachusetts. 2003

Berridge , Kent C. Motivation concepts in behavioural neuroscience. *Physiology & Behaviour* 81, 179-209,2004,.

Breazeal C. *Designing Sociable Robots*. The MIT Press. 2002

Cañamero, D. Modeling Motivations and Emotions as a Basis for Intelligent Behavior. *In W. Lewis Johnson, ed., Proceedings of the First International Symposium on Autonomous Agents (Agents'97)*, 148-155. New York, NY: The ACM Press. 1997.

Cañamero, D. Designing emotions for activity selection in autonomous agents. In: *Emotions in Humans and Artifacts.* (Robert Trappl, Paolo Petta and Sabine Payr), pp 115- 148. The MIT Press. Cambridge, Massachusetts. 2003

Damasio, Antonio. *Descartes' Error – Emotion, reason and human brain*. Picador, London. 1994

Fong. T., Nourbakhsh. I., Dautenhahn K. *A survey of socially interactive robots: Concepts, design, and applications.* Technical Report CMU-RI-TR-02-29. 2002

Frijda, N. and Swagerman, J. Can computers feel? Theory and design of an emotional model. *Cognition and Emotion.* 1 (3). pp 235-357. 1987

Fujita, Masahiro AIBO: Toward the Era of Digital Creatures. *The International Journal of Robotics Research*. Vol 20, Nº 10, pp 781-794. October 2001

Gadanho, Sandra Clara. Emotional and Cognitive Adaptation in Real Environments. *In: Symposium ACE'2002 of the 16th European Meeting on Cybernetics and Systems Research*, Vienna, Austria. 2002

Gadanho, Sandra Clara. Learning behavior-selection by emotions and cognition in a multi-goal robot task. *The Journal of Machine Learning Research.* Volume 4 Pages: 385 – 412. MIT Press Cambridge, MA, USA. 2003

Littman, M. L. Markov games as a framework for multiagent learning. *In Proceedings of the Eleventh International Conference on Machine Learning*, San Francisco, California, pp. 157--163. 1994

Littman, M. L. Friend-or-foe Q-learning in general-sum games. *In Proceedings of the Eighteenth International Conference on Machine Learning*, pages 322--328, Williams College, June 2001.

Lorentz K, Leyhausen P. *Motivation of human and animal behaviour; an ethological view*. New York: Van Nostrand-Reinhold; xix, 423 pp. 1973

Malfaz, M. and Salichs, M.A.. A new architecture for autonomous robots based on emotions. *Fifth IFAC Symposium on Intelligent Autonomous Vehicles*. Lisbon. Portugal. Jul, 2004.

Ortony, A., Clore, G. L., and Collins, A.. *The Cognitive Structure of Emotions.* Cambridge University Press. Cambridge, UK. 1988

Picard, Rosalind W. *Affective computing.* Ed. Ariel S.A. 1998

Picard, Rosalind W. What does it mean for a computer to have emotions?. In. *Emotions in Humans and Artifacts*. (Robert Trappl, Paolo Petta and Sabine Payr), pp 213-235. The MIT Press. Cambridge, Massachusetts. 2003

Rolls, Edmund, T. A Theory of emotion, its functions, and its adaptive value. In: *Emotions in Humans and Artifacts.* (Robert Trappl, Paolo Petta and Sabine Payr), pp 11-35. The MIT Press. Cambridge, Massachusetts. 2003

Salichs, M. et al. Maggie: A Robotic Platform for Human-Robot Social Interaction. In *2006 IEEE International Conferences on Cybernetics & Intelligent Systems (CIS) and Robotics, Automation & Mechatronics.* Bangkok, Thailand. 2006

Shibata T., Tashima T.,Arao M., Tanie K. Interpretation in Phisical Interaction between human and artificial emotional creature. *Proceedings of the 1999 IEEE. International Workshop on Robot and Human Interaction*. Pisa, Italy – September 1999

Shoham, Y., Powers, R. and Grenager, T. *Multi-agent reinforcement learning: a critical survey.* Technical report, Computer Science Department, Stanford University, Stanford. 2003.

Sloman, Aaron. How many separately evolved emotional beasties live within us. In: *Emotions in Humans and Artifacts*. (Robert Trappl, Paolo Petta and Sabine Payr), pp 35-115. The MIT Press. Cambridge, Massachusetts. 2003

Sutton, Richard S. and Barto, Andrew G. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, A Bradford Book.1998

Velásquez, J. When Robots Weep: Emotional Memories and Decision Making. *In: Proceedings of AAAI-98*. 1998

Yang E. and Gu D. *Multiagent Reinforcement Learning for Multi-Robot Systems: A Survey.* Technical Report CSM-404.University of Essex. (2004)