A biologically inspired architecture for an autonomous and social robot

María Malfaz, Álvaro Castro, Ramón Barber , Miguel A. Salichs

***Abstract*—Lately, lots of effort has been put into the construction of robots able to live among humans. This fact has favored the development of personal or social robots, which are expected to behave in a natural way. This implies that these robots could meet certain requirements, for example: to be able to decide their own actions (autonomy), to be able to make deliberative plans (reasoning), or to be able to have an emotional behavior in order to facilitate human-robot interaction.**

In this paper, the authors present a bio-inspired control architecture for an autonomous and social robot, which tries to accomplish some of these features. In order to develop this new architecture, authors have used as a base a prior hybrid control architecture (AD) that is also biologically inspired. Nevertheless, in the later, the task to be accomplished at each moment is determined by a fix sequence processed by the Main Sequencer. Therefore, the Main Sequencer of the architecture coordinates the previously programmed sequence of skills that must be executed. In the new architecture, the Main Sequencer is substituted by a decision making system based on drives, motivations, emotions, and self-learning, which decides the proper action at every moment according to robot's state. Consequently, the robot improves its autonomy since the added decision making system will determine the goal and consequently the skills to be executed. A basic version of this new architecture has been implemented on a real robotic platform. Some experiments are shown at the end of the paper.

***Index Terms*—Cognitive robotics, control architectures, autonomy, decision making systems, motivations, emotions.**

I. Introduction

In the Nineties, the term "cognitive robotics" was first introduced by Ray Reiter and his colleagues, who have a research group on this topic at the University of Toronto. According to them, cognitive robotics is concerned with endowing robotic or software agents with higher level cognitive functions that involve reasoning about goals, perception, actions, mental states of other agents, collaborative task execution, etc.

Since the Seventies, robotics has evolved trying to provide useful services to humans. Today, robots which carry out dangerous [1], assistance [2], or transportation tasks [3], among others, are a reality. Traditionally, robotic research has been centred on control architectures, planning, navigation, etc. Nevertheless, during the last few years, the interest in robots which are integrated in our everyday environment, personal robots, has increased [4]. Human-robot interaction is one of the main characteristics of these robots.

The authors are with the RoboticsLab at the Carlos III University of Madrid. 28911, Leganés, Madrid, Spain e-mail: mmalfaz@ing.uc3m.es; acgonzal@ing.uc3m.es; rbarber@ing.uc3m.es; salichs@ing.uc3m.es

Some researchers in cognitive robotics have begun to use their old architectures as a base for their cognitive robotics programs. The idea is to extend these architectures in order to implement some of the high level cognitive functions. The first control architecture developed by the authors, a hybrid architecture named AD (Automatic/Deliberative), took the theories of the modern psychology expressed by Shiffrin and Schneider [5], [6] as a base. According to these authors, two mechanisms of processing information are established: automatic processes and controlled ones. Therefore, we can differentiate between two levels of activity in human beings: automatic and deliberative.

Moreover, at the beginning of the Sixties, the artificial intelligence precursor Herbert Simon was convinced that including emotions in the cognitive model to approximate the human mind was necessary [7]. Later, near the mid Nineties, Antonio Damasio published *Descartes's Error* [8]. His studies proved that damage to the brain's emotional system caused the patient to make poor judgments despite intact logical reasoning skills. As a consequence, the positive role of human emotions in cognition started to gain prominence among a group of researchers from the scientific community. Later, other studies showed that emotions have influence on many cognitive mechanisms, such as memory, attention, perception, and reasoning [9], [10], [11], [12]. Besides, emotions play a very important role in survival, social interaction and learning of new behaviors [13], [14], [15].

Therefore, in recent years, the role of emotional mechanisms in natural and artificial cognitive architectures, in particular in cognitive robotics, has been considered. According to [16], in relation to the main question: do robots need emotions? many researchers have answered positively, mainly considering the two aspects of emotion: the external (social) one and the internal (individual) one. It seems to be obvious that in human-robot social interaction, expression of emotions helps to make interaction more natural [17]. On the other hand, the internal aspects of emotion, i.e., its role in the behavioral organization of an individual cognitive agent, are essential for the autonomy issue, and this is the main concern of many researchers.

The concept of autonomy has been treated by several authors such as Arkin [18], Gadanho [10], Bellman [14], or Cañamero[15]. In general, they state that an autonomous agent must be self-sustained, which implies a decision making system. Moreover, it must have some goals and motivations and these must be oriented to maintain its internal equilibrium (homeostasis).

In fact, in recent years, several authors have argued that a truly biologically inspired and truly cognitive robotics would need to take into account homeostatic/emotional dynamics, i.e., the interplay between constitutive and interactive aspects of autonomy; for example, the need to keep essential system-internal variables within certain viability ranges [19].

The work presented in this paper tries to consider all the previous requisites in order to design a new biologically inspired architecture for an autonomous and social robot. **Although the current context of this robot is a laboratory, this proposed architecture will be implemented on social robots living with human beings and sharing common spaces with complex configurations. In these situations, autonomy and friendly human-robot interaction are essential. Therefore, as previously stated, in order to implement those features in our robot, this new bio-inspired architecture is required.** In order to fulfill this goal, we started our approach using the previous hybrid control architecture developed by the authors, the AD architecture [20]. This architecture has been modified, replacing the Main Sequencer that manages the global behavior of the robot with a decision making system based on drives, motivations, emotions, and self-learning [21]. Following a homeostatic approach, the goal of the robot can be to satisfy its needs maintaining its necessities within an acceptable range. The learning process is made using a well-known reinforcement learning algorithm, Q-learning [22]. By using this algorithm, the robot learns the value of every state-action pair through its interaction with the environment. This means, it learns the value that every action has in every possible state. The highest value indicates that the correspondent action is the best one to be selected in that state. At the beginning of the learning process these values, called the q-values, can all be set to zero, or some of them can be fixed to another value. In the first case, this implies that the robot will learn from scratch, and in the second, that the robot has some kind of previous information about the behavior selection. These initial values will be updated during the learning process.

Before implementing this system on a real robot, as a previous step, this work was successfully implemented on virtual agents [23], [24], [25]. In those works, how the agent learned the right action to execute in every situation presented in the environment by using reinforcement learning algorithms with no previous knowledge is shown.

In this paper, we present the first results of our work by showing the implementation on a real robotic platform of a simplified version of this new architecture, with neither emotions nor self-learning. The robot is working with a decision making system based on drives and motivations, which sends the information about what to do in every moment to the AD architecture.

The paper is organized as follows. The next section, section II, introduces some basic concepts needed for the decision making system. Next, in section III, the question *why do robots need emotions?* is answered from a functional point of view. In section IV, a brief review of biologically inspired control architectures, some of them based on traditional mechanisms of processing information, and others based on emotions and motivations, is given. Section V shows the approach proposed in this paper: the AD architecture, modified by adding a decision making system. In this section, both the architecture and the system, are described. Next, section VI shows how a basic version of this new architecture is being implemented on a real robotic platform. Moreover, this platform is also briefly described. Section VII presents some results of the experiments carried out. And finally, the conclusions and future works are summarized in section VIII.

mds
February 25, 2010

II. Homeostasis, drives, and motivations

In this section we give a brief review of the basic concepts related to the decision making system. As stated previously, we will follow an homeostatic approach when designing this system and terms such as drives and motivations must be introduced.

Homeostasis was discovered by Claude Bernard in the middle XIX century when he observed that the body variations had as an objective to give the stability back to the body. In other words, we could say that homeostasis means maintaining a stable internal state, see [26]. According to the homeostatic approach, the human behavior is oriented to the maintenance of the internal equilibrium.

One of the oldest theories about drives was proposed by Hull in [27]. Hull suggested that privation induces an aversion state in the organism, which is termed drive. According to his theory, the drives increase the general excitation level of an animal and they are considered as properties of deficit states which motivate behavior.

The word motivation derives from the Latin word *motus* and indicates the dynamic root of the behavior, that means those internal, more than external, factors that urge to action [28]. In other words, the motivational state is a tendency to correct the error (drive) through the execution of behaviors.

Many drive theories of motivation between 1930 and 1970 posited that drive reduction is the chief mechanism of reward. If motivation is due to drive, then, the reduction of deficit signals should satisfy this drive and essentially could be the goal of the entire motivation [26].

Hull [27] also proposed the idea that motivation is determined by two factors. The first factor is the drive. The second one is the incentive, that is, the presence of an external stimulus that predicts the future reduction of the need. For example, the presence of food constitutes an incentive for a hungry animal.

In our work, the robot has certain needs (drives), that need to be satisfied, and motivations. Following the homeostatic approach, the proposed decision making system will be oriented to maintain those needs within an acceptable range. These needs will not be just limited to physical ones (as it is stated in the classical point of view of the homeostasis), but psychological and social necessities too. In the next section, the reasons why emotions should be included in this decision making system will be analyzed.

III. Why do robots need emotions?

Several authors have expounded their reasons to include emotions in robots besides their importance in the human-robot interaction. Moreover, others have studied their generation as well as the optimal number that should be implemented on virtual agents or real robots. In this section, a brief review of these ideas is given.

According to Arkin, motivations/emotions provide two potential crucial roles for robotics: survival and interaction [13]. Cañamero 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, similarly to Arkin, that it could be useful to exploit this role of emotions to design mechanisms for an autonomous agent [15].

On the other hand, Ortony explains that robots need emotions for the same reason as humans do: one of the fundamental functions of emotions is that they are a requisite for establishing long-term memories. The second function is that emotions provide opportunities for learning, from simple forms of reinforcement learning to conscious and complex planning [29].

In the same line, Bellman [14], Fellows [30], and Kelley [31] state that, since emotions allow animals with emotions to survive better than others that lack emotions, robots should be provided with features related to emotions in a functional way.

Different models of emotion systems have been proposed to be implemented in artificial agents. One of the main differences among them is the mechanism used to generate emotions.

Currently, most experts agree that emotions are produced by an appraisal of the situation of the agent in its relation with the world. Therefore, different emotions are associated to different situations. Many researchers think that the relation between situations and emotions is mediated by a set of intermediate variables. These variables act as dimensions of an affective space and each emotion is associated to a different zone of the affective space [17] [32]. Lazarus [33], on the other hand, considers emotions as discrete categories. In the discrete emotional approach, dimensions of emotional intensity are still employed, but these are applied within each emotional category.

In our work, we follow a discrete emotional approach and we consider that the relation between situations and emotions is specific for each emotion. Therefore, each emotion requires a particular study to establish this relationship.

According to Spinola and Queiroz [34], another important issue related to the implementation of artificial emotions in robots is: How many and which emotions must be selected? In this work many different approaches are described, from authors that defended the idea of implementing a varying number of “basic” or “primary” emotions, from 4 to 22, [35] [36] [37], to others that decided to implement just one or two emotions [38], [39]. Finally, one very different point of view is presented by Cañamero in [40]: “Do not put more emotion in your system than what is required by the complexity of the system-environment interaction”.

Following this last point of view, currently, our research focuses on three emotions: happiness, sadness, and fear. Until now, the implementation of emotions has been done on virtual agents which live in a simple environment. In a near future, our intention is to define new emotions when the complete control architecture is implemented on a real robot. Therefore, as the functionality of our robot and its environment become more complex, it will have to cope with new situations and maybe new emotions, or a redefinition of the existing ones, will be needed.

IV. Related work

A. Classical control architectures

Classical architectures are mainly focused in navigation tasks. Early robots, in the Sixties and Seventies, used planning-based architectures [41][42]. Any movement of these robots had to be planned in advance. Planners needed models to predict the results of each action. The main goal of these robots was motion, and therefore, the models were maps of the environment and the planners were motion planners.

Poor results of planning-based architectures obliged to search for other alternatives. In the mid Eighties, reactive architectures began to be developed [43], [44], [45]. In reactive architectures, the use of planners and models was minimized. In fact, there were neither maps nor planners in most robots with reactive architectures. Decisions were based on real time information from sensors, making the creation of maps with that information unnecessary. At this time, this approach produced very good results in comparison with planning based architectures. For instance, robots were able to move quite fast in dynamic environments, avoiding obstacles.

Reactive architectures meant a significant advance in the development of robots, although not everything was positive in reactive architectures. They also have some drawbacks. Behaviors of robots with reactive architectures usually do not include the achievement of an explicit goal as in the planning-based architectures.

In sum, both approaches, reactive and planning-based, offer some advantages, but they also show some drawbacks. Trying to get the best of both, in the mid Nineties hybrid architectures began to appear. These architectures usually adopt a reactive approach at the low level (the modules closer to sensors and actuators), and a planning-based approach at high level. That means that motion control loops are close to low level producing different behaviors, and at the same time it is possible to reach planned decisions based on models. Reactive modules make short term decisions in local areas (e.g. immediate movements in the area close to the robot) and planning modules make mid and long term decisions at global areas (e.g. future movements to distant areas).

Among hybrid architectures, the next ones can be highlighted. Firby establishes three levels: a planner which makes plans according to the goal to be reached, a controller which interacts with the environment, and an executor (RAP) that links the planner and the controller, giving the detailed information which the controller requests from the planner’s information [46] [47]. Bonasso, in the 3T architecture, considers a layer consisting of reactive skills, a sequencer that enables or disables the skills, and a deliberative planner capable of guiding the robot to the target goal [48] [49]. Gat, in the

ATLANTIS architecture, distinguishes among a controller of reactive primitive activities, a sequencer that manages those primitive activities according to the deliberative computations, and a deliberator which is in charge of the planning [50] [51] [52]. Lastly, Chatila considers a functional level which includes perceptive and motor capacities, an execution level, without reaction capacity, which controls them, and a decisional level in which the planning and supervision are included [53].

The three layer architectures mentioned above have the sequencing layer between the deliberative and reactive ones. This fact leads to a rigidity in the planning-sequencing-acting paradigm.

On the other hand, our AD (Automatic/Deliberative) architecture [54] was designed trying to avoid rigidity in the mentioned planning-sequencing-acting paradigm. It is composed by only two levels: one for deliberative activities and a second one for automatic activities. The sequencing processes are distributed between the Deliberative and Automatic levels, providing more flexibility to the hybrid architecture.

B. Control architectures based on motivations and emotions

More recently, as previously stated, some authors have implemented cognitive-related concepts in their control architectures, such as motivations, emotions, learning, etc. In this section, we present a review of the works that have inspired our research.

The work developed by Lola Cañamero is one of the first researches done in this area [55], [56], [15]. The original idea was that the behaviors of an autonomous agent are directed by motivational states and its basic emotions. As it is said before, motivations can be viewed as homeostatic processes that maintain inner variables controlled within a certain range. A detector of errors generates an error signal, the drive, when the value of this variable is not equal to its ideal value. Each motivation is modelled as a function of its related drive and an external or incentive stimulus. The motivation with the highest value becomes active and it will organize the behavior of the agent in order to satisfy the drive. Emotions in this approach influence the decision making process in two ways. First, they can modify the intensity of the current motivation and, as a consequence, the intensity of the related behavior. In fact, in extreme cases, they can avoid the execution of the behavior. Second, they can modify the reading of the sensors that monitors the variables affected by emotions. Therefore, they can alter the perception of the state of the body. The implemented emotions work as monitoring mechanisms to cope with important situations related to survival.

Another interesting approach is the one presented by Gadanho [10], [57]. In this work, the research is focused on how artificial emotions can improve the behavior of an autonomous robot. In her approach, the robot adapts to its environment using an adaptive controller adjusted by using reinforcement learning. Emotions are used to influence perception, as Cañamero does, and to provide a reinforcement function. This is because, according to the authors, it is frequently assumed that the human decision making process consists on maximizing the positive emotions and minimizing the negative ones. In later works, [58], the emotional system was substituted by a goal system. This system is based on a set of homeostatic variables which must be maintained within a certain range. The goals are explicitly associated to the homeostatic variables.

Another approach was developed by Velásquez [59], [60], who proposed an architecture called Cathexis. This architecture was developed for autonomous agents and contains an emotion generation model. Moreover, it also has simple models for other motivations and a decision making algorithm. Later, this architecture was completed by a drive system in order to develop a decision making model based on emotions. In this model, the emotional system is the main motivation of the agent. The drive system even exploits its influence in order to select specific behaviors. For example, the Hunger drive and the Distress caused by it motivate the agent to obtain food. In this model, the behaviors compete among each other to take the control. Therefore, only one behavior is active at a time.

Currently, this work has been continued by Cynthia Breazeal, whose main research interest is the study of human-robot interaction. The developed robots, Kismet and Leonardo, have a cognitive and an emotional system. The cognitive system is formed by the perception, the attention, the drives, and the behavior systems. The behaviors are selected based on the values of the drives and the external stimuli. These behaviors are also related to every drive and they compete to determine which need must be satisfied. The role of the emotional system is to influence the cognitive system to promote appropriate and flexible decision making, and to communicate the robot internal states, see [61], [62], and [63].

Nowadays, many efforts have also been put on autonomous agents for characters in computer games. Sevin, in [64], developed a motivational model of action selection for virtual humans. The model chooses the appropriate behavior according to the motivations and the environmental information. In other works, there are actions associated to motivations. Therefore, the actions related to the highest motivation become active. Most of the time, the action receiving activity from the highest internal variable is the most active one, and then it is chosen by the action selection mechanism.

Until this point, most of the presented works use a motivational system in order to select the behaviors and, in some of them, emotions are only used to influence this decision making in one way or another. Nevertheless, there are other approaches that use emotions as the central aspect of the decision making system. This is the case, for example, of the work presented by Hirth et al, [65]. They propose an emotion-based control architecture which consists of three main parts: behavior, emotion, and cognition. All possible movements of the robot, from simple reflexes up to high level motor skills, are located in the behavior group. These behaviors are activated in different ways, e.g., directly depending on sensor data, depending on the emotional state, or deliberately by the cognition part. In this architecture, the high level behaviors are mostly activated by the emotions and specially by the cognitive part, whereas low level behaviors are activated directly by the sensor input.

In [32], Hollinger et al present another robot using emotion-based decision mechanisms. These mechanisms are based on the Mehrabian PAD (pleasure, arousal, and dominance) scale that determines emotions using an affective space. The robot state is translated into a particular set of sound and movement responses. In this approach, the emotion state of the robot varies according to its interaction with people. In fact, this gets modified when the robot sees different color shirts.

Finally, another approach is the one presented by Lisseti and Marpaung in [66], where the behavior of the robot is selected according to its current emotional state. They generate this emotional state based on the data received from the input sensors of the robot. In fact, each emotion is related to certain external events, e.g., the parameter of the Sad emotion is increased if the door is closed or the robot does not recognize someone. Once the emotional state is determined, the robot will execute the proper action tendency, i.e., the robot identifies the most appropriate (or a set of) actions to be taken from that emotional state.

The presented work has been inspired mainly by Cañamero's, Gadanho's, and Velásquez's works. As will be shown in the next section, we use homeostatic drives that are related to motivations, as those authors do. In our approach, the motivations, and not the behaviors (as referred to in Velásquez's and Breazeal's approaches) compete among each other following the point of view of Cañamero. Nevertheless, in her approach, the winner motivation has a related behavior that satisfies the associated need, as Sevin also proposes.

In fact, one of the main differences of our work with other authors's approaches is that the behaviors are not necessarily previously linked with a need or an emotion. This means that there are no motivational or emotional behaviors. The agent/robot can learn, using a reinforcement learning algorithm, which behavior to select in order to satisfy the most urgent drive. In Cañamero's and Sevin's works, it is assumed that there is only one behavior able to satisfy one need. This fact can be seen as a disadvantage, since it limits the flexibility of the decision making system. It could happen, as in our approach, that several behaviors satisfy the same need. This point of view seems to be more bio-inspired since, in nature, in order to satisfy for example, hunger, we can eat something but also, drinking some water can reduce this need.

The second difference is that in our approach, the way each emotion is defined in the architecture is different. This means that emotions are not defined as a whole as most authors do. As can be observed, there are two points of view in relation to the role of emotions in the decision making process. Cañamero, Gadanho, Velásquez, and Breazeal used emotions to influence the decision making process, not for selecting the behavior directly according to them. On the contrary, others, such as Hirth et al, Hollinger et al, and Lisseti and Marpaung consider emotions as the central aspect of their decision making system so, in some cases, the behavior is selected according to the current emotional state. In our approach, we do not limit the role of emotions to one of them, but we exploit both points of view. On one hand, some emotions are used as the reinforcement function in the learning process, as Gadanho also proposed, not determining directly the action selection. On the other hand, other emotions are defined as motivations so, the behaviors will be completely oriented to cope with the situation that generated those emotion.

V. Our approach: AD architecture with a biologically inspired decision making system

A. AD architecture

As stated in section I, the previous control architecture developed by the authors is the AD architecture. This biologically inspired architecture is based on the ideas of the modern psychology expressed by Shiffrin and Schneider [5], [6], so it considers two levels, the automatic and the deliberative levels, as shown in Figure 1.

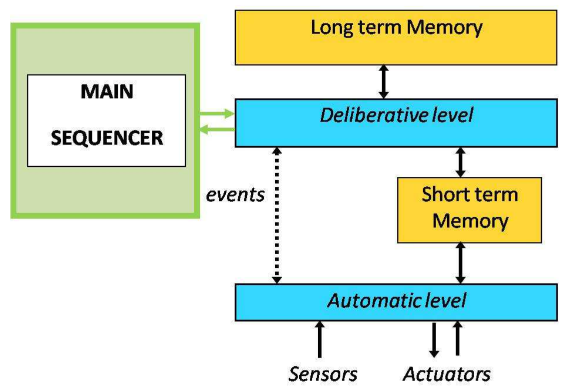

Fig. 1. AD architecture levels

In AD architectures [54], both levels are formed by skills, which endow the robot with different sensory and motor capacities, and process information. Skills can be coordinated by sequencers and the Main Sequencer manages the deliberative skills according to a predefined sequence. This sequencer will be explained later in more detail.

1) Deliberative level: In the natural world, humans deliberative activities are characterized by the fact that these are carried out in a conscious form. Moreover, temporal dimension is an important property: deliberative processes require a large quantity of time to be dedicated to the analysis. These activities are carried out sequentially, that is, one after another, and it is not possible to carry out more than one deliberative activity at a time.

In our AD architecture implementation, deliberative skills are based on these activities and the authors consider that only one deliberative skill can be activated at once.

2) Automatic level: Living beings' automatic activities are characterized by the fact that their actions and perceptions are carried out without the necessity of having consciousness of the processes responsible for controlling those activities. Examples of this would be the heart beat, the hand movement when writing, or that of legs when walking. An automatic activity can be carried out in parallel with other automatic activities and with a deliberative activity. For example, a person can be driving a vehicle and maintaining a conversation simultaneously. The level of complexity of automatic activities may be very variable and goes from the "simplicity" of moving a finger to the complexity of playing a sonata previously memorized on the piano.

In the AD implementation, the automatic level [67] is mainly formed by skills which are related with sensors an actuators. Automatic skills can be performed in a parallel way and they can be merged in order to achieve more complex skills.

3) AD Memories: One of the main characteristics of human beings is their ability to acquire and store information from the world and from their own experiences. Memory can be defined as the capacity to recall past experience or information in the present [**?**].

Based on the memory model proposed by Atkinson and Shiffrin [68], the AD architecture considers two different memories: the Short-Term Memory and the Long-Term Memory, see Figure 1. In our architecture, Short-Term Memory is defined as a temporary memory. This memory is regarded as a working memory where temporal information is shared among processes and skills. On the other hand, Long-Term Memory is a permanent repository of durable knowledge. This knowledge can come from learning, from processing the information stored in Short-Term Memory, or it can be given *a priori*. In AD architecture this memory refers to a permanent memory where stable information is available only for deliberative skills.

4) The Main Sequencer: The Main Sequencer, as it is shown in Figure 1, is the element in charge of coordinating deliberative skills in order for a robot to fulfil a task. The Main Sequencer performs a sequence of skills that must be carried out by the robot. This sequence is a fixed script where all possible situations that the robot can face are considered. This means that this script has been programmed in advance and it is exclusive for certain objectives.

A relevant feature of this architecture is, as already stated, that all possible options must be considered in the sequence a priory. Depending on the definition of autonomy, this can be considered a negative factor since, in bio-inspired systems, the fact that it is the proper agent/robot who must decide its own objectives it is assumed. Therefore, since this is our objective, the Main Sequencer has been replaced with a decision making system based on drives, motivations, emotions, and self-learning. This system is described in the next section.

B. Adding the biologically inspired decision making system

As shown in Figure 2, the decision making system has a bidirectional communication with the AD architecture. On one side, the decision making system will select the behavior the robot must execute according to its state. This behavior will be taken by the AD architecture activating the corresponding skill/s (deliberative or automatic one). On the other side, the decision making system needs information in order to update the internal and external state of the robot.

The general idea of the proposed decision making system is shown in Figure 3. As explained in section II, the term homeostasis means maintaining a stable internal state [26]. This internal state can be configured by several variables, which must be at an ideal level. When the value of these

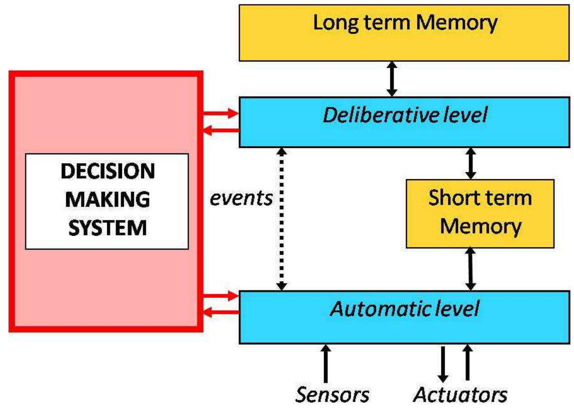

Fig. 2. AD architecture with the decision making system

variables differs from the ideal one, an error signal occurs: the drive [55].

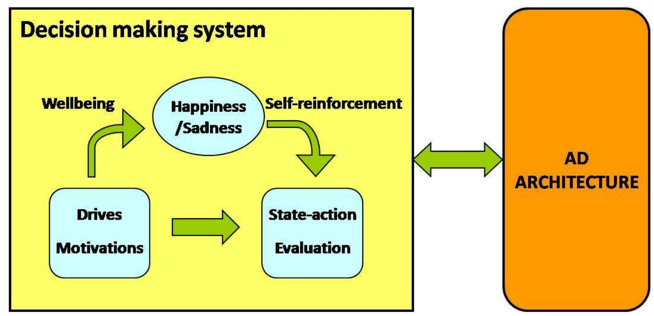

Fig. 3. The decision making system

In our approach, the autonomous robot has certain needs (drives) and motivations, and following the ideas of Hull [27] and Balkenius [69] [70], the intensities of the motivations of the robot are modeled as a function of its drives and some external stimuli. For this purpose we used Lorentz's hydraulic model of motivation as an inspiration [71]. In Lorenz's model, the internal drive strength interacts with the external stimulus strength. If the drive is low, then a strong stimulus is needed to trigger a motivated behavior. If the drive is high, then a mild stimulus is sufficient [26]. The general idea is that we are motivated to eat when we are hungry and also when we have food in front of us, although we do not really need it. Therefore, the intensities of the motivations are calculated as shown in (1)

$$\begin{array}{l}\text{If } D_i < L_d \text{ then } M_i = 0 \\ \text{If } D_i \geq L_d \text{ then } M_i = D_i + w_i\end{array} \quad (1)$$

where M_i are the motivations, D_i are the related drives, w_i are the related external stimuli, and L_d is called the activation level.

According to Balkenius [69] [70], all excited motivational states can not be allowed to direct the robot at once since this would generate incoherent behaviors. In his opinion, this problem cannot be handled solely by behavioral competition but must be resolved at an earlier stage of processing. The solution proposed is a motivational competition, as Cañamero also proposed in [55]. Therefore, in our approach, once the intensity of each motivation is calculated, they compete among themselves for being the dominant one, and this one determines the inner state of the robot. It could happen that if none of the drives is greater than the activation level L_d then, there is no dominant motivation.

As stated in previous sections, in this decision making system, there are no motivational behaviors. This means that the robot does not necessary know in advance which behaviors to select in order to satisfy the drive related to the dominant motivation. There is a repertory of behaviors and they can be executed depending on the relation of the robot with its environment, i.e. the external state. For example, the robot will be able to interact with people as long as it is accompanied by someone.

The objective of this decision making system is having the robot learn how to behave in order to maintain its needs within an acceptable range. For this purpose, it uses the Q-learning algorithm to learn from its bad and good experiences. As previously stated, the autonomous robot can learn, from scratch or using some a priori information about some q-values of the state-action pairs, the proper behavior to select in every state through its interaction with the environment.

Besides, as shown in Figure 3, happiness and sadness are used in the learning process as the reinforcement function and they are related to the wellbeing of the robot. Next, we justify this decision but first, let us introduce this concept: the wellbeing of the robot is defined as a function of its drives and it measures the degree of satisfaction of its internal needs.

$$Wb = Wb_{ideal} - \sum_i \alpha_i \cdot D_i, \tag{2}$$

where α_i is the set of the personality factors that weight the importance of each drive on the wellbeing of the robot and Wb_{ideal} is the ideal value of the wellbeing of the robot. As observed, as the values of the needs of the robot increase, its wellbeing decreases.

In order to define happiness and sadness, we took the definition of emotion given by Ortony [36] into account. In his opinion, emotions occur due to an appraised reaction (positive or negative) to events. According to this point of view, in [72], Ortony proposes that happiness occurs because something good happens to the agent. On the contrary, sadness appears when something bad happens. In our system, this can be translated into the fact that happiness and sadness are related to the positive and negative variations of the wellbeing of the robot.

On the other hand, the role of happiness and sadness as the reinforcement function was inspired by Gadanho's works, as shown in section IV-B, but also by Rolls [73]. He proposes that emotions are states elicited by reinforcements (rewards or punishments), so our actions are oriented to obtaining rewards and avoiding punishments. Following this point of view, in this proposed decision making system, happiness and sadness are used as the positive and negative reinforcement functions during the learning process, respectively. Moreover, this approach seems consistent with the drive reduction theory introduced in section II where, according to this theory, the drive reduction is the chief mechanism of reward.

In summary, the decision making process is cyclic and it can be described in the following points:

1. Updating the drives and motivation intensities.
2. Motivation competition and selection of the inner state.
3. Determining the external state.
4. Updating the wellbeing function.
5. Generating the reinforcement function (happiness/sadness).
6. State-action evaluation (reinforcement learning).
7. Behavior selection.

As said at the beginning of this paper, this decision making system has been successfully implemented on autonomous virtual agents [23], [24], [25]. These agents live in a virtual environment created using a text-based game available online and called CoffeMud [74]. Next, in this section, we give a brief review of this implementation and the results obtained.

The environment where the agent has to live is a simple rooms-corridor stage. In these rooms, it can find several objects, such as food, water, etc., which are needed in order to satisfy the drives of the agent. Moreover, the agent has a limited set of actions related to every object, for example, "to eat food", or "to take water". In this implementation, all the initial q-values are set to zero, therefore, the agent does not have any previous information about the behavior selection. It is important to note again, that the actions are not related to motivations. This means that the agent does not know in advance that, for example, it must eat in order to satisfy its hunger. The drives and motivations implemented are: Hunger, Thirst, Weakness, Loneliness, and Fear. Hunger, Thirst, and Weakness, are related to the consumption of food, water, and medicine respectively. Loneliness is related to social interaction and in order to satisfy it, the agent must interact with other agents that will be sharing the same environment. Moreover, those agents (opponents) are able to behave badly or kindly with our agent.

In relation to the emotions, happiness and sadness are the reinforcement function, as previously explained, and the emotion fear, based on some theories that state that emotions can motivate behaviors [12], [17], [75], is defined as a motivation. Therefore, according to our decision making process, fear could be the dominant motivation and, in that case, the agent would be "scared". When this happens, the agent must learn the right action to execute in order to cope with the situation that caused this inner state.

The results obtained showed that the agent, using this decision making system, is able to learn the right sequence of actions in order to satisfy its needs by maximizing its wellbeing. This means, for example, that in the case the agent is hungry, it learns that it must go where the food is, then take it, and finally eat it. These results can be seen as obvious, but it was the right selection of the reinforcement function which allows the agent to learn properly, without any previous knowledge about which action to select at every moment. Another important result is the one obtained with the emotion fear. How the agent is able to generate a "run-away" behavior that was not previously programmed is shown

Fig. 4. Our social robot Maggie interacting with children

in [25]. Moreover, the agent is able to identify the situation that scared it. This fact is quite important since most authors have an emotional releaser, as for example, to be in presence of an enemy, but in our case the agent, after several trials, learns to identify that dangerous situation.

VI. Implementation on a social robot

In this section, the developed system is presented. First, the robotic platform is briefly introduced. Then, the general elements in the architecture are presented. Later on, the decision making module is shown, explaining how it interacts with the architecture. As already stated, the decision making system implemented on the robot is a basic version, and currently it is being improved and extended. In this first approach, neither emotions nor learning have been implemented on the robot.

A. Framework

The presented work has been implemented on the research robotic platform named Maggie [76]. Maggie is a social and personal robot intended for performing research on human-robot interaction and improving robots autonomy (Figure 4). It was conceived for personal assistance, for entertainment, to help handicapped people, to keep people accompanied, etc. Its external friendly look facilitates its social robot task. Both software and hardware have been developed by the Robotics Lab research group from the University Carlos III of Madrid.

In relation to its hardware, Maggie is a computer-controlled system with a wheel base which allows the robot to move through the environment. Its arms, neck, and eyelids movements show signs of life. The vision system uses a camera in the head and, thanks to it, Maggie can recognize people and play several games. Laser telemeter and ultrasound sensors are used by the navigation system to avoid collisions. By means of an infrared emitter/receiver, Maggie also operates different home appliances such as televisions. Touch sensors on the surface of the body and a touch screen situated in the breast are used for a direct interaction with people. Inside the head, an RFID antenna is placed for identifying objects. In order to provide verbal interaction, our robot is equipped with a text-to-speech module and an automatic speech recognition system.

The required energy for all devices is received from two batteries which provide a power supply of 25 V. During its working life, the robot needs at least 20 V. The purpose is to achieve a robot working continuously in a never-ending working life. This means that the battery should always be over this threshold.

B. AD architecture

Considering the ideas previously stated, the software is based on the two levels of the Automatic-Deliberative architecture [20], [54], previously described in section V. The automatic level is linked to modules that communicate with hardware, sensors, and motors. At the deliberative level, reasoning processes are placed. As shown in Figure 2, the communication between both levels is bidirectional and it is carried out by the Short-Term Memory and events [77].

Events are the mechanisms used by the architecture for working in a cooperative way. An event is an asynchronous signal for coordinating processes by being emitted and captured. The design is accomplished by the implementation of the publisher/subscriber design pattern so that an element that generates events does not know whether these events are received and processed by others or not.

The Short-Term Memory is a memory area which can be accessed by different processes, where the most important data is stored. Different data types can be distributed and are available to all elements of the AD architecture. The current and the previous value, as well as the date of the data capture, are stored. Therefore, when writing new data, the previous data is not eliminated, it is stored as a previous version. The Short-Term Memory allows to register and to eliminate data structures, reading and writing particular data, and several skills can share the same data. It is based on the blackboard pattern.

On the other hand, the Long-Term memory has been implemented as a data base and files which contain information such as data about the world, the skills, and grammars for the automatic speech recognition module.

As already stated, the essential component in the AD architecture is the skill [77] and it is located in both levels. In terms of software engineering, a skill is a class that hide data and processes that describes the global behavior of a robot task or action. The core of a skill is the control loop which could be running (skill is activated) or not (skill is blocked).

Skills can be activated by other skills, by a sequencer, or by the decision making system. They can give data or events back to the activating element or other skills interested in them. Skills are characterized by:

- They have three states: ready (just instantiated), activated (running the control loop), and locked (not running the control loop).

- Three working modes: continuous, periodic, and by events.
- Each skill is a process. Communication among processes is achieved by Short-Term Memory and events.
- A skill represents one or more tasks or a combination of several skills.
- Each skill has to be subscribed at least to an event and it has to define its behavior when the event arises.

The AD architecture allows the generation of complex skills from atomic skills (indivisible skills). Moreover, a skill can be used by different complex skills, and this allows the definition of a flexible architecture.

C. The decision making system

The decision making system proposed in preceding sections is intended for achieving a full autonomous robot. Therefore, the decision making module is the one in charge of selecting the most appropriated skill at each moment for maximizing the robot wellbeing. Choosing the right skill depends on the value of the motivations, previous experiences, and the relationship with the environment. All these elements have been modelled in order to be processed by the implemented decision making module.

The whole process can be summarized in the next steps:

1) Selecting the dominant motivation
2) Determining the state in the world
3) Selecting the feasible skills and executing the best one

In the following sections, these steps will be explained.

All the parameters set in this implementation will shape a specific personality for the robot. Changing these parameters, new personalities will be exhibited by the robot. The performance with different personalities will be studied in the future.

1) Which drives and motivations?: As expressed by equation (1), each motivation is represented by an integer value and it is affected by two factors: internal needs and external stimuli. Internal needs are the drives and their values depend on inner parameters. External stimuli are the objects situated in the environment that alter the robot motivations. In addition, each drive has its activation level: below it, motivations values will be set to zero and hence, they will not be considered for being the dominant motivation.

As mentioned, the internal needs, the drives, represent an internal value. Each motivation is connected to a drive. The choice about which drives (and consequently motivations too) must be implemented, were made at design time. Since the system has to be running on a robot intended to interact with people, some *social* motivation is needed to "push" the robot into human-robot interaction. Moreover, the authors want the robot to be endowed with play-oriented aspects, hence, a *recreational* nature is required by the robot. Nevertheless, the first primitive drive for all entities is to *survive* and, in our case, it is translated to the need of energy.

Therefore, the selected drives are:

- loneliness: the need of companion.
- boredom: the need of "fun" or entertainment.
- energy: this drive is necessary for survival.

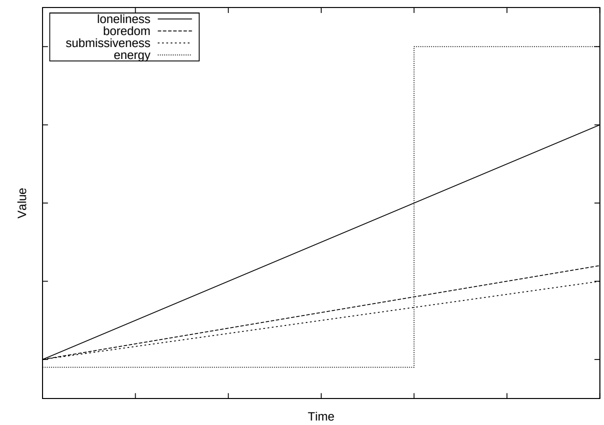

Fig. 5. Temporal evolution for all drives.

Based on the previous drives, the following non-conventional motivations have been defined:

- social: it means the need of interaction with a human and its drive is *loneliness*.
- recreational: this motivation is related to entertainment purposes. Its associated drive is *boredom*.
- survival: it refers to the energy dependence. This motivation is connected to the *energy* need.

All drives, and consequently motivations too, temporally evolve from their initial values. In our implementation (Figure 5), each drive can have a *satisfaction time*. This represents the period of time the drive remains at its initial value after it has been satisfied (look at the beginning of *loneliness* drive in Figure 5). During this time the drive does not evolve. After the satisfaction time, *loneliness* and *boredom* drives linearly increase but with different parameters. It means that, as time goes by, these drives become bigger and bigger, and so do the corresponding motivations. *Loneliness* is the fastest drive and *boredom* evolves slighter. This is because in social robots, as ours, interaction with people is one of the most relevant aims. Hence satisfaction time is very short and it is likely that *social* motivation will become the dominant motivation.

The *energy* drive is significantly different. This is the most relevant inner need due to the implicit necessity of *survival*. If we want to achieve a fully autonomous robot, power autonomy is the first step. Therefore, it will keep its initial value until a low battery level is detected. Then, at this point, its value will suffer a drastic raise.

In order to avoid an unstopped increase in the value of one of the motivations, a saturation level is defined for each one: once a motivation has reached its saturation value, it will not grow more. Different motivations have different saturation values which will determine the priority of the dominant motivation in case of a never-ending expansion of the motivations. In our implementation, *survival* is the first one, and the *social* and *recreational* motivations go after.

2) Sensing the world: The world is perceived by the robot in terms of objects and the states in relation to these objects (the external state). As a first approach, the world where Maggie is living in is limited to the laboratory. In this environment three objects have been defined: the people living

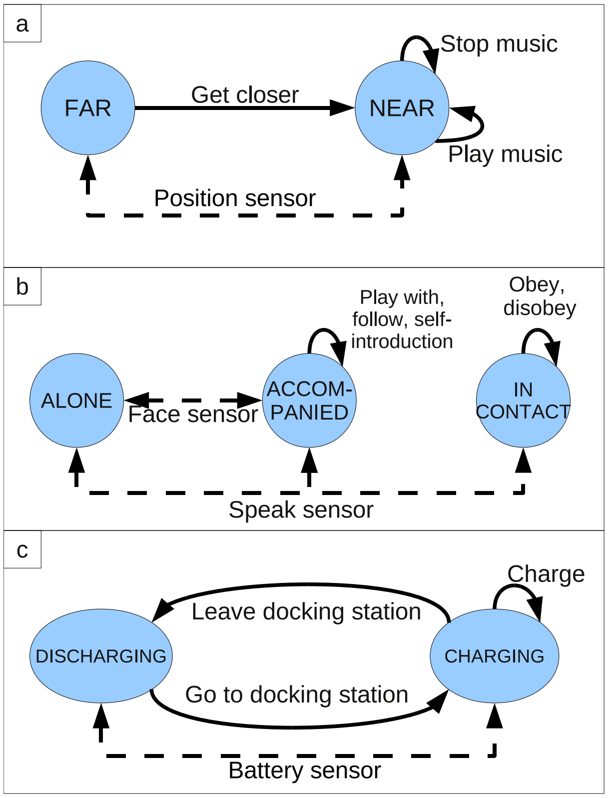

Fig. 6. States and actions for items: (a) TV, (b) person, and (c) docking station. The dashed arrows represent the skills that monitor the states, the continuous ones mean the actions executed with the objects, and the circles are the states related to each item.

around the robot, a television/radio appliance, and the docking station for supplying energy.

In Figure 6, the states related to each object, the actions, and the transitions from one state to another are shown. Dashed arrows represent the skills that monitor the states, continuous ones mean the actions executed with the objects, and the circles are the states related to each item. If an action does not appear at one state, it means that it is incoherent to execute it from that state, e.g., Maggie cannot *play music* if it is *far* from the TV or it cannot *follow* a person if it is *alone*. Figures 6.a, 6.b, and 6.c represent *TV*, *person*, and *docking station* objects, and their states and actions, respectively.

Television object

Since Maggie can control the TV appliance by means of an infrared interface [78], the robot must be placed at a certain distance and facing the appliance in order to be able to operate it. Therefore, two states related to the *television* item are defined: *close* and *far*, which symbolize the position where the robot is able to command the TV or not, respectively (Figure 6.a). In this case, these states are monitored by a skill that reads information from the navigation system which knows where the TV position is. From the *far* state, no other action than getting closer to the TV is possible; then, it is able to operate the appliance.

Person object

For the *person* item, three states are determined: *alone*, *accompanied*, and *commanded* (Figure 6.b). *Alone* represents when no people are around the robot, and this is the initial state with respect to people; *accompanied* means that someone is around Maggie but no direct interaction exists; and, finally, *commanded* state corresponds to a direct interaction between a person and Maggie where the user asks the robot to do something. The transitions from one state to another are detected by two skills integrated in the control architecture: a face detection skill and a speech recognition skill. The face detection skill notifies if the robot is alone or accompanied by searching for faces. The speech recognition skill listens to any word through a microphone and it distinguishes if the dialogue is directed to Maggie or not, determining if the robot is *accompanied* or *commanded*. If no words are received and no faces are found for a long time, the robot changes its state related to *person* to *alone*. In the experiments, this time is set to 10 seconds.

It needs to be mentioned that actions at *commanded* state, *obey* and *disobey*, are a bit special. These functions are designed to execute what a person has ordered to do or to reject the user's command. This implies that Maggie can disobey a instruction from the user and the robot will inform the user about it, reason why a direct dialogue between Maggie and the user is needed. Therefore, the *commanded* state is required for these actions.

Docking station object

In relation to the docking station, Maggie has two states: *charging* and *discharging*. If Maggie is connected to the station, then the battery level is increasing, so it is *charging*. Otherwise, if the robot is unplugged, it is *discharging* and the battery level decreases. This information is read by the battery sensor skill. When the robot is *discharging*, it can just *go to* charger. After that, it is plugged in and, once in the *charging* state, just one action is possible: to *charge* the batteries. If a skill that moves the robot around is selected when the robot is *charging*, it will leave the docking station and *discharging* will be the state.

External stimuli

Just like human beings can feel thirst when they see water, the motivations can be influenced by some objects present in the environment. These are called the external stimuli or incentives. These stimuli may have more or less influence: their values depend on the states related to the objects (this means, if they are near or far from the robot). In our implementation, all external stimuli values have been fixed empirically, and these values are shown in table I.

Due to the fact that Maggie is a very friendly robot and it loves people, the *social* motivation is affected when a person is near the robot, i.e., when the state of Maggie with respect to the person item is *accompanied* or *commanded*. Therefore, the total *social* motivation value will be increased by five units. Since our robot likes playing with people and dancing while listening to music from TV, the *recreational* motivation is increased by 10 units when people are around the robot or

TABLE I
THIS TABLE SHOWS EXTERNAL STIMULI, OBJECTS STATES LINKED TO THEM, THEIR VALUE, AND THE AFFECTED MOTIVATIONS

Motivation	Ext. stim.	State related to ext. stim.	Value
social	person	accompanied	5
		commanded	5
recreational	person	accompanied	10
	TV	near	10
survival	-	-	-

TABLE II
ACTIONS EFFECTS

Action	Object	Effect	Drive
follow	person	set to 0	loneliness
		-5	boredom
self-introduction	person	-10	loneliness
		-1	boredom
play with	person	set to 0	loneliness
		set to 0	boredom
Obey	person	-15	loneliness
Disobey	person	+15	loneliness
Get closer to	TV		
Play music	TV	set to 0	boredom
Stop music	TV	+5	boredom
Go to	docking station		
Charge	docking station	set to 0	energy
Leave	docking station		

when it is close to the TV. Up to now, the *survival* motivation does not have any external stimuli, but in the future, when the docking station is seen by the robot, it could feel the need of energy.

3) Acting in the world: Maggie interacts with the world through the objects and their potential actions. These actions are implemented as skills in the AD architecture. The possible actions with the *person* item are:

- following a person: Maggie will move following the closest person to it.
- self-introduction: the robot will introduce itself informing about its history and abilities.
- playing with: our robot will play several games with the user, such as tic-tac-toe, hangman, and animal-trivial [**?**].
- obey: Maggie will comply with the user request. The user can ask the robot to execute one of the previous actions.
- disobey: Maggie will deny the user request and it will inform the user about it.

About the TV appliance, its actions are:

- getting closer to: the robot moves towards TV.
- play music: Maggie turns on the TV and changes to a music channel.
- stop music: music is stopped and the TV is switched off.

In relation to the docking station, the possible actions are:

- go to: the robot plugs itself in the station.
- charge: Maggie keeps connected until batteries are full.

The actions cause effects over the drives. When the actions have ended, i.e., when the associated skill has been blocked because it has reached its goal, the effects are applied. If an error occurs during a skill execution, or it is not successful, this situation is notified and its effect is not applied. In the experiments presented in the last section, most of the effects affect one or more drives, which become zero, decrease or increase their value.

All effects are presented in Table II. These effects have been defined by the designer and any other values could have been selected. As it is shown, the *loneliness* drive is satisfied after the *follow* and *play with* actions are executed because both actions suggest a bidirectional interaction between the robot and a person. However, *self-introduction* could be accomplished without any response from the person so its effect decreases the *loneliness* drive by 10 units. Our robot's hobbies are playing and listening music, therefore, *play with* and *play music* satisfy the *boredom* drive. Moreover *stop music* increases *boredom* by five units. Other actions where a person is required (*follow* and *self-introduction*) affect in a lower factor (minus five and minus one respectively). **Besides *loneliness* is influenced by *obey* and *disobey* as well. When the robot complies with a user's request *loneliness* is reduced by fifteen units. On the other hand, *disobeying* an instruction rises the *loneliness* by fifteen units.** At last, the *energy* drive is satisfied when the batteries are recharged and this happens when the robot is at *charge* state.

4) What does Maggie do now?: Once the world has been presented, how the decision making system operates will be explained. First of all, when the system starts, the drives begin to evolve independently from their initial value, and the skills start monitoring the states related to items. When a new state transition is detected, a specific event is emitted and the states are written in the Short-Term Memory. The decision making module receives this event and the data of the states is updated. Within robot lifetime, the action selection loop is executed in order to determine the next skill to be activated. At each iteration, the dominant motivation is computed as the maximum motivation whose value (internal needs plus external stimulus) is over the activation level. This parameter has been fixed to 10 for every motivation. Using the dominant motivation, the current states related to objects, and the learnt values, the next action will be chosen.

As briefly described in section V, this approach has already been implemented on virtual agents. During these simulations and using the Q-learning reinforcement learning algorithm [22], the agent learnt the right q-values for maximizing its wellbeing. Taking those values as an inspiration, for the current implementation on Maggie, we propose a set of initial q-values that represent the best possible actions at each world configuration (the dominant motivation plus the state related to each object). The tuple formed by the dominant motivation, the object, the state related to the object, and the action with the highest q-value, will decide the selected action. In future implementations, the learning process will be carried out on-line by the robot itself and the initial q-values will be updated through the learning process.

As already stated, the available actions at each state depend on the state itself. Hence, each object-state pair will be associated to different actions. For example, for *playing with* people Maggie has to be *accompanied* with a person. Therefore, the *play with* action can not be activated when the robot is *alone*. At this point, these combinations do not exist and therefore, they will never be selected for execution.

VII. Experiments

In this section, some preliminary experiments are presented showing the performance of our system.

A. Policy of behavior: action selection

As previously mentioned, in this first implementation, the robot will use a predefined set of initial q-values for maintaining its internal needs within a determined range. In the real world, each action is connected to a skill. Therefore, depending on the configuration of the world, Maggie will execute one action or skill: it chooses the right action from all possible ones according to a *softmax* action selection rule [79]. The Softmax algorithm endows the decision making system with a certain randomness: the bigger the q-value for an action, the more likely this action will be selected. Again, in the future, when learning is done on-line, the q-values will change during the working life.

In Table III all the initial q-values are presented. These values represent the value for each action at each state. The actions in the table are the possible actions according to the objects states. Depending on the states, some actions have not been shown because they are not feasible: *follow* a person is not possible if the robot is alone, or music cannot be played if it is far from *TV*. For that reason, they will not be chosen.

If *social* is the dominant motivation, all actions related to the *person* item have high values. Actions connected to *TV* will be executed just when robot is *alone*, so no actions with persons are possible.

Focussing on the *recreational* motivation, "fun" skills will be likely executed whether a person is nearby the robot or Maggie is *near TV*. Fun skills have been defined by the authors and they are: *follow*, *self-introduction*, *play with*, and *play music*.

Concerning obedience, *obey* will probably be executed just in the *commanded* state. In other cases, all possible actions have the same probability of being run.

According to the good sense, if the dominant motivation is *survival*, actions concerning the docking station item will be probably selected. So, if the robot is *discharging*, it will *go to* the docking station and afterwards, when it is *charging*, it will be plugged in until its batteries are charged, i.e., *charge* skill.

The robot also must consider what to do when all its needs are satisfied. In our case, when a dominant motivation does not come up, depending on the state, the most reasonable skills will be to *charge* its batteries or to *go to* the docking station.

B. Evolution on Maggie

This experiment presents an example of how the motivation values change with time during Maggie's lifetime, see Figure 7. Most of motivations grow uniformly but, sometimes, jumps appear. These jumps are because of the presence of external stimuli as well as due to the effects of the actions on the drives (the numbers located on top of Figure 7 represent the executed action).

For example, focusing on the *recreational* motivation, we can observe a little increase at the beginning of the lifetime. This is because Maggie changed its state to *near* the TV. This is an external stimulus of this motivation and the *recreational* value is increased by ten units.

During the execution of action number four, *social* and *recreational* motivations raise at the same time. This is because the robot has detected the presence of some people (*accompanied* state).

TABLE III
PROPOSED Q-VALUES

Dominant motivation	**Object**	**States**	**Actions**	**Initial Q-value**
social	person	alone	-	-
		accompanied	follow	10
			self-introduction	10
			play with	10
		commanded	obey	10
			disobey	10
	TV	near	play music	5
			stop music	5
		far	get closer	5
	Docking st.	charging	charge	5
		discharging	go to	5
recreational	person	alone	-	-
		accompanied	follow	10
			self-introduction	10
			play with	10
		commanded	obey	5
			disobey	5
	TV	near	play music	10
			stop music	5
		far	get closer	10
	Docking st.	charging	charge	5
		discharging	go to	5
survival	person	alone	-	-
		accompanied	follow	1
			self-introduction	1
			play with	1
		commanded	obey	1
			disobey	1
	TV	near	play music	1
			stop music	1
		far	get closer	1
	Docking st.	charging	charge	10
		discharging	go to	10
none	person	alone	-	-
		accompanied	follow	0
			self-introduction	0
			play with	0
		commanded	obey	0
			disobey	0
	TV	near	play music	0
			stop music	0
		far	get closer	0
	Docking st.	charging	charge	5
		discharging	go to	5

Looking at the middle, the *social* and *recreational* motivations jump down quickly. At this point, Maggie was *accompanied* by a person and the executed action was *play*. The effects of this action are to set the *loneliness* and *boredom* drives to zero, and hence the *social* and *recreational* motivations fall, respectively. These motivations are not zero at this point because of the external state of the robot: *accompanied* state adds five points to social motivation and ten to recreational one. When the robot changes to the *commanded* state the external stimulus for these motivations disappear but a new equivalent one appears for social motivation so no jump comes up in this motivation.

Besides, we would like to mention the final part of the graph. Here, how the transition to the *far* state affects just the recreational motivation because this makes its external stimulus disappear is shown. Afterwards, the effect of *follow* action is pointed: it reduces the *recreational* motivation and satisfies the *loneliness* drive.

As shown, an action or state transition can affect several motivations, which means that effects are not attached to just one.

According to our initial goals, it is easy to appreciate that *social* is the fastest motivation.

At the multicolored band indicating the dominant motivation (upper Figure 7), a short black band stands out several times. This is the period of time when all motivations are satisfied: all drives are below their activation limits, and there is no dominant motivation. At this time, the *go to* action is executed because it is the most likely.

Finally, in relation to the *survival* motivation, since the most part of this experiment was executed with full batteries, this motivation is stable and very low almost all the time. But at the end, the battery level exceeds the limit and the survival motivation becomes the dominant one.

According to the initial q-values, Maggie knows that the best possible actions are *go to* and *charge*. Since Maggie is *charging* at that moment, the most probable action is *charge*. Finally, this is selected and executed.

Moreover, satisfaction time for *social* motivation is pointed at the end of the graph.

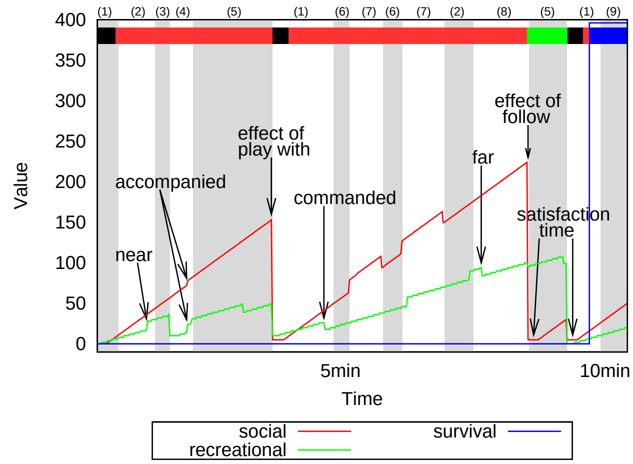

Fig. 7. Temporal evolution of motivations. Numbers on top represent the executed actions: (1)go to the docking station, (2)get closer to tv, (3)play music on tv, (4)stop music on tv, (5)play with a person, (6)disobey, (7)obey, (8)follow and (9)charge. The vertical white-grey bands at the background correspond to the execution time of each action. The upper colored band indicates the dominant motivation. Some action effects and changes of states are pointed.

VIII. Conclusions and Future works

In the last years, due to the increasing interest on social robots, cognitive systems have served as an inspiration for a new design of control architectures. In this work, we have presented a biologically inspired control architecture, in which the main decisions are made based on motivations and emotions. This architecture is an evolution from a previous one, the AD architecture, where the current goal of the robot was decided by an external operator, and some predefined sequences coordinated the robot behavior.

In the presented work, the new control architecture has been endowed with a decision making system based on biologically inspired concepts such as drives, motivation, emotions, and self-learning. Those concepts are included in order to improve the autonomy of the robot, as well as to try to imitate a living-being behavior. Up to now, the complete decision making system has been successfully tested on virtual agents.

This paper shows the implementation on a real robotic platform of a basic version of this architecture, where neither emotions nor self-learning are included in the decision making system. The experiments made on Maggie, a social robot designed for the interaction with humans, show that the robot is able to select the most appropriate skill autonomously, based on its own drives and motivations.

The next step of our work is to implement the learning process on Maggie. Therefore, the skill selection will be learnt by the robot through its interaction with the environment. Besides, the initial q-values could be set all to zero or it would be useful to give some prior knowledge. The last could be useful to improve the learning time based on some inheritance knowledge as living beings have.

Moreover, the learning process will take the role of some emotions as the reinforcement function (happiness, sadness) into account. On the other hand, fear will be implemented as another motivation for the robot. Nevertheless, it is expected that, due to the complexity of a real environment, the definition of new emotions, or re-definition of the old ones, will be needed.

Acknowledgment

The authors gratefully acknowledge the funds provided by the Spanish Government through the project called “Peer to Peer Robot-Human Interaction” (R2H), of MEC (Ministry of Science and Education), and the project “A new approach to social robotics” (AROS), of MICINN (Ministry of Science and Innovation). Moreover, the research leading to these results has also received funding from the RoboCity2030-II-CM project (S2009/DPI-1559), funded by Programas de Actividades I+D en la Comunidad de Madrid and cofunded by Structural Funds of the EU.

REFERENCES

[1] F. Littmann and J. Riviere, "A remote-operated system for interventions on explosives," in *Proceedings of the ANS Seventh Topical Meeting on Robotics and Remote Systems*, vol. 2, 1997, pp. 1038–42.

[2] A. Casals, R. Merchan, E. Portell, X. Cuf, and J. Contijoch, "Capdi: a robotized kitchen for the disabled and elderly," in *Proceedings of the Assistive Technology an the Threshold of the New Millennium. AAATE 99*, Dusseldorf, Germany, 1999, pp. 346–351.

[3] K. Schilling, M. Mellado, J. Garbajosa, and R. Mayerhofer, "Design of flexible autonomous transport robots for industrial production," in *Proceedings of the IEEE International Symposium on Industrial Electronics*, vol. 3, 1997, p. ISIE '97.

[4] N. Kubota, Y. Nojima, N. Baba, F. Kojima, and T. Fukuda, "Evolving pet robot with emotional model." Proceedings of the 2000 Congress on Evolutionary computation, 2000.

[5] R. M. Shiffrin, "Atention," *Stevens' Handbook of Experimental Psychology. Second Edition.*, vol. 2, 1988.

[6] R. M. Shiffrin and W. Schneider, "Controlled and automatic human information processing: In perceptual learning, automatic attending and a general theory," *Psicological Review*, pp. 127–190, 1997.

[7] J. LeDoux, *El cerebro emocional*. Ariel/Planeta, 1996.

[8] A. Damasio, *Descartes' Error - Emotion, reason and human brain*. Picador, London, 1994.

[9] S. C. Lewis, "Computational models of emotion and affect," Ph.D. dissertation, University of Hull, 2004.

[10] S. Gadanho, "Reinforcement learning in autonomous robots: An empirical investigation of the role of emotions," Ph.D. dissertation, University of Edinburgh, 1999.

[11] R. W. Picard, *Los ordenadores emocionales*. Ed. Ariel S.A., 1998.

[12] E. Rolls, *Emotion Explained*. Oxford University Press, 2005.

[13] R. C. Arkin, *Who needs emotions? The brain meets the robots*. Oxford University Press, 2004, ch. Moving up the food chain: Motivation and Emotion in behavior-based robots.

[14] K. L. Bellman, *Emotions in Humans and Artifacts*. MIT Press, 2003, ch. Emotions: Meaningful mappings between the individual and its world.

[15] L. Cañamero, *Emotions in Humans and Artifacts*. MIT Press, 2003, ch. Designing emotions for activity selection in autonomous agents.

[16] T. Ziemke and R. Lowe, "On the role of emotion in embodied cognitive architectures: From organisms to robots," *Cognitive Computation*, vol. 1, no. 1, pp. 104–117, 2009.

[17] C. Breazeal, *Designing Sociable Robots*. The MIT Press, 2002.

[18] R. C. Arkin, "Homeostatic control for a mobile robot: Dynamic replanning in hazardous environments," in *SPIE Conference on Mobile Robots, Cambrige, MAA*, 1988.

[19] T. Ziemke, "On the role of emotion in biological and robotic autonomy," *Biosystems*, vol. 91, no. 2, pp. 401–408, 2008.

[20] R. Barber, "Desarrollo de una arquitectura para robots moviles autonomos. aplicacion a un sistema de navegacion topologica." Ph.D. dissertation, Universidad Carlos III de Madrid, 2000.

[21] M. Malfaz, "Decision making system for autonomous social agents based on emotions and self-learning." Ph.D. dissertation, Carlos III University of Madrid, 2007.

[22] C. J. Watkins, "Models of delayed reinforcement learning," Ph.D. dissertation, Cambridge University, Cambridge, UK, 1989.

[23] M. Malfaz and M. Salichs, "Learning behaviour-selection algorithms for autonomous social agents living in a role-playing game," in *Proccedings of the AISB'06: Adaptation in Artificial and Biological Systems. University of Bristol, Bristol, England*, April 2006.

[24] ——, "Emotion-based learning of intrinsically motivated autonomous agents living in a social world," in *Proccedings of the ICDL 5: The Fifth International Conference on Development and Learning, Bloomington, Indiana*, June 2006.

[25] ——, "The use of emotions in an autonomous agent's decision making process." in *Ninth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems (EpiRob09). Venice. Italy*, 2009.

[26] K. C. Berridge, "Motivation concepts in behavioural neuroscience," *Physiology and Behaviour*, no. 81, pp. 179–209, 2004.

[27] C. L. Hull, *Principles of Behavior*. New York: Appleton Century Crofts, 1943.

[28] J. Santa-Cruz, J. M. Tobal, A. C. Vindel, and E. G. Fernández, "Introduccin a la psicologa," 1989, facultad de Psicologa. Universidad Complutense de Madrid.

[29] A. Ortony, D. A. Norman, and W. Revelle, *J.M. Fellous and M.A. Arbib, Who needs emotions: The brain meets the machine*, 2005, ch. Affect and proto-affect in effective functioning.

[30] J. Fellows, "From human emotions to robot emotions," AAAI 2004 Spring Symposium on Architectures for Modelling Emotion: Cross-Disciplinary Foundations.SS-04-02. AAAI Press., Tech. Rep., 2004.

[31] A. Kelley, *Who Needs Emotions? The Brain Meets the Robot*. Oxford University Press, 2005, ch. Neurochemical networks encoding emotion and motivation: an evolutionary perspective.

[32] G. Hollinger, Y. Georgiev, A. Manfredi, B. Maxwell, Z. Pezzementi, and B. Mitchell, "Design of a social mobile robot using emotion-based decision mechanisms," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, 2007, pp. 3093–3098.

[33] R. S. Lazarus, *Appraisal processes in emotion: Theory, methods, research*. New York: Oxford University Press, 2001, ch. Relational meaning and discrete emotions, pp. 37–67.

[34] J. S. de Freitas and J. Queiroz, *Advances in Artificial Life*, 2007, ch. Artificial Emotions: Are We Ready for Them?, pp. 223–232.

[35] S. Gadanho, "Learning behavior-selection by emotions and cognition in a multi-goal robot task," *The Journal of Machine Learning Research. MIT Press Cambridge, MA, USA*, no. 4, pp. 385–412, 2003.

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

[37] R. Plutchik, *Emotion: a psycho evolutionary synthesis*. Harper & Row, New York, 1980.

[38] M. Scheutz, "Useful roles of emotions in artificial agents: a case study from artificial life." in *AAAI 2004*. AAAI press, Menlo Park, 2004, pp. 42–48.

[39] C. Delgado-Mata and R. Aylett, "Emotion and action selection: Regulating the collective behaviour of agents in virtual environments," in *3rd International Joint Conference on Autonomous Agents and Multiagent Systems*, 2004.

[40] D. Cañamero, "Emotions and adaptation in autonomous agents: a design perspective," *Cybernetics and systems: International Journal*, vol. 32, pp. 507–529, 2001.

[41] J. S. Albus, C. R. McLean, A. J. Barbera, and M. L. Fizgerald, "Hierarchical control for robots and teleoperatrors," in *Proceedings of the Workshop on Intelligent Control*, 1985.

[42] G. Saridis, "Control prfomance as an entropy: an integrated theory for inteligent machines," Robotics and Automation Laboratory, New York., Tech. Rep., 1983.

[43] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal of Robotics and Automation*, pp. 14–24, April 1986.

[44] ——, "Intelligence without reason," *Artificial Intelligence*, pp. 139–159, 1991.

[45] M. Mataric, "Behaviour-based architectures for inteligent control," in *Workshop on Inteligent Autonomours Control System*, November 1992.

[46] J. R. Firby, "Building symbolic primitives with continuous control routines," in *In First International Conference on AI Planning Systems.* College Park, M.D., June 1992.

[47] ——, "Task networks for controlling continuous processes," in *In sensorial International Conference on AI Planning Systems.*, E. K. Hammond., Ed. AAAI Press, 1994, pp. 49–54.

[48] R. P. Bonasso, J. Firby, E. Gat, D. Kortenkamp, D. P. Miller, and M. G. Slack, "Experiences with an architecture for inteligence, reactive agents," *Journal of Experimental Theory of Artificial Intelligence.*, vol. 9, pp. 237–256, 1997.

[49] D. Schreckenghost, P. Bonasso, D. Kortenkamp, and D. Ryan, "Three tier architecture for controlling space life support systems," in *IEEE Symposium on Intelligence in Automation and Robotics*, May 1998, pp. 195–201.

[50] E. Gat, "Alfa: A language for programming reactive robotic control systems," in *IEEE Conference on Robotics and Automation*. Los Alamitos, CA.: IEEE Computer Society Press., 1991.

[51] ——, "Integrating planning and reacting in a heterogeneous asynchronous architecture for controlling real-world mobile robots," in *In Tenth National Conference on Artificial Intelligence.* AAAI. San Jose, CA., July 1992.

[52] ——, "Esl: A language for suporting robust plan execution in embedded autonomous agents," in *In Proceedings of the IEEE Aerospace Conference.* Los Alamitos, CA.: IEEE Computer Society Press., 1997.

[53] R. Alami, R. Chatila, S. Fleury, M. Ghallab, and F. Ingrand, "An architecture for autonomy," *The International Journal of Robotics Research*, vol. 17, no. 4, pp. 315–337, 1998.

[54] M. A. Salichs and R. Barber, "A new human based architecture for intelligent autonomous robots." in *4th IFAC Symposium on Intelligent Autonomous Vehicles*, 2001, pp. 85–90.

[55] L. Cañamero, "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.

[56] ——, “Designing emotions for activity selection,” Dept. of Computer Science Technical Report DAIMI PB 545, University of Aarhus, Denmark, Tech. Rep., 2000.
[57] S. Gadanho and J. Hallam, “Emotion- triggered learning in autonomous robot control,” *Cybernetics and Systems*, vol. 32(5), pp. 531–59, 2001.
[58] S. Gadanho and L. Custodio, “Asynchronous learning by emotions and cognition,” in *From Animals to Animats VII, Proceedings of the Seventh International Conference on Simulation of Adaptive Behavior (SAB’02), Edinburgh, UK*, 2002.
[59] J. Velásquez, “Modeling emotions and other motivations in synthetic agents,” in *Fourteenth National Conf. Artificial Intelligence*, 1997.
[60] ——, “Modelling emotion-based decision-making,” in *1998 AAAI Fall Symposium Emotional and Intelligent: The Tangled Knot of Cognition*, 1998.
[61] C. Breazeal and L. Aryananda, “Recognition of affective communicative intent in robot- directed speech,” *Autonomous Robots*, vol. 12, pp. 83–104, 2002.
[62] C. Breazeal and R. Brooks, *Who Needs Emotions: The Brain Meets the Robot*. MIT Press, 2004, ch. Robot Emotion: A Functional Perspective.
[63] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, and B. Blumberg, “Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots,” *Artificial Life*, vol. 11, pp. 1–32, 2005.
[64] E. D. Sevin and D. Thalmann, “A motivational model of action selection for virtual humans,” in *Proceedings of Computer Graphics International*, 2005, pp. 213– 220.
[65] J. Hirth, T. Braun, and K. Berns, *KI 2007: Advances in Artificial Intelligence*, 2007, ch. Emotion Based Control Architecture for Robotics Applications, pp. 464–467.
[66] C. L. Lisetti and A. Marpaung, *KI 2006: Advances in Artificial Intelligence*, 2007, ch. Affective Cognitive Modeling for Autonomous Agents Based on Scherer’s Emotion Theory, pp. 19–32.
[67] M. J. L. Boada, R. Barber, and M. A. Salichs, “Visual approach skill for a mobile robot using learning and fusion of simple skills,” *Robotics and Autonomous Systems*, vol. 38, pp. 157–70, March 2002.
[68] R. C. Atkinson and R. M. Shiffrin, *The Psychology of Learning and Motivation*. K. W. Spence and J. T. Spence. New York: Academic Press, 1968, vol. 2, ch. Human Memory: A Proposed System and Its Control Processes, pp. 89–195.
[69] C. Balkenius, “Motivation and attention in an autonomous agent,” in *Workshop on Architectures Underlying Motvation and Emotion WAUME 93, University of Birmingham*, 1993.
[70] ——, “Natural intelligence in artificial creatures,” Ph.D. dissertation, Lund University Cognitive Studies 37, 1995.
[71] K. Lorenz and P. Leyhausen, *Motivation of human and animal behaviour; an ethological view*. New York: Van Nostrand-Reinhold, 1973, vol. xix.
[72] A. Ortony, *Emotions in Humans and Artifacts*. MIT Press, 2003, ch. On making Believable Emotional Agents Believable, pp. 188–211.
[73] E. Rolls, *Emotions in Humans and Artifacts*. MIT Press, 2003, ch. Theory of emotion, its functions, and its adaptive value.
[74] B. Zimmerman, “http://www.coffeemud.org/,” 2007.
[75] L. Cañamero, “Emotion understanding from the perspective of autonomous robots research,” *Neural Networks*, vol. 18, pp. 445–455, 2005.
[76] M. A. Salichs, R.Barber, A. M.Khamis, M.Malfaz, J. F.Gorostiza, R.Pacheco, R.Rivas, A.Corrales, and E.Delgado, “Maggie: A robotic platform for human-robot social interaction,” in *IEEE International Conference on Robotics, Automation and Mechatronics (RAM 2006). Bangkok. Thailand*, 2006.
[77] R. Rivas, A. Corrales, R. Barber, and M. A. Salichs, “Robot skill abstraction for ad architecture,” in *6th IFAC Symposium on Intelligent Autonomous Vehicles*, 2007.
[78] J. Salichs, A. Castro-Gonzalez, and M. A. Salichs, “Infrared remote control with a social robot,” in *FIRA RoboWorld Congress 2009*, Springer, Ed. Incheon, Korea.: Springer, August 2009.
[79] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, A Bradford Book, 1998.