

Simulating empathic behavior in a social assistive robot

Berardina De Carolis¹ · Stefano Ferilli¹ ·
Giuseppe Palestra¹

Received: 10 December 2015 / Revised: 15 July 2016 / Accepted: 19 July 2016 /

Published online: 27 September 2016

© Springer Science+Business Media New York 2016

Abstract When used as an interface in the context of Ambient Assisted Living (AAL), a social robot should not just provide a task-oriented support. It should also try to establish a social empathic relation with the user. To this aim, it is crucial to endow the robot with the capability of recognizing the user's affective state and reason on it for triggering the most appropriate communicative behavior. In this paper we describe how such an affective reasoning has been implemented in the NAO robot for simulating empathic behaviors in the context of AAL. In particular, the robot is able to recognize the emotion of the user by analyzing communicative signals extracted from speech and facial expressions. The recognized emotion allows triggering the robot's affective state and, consequently, the most appropriate empathic behavior. The robot's empathic behaviors have been evaluated both by experts in communication and through a user study aimed at assessing the perception and interpretation of empathy by elderly users. Results are quite satisfactory and encourage us to further extend the social and affective capabilities of the robot.

Keywords Social assistive robots · Affective computing · Empathy

1 Introduction

A recent trend in the Ambient Assisted Living (AAL) domain is to employ social robots (<http://ksera.ieis.tue.nl>; http://www.aal-europe.eu/wp-content/uploads/2012/04/AALCatalogue_onlineV4.pdf) [5, 35, 37, 41]. Usually, such robots support elderly people in

✉ Berardina De Carolis
berardina.carolis@uniba.it

Stefano Ferilli
stefano.ferilli@uniba.it

Giuseppe Palestra
giuseppe.palestra@uniba.it

¹ Dipartimento di Informatica, Univerisita' di Bari, Aldo Moro, Via Orabona 4, 70126 Bari, Italy

their daily living and provide them an interface toward the technology that is present in the environment. Due to their potential to involve users in human-like conversations, showing empathy and emotions in their behavior [43], social robots seem appropriate to support the interaction in AAL environments where people live and stay for a long time. For this reason, besides providing task-oriented support, AAL environments should establish a social long-term empathic relation with the user and increase the user's trust. Empathy, broadly defined as the feeling by which one understands and shares another person's experiences and emotions, facilitates the development of social relationships, affection and familiarity [4, 29].

In the context of the NICA¹ (Natural Interaction with a Caring Agent) project, we developed the behavioral architecture of a social robot, embodied in the NAO robot by Aldebaran (<http://www.aldebaran-robotics.com>), that is able to act as a natural and social interface between the user and the smart home environment, by implementing an easy access to the (digital or physical) services of the environment [see 2 and 47 for more details]. Moreover, we endowed the robot with the capability of being empathic toward the user. Indeed, it has to act as a virtual carer and, according to [6, 16, 36], empathy plays a key role for therapeutic improvement.

Early definitions of empathy are grounded on the concept of taking the role of another or as “an affective response more appropriate to someone else's situation than to one's own” [17]. Davis defines empathy as the “reactions of one individual to the observed experiences of another” [9]. Recently empathy has been conceptualized as having two dimensions: an affective (emotional response to the perceived emotional experience of others) and a cognitive (understanding of another person's situation) one. So, the expression of empathy should demonstrate that the other's feelings are understood or shared and show an emotional response that is appropriate to the situation [1]. Leite et al. concluded that affective and cognitive empathy, together, are effective in making the robot appear as empathic [22].

In this paper the model of empathy defined in [2] was considered as a guideline for developing the empathic behavior in NICA. In this model, the distinction between cognitive empathy, i.e. understanding how another feels, and affective empathy, that is an active emotional reaction to another's affective state, is considered to be important for a complete definition of empathic behavior.

To simulate the cognitive aspect of empathy in NICA, the robot had to be endowed with the capability of recognizing and understanding the affective state of the user. Due to the conversational nature of the interaction between the user and the robot, the user's affective state is recognized by analyzing communicative signs extracted from speech and facial expressions. Then, in order to show an appropriate empathic reaction to the user's affective state, NICA reasons on it for triggering its own social emotion and, consequently, for deciding which empathic goal to achieve by executing the most appropriate plan of actions. Reasoning has been modeled using Dynamic Belief Networks (DBNs) [18], a formalism that is suitable for representing situations that gradually evolve from a step to the next one, allowing one to simulate probabilistic reasoning and to deal with uncertainty which is typical of natural situations. In this way NICA is able to simulate the affective component of empathy, since a social emotion is triggered in the robot as a consequence of its perception of the user's state.

The resulting prototype was tested both by expert evaluation and through user study. Results of both evaluations showed that the behaviors of the robot were appropriate for conveying the intended empathic response. In particular, the user study showed that the cognitive component of empathy was clearly recognized, while its affective component was

¹ The name of the project, NICA, is also used to indicate the robot.

not always perceived as expected. Indeed, sometimes the robot's expression, related to the perception of the user's state, was considered as inappropriate.

This paper is structured as follows. After providing an overview of related work in Section 2, in Section 3 we describe how the empathic behavior in the robot was modeled and simulated. Then, in Section 4, a brief illustration of a case study is provided. Section 5 explains how the evaluation was conducted and presents its results. Finally, we conclude the paper with discussion and directions for future work in Section 5.

2 Related work

The problem of taking care of the elderly is becoming extremely relevant, because significant demographical and social changes affected our society in the last decades, and population aged between 65 and 80 will increase significantly over the next decade. The aging of population affects both our society and the elderly themselves: on one hand, there is a growing demand for caregivers; on the other hand, the ability of elderly people to live independently becomes more difficult to achieve. Moreover, the elderly usually prefer to live at home, possibly in an environment they feel comfortable in. In this perspective, ambient intelligence provides a great opportunity to improve their quality of life by providing cognitive and physical support, and easy access to the services of the environment (http://www.aal-europe.eu/wp-content/uploads/2012/04/AALCatalogue_onlineV4.pdf) [6, 27, 36]. In this context, interaction with environment's services may be provided mainly in two ways: i) in a seamless way, using proactive and pervasive technologies for supporting the user's interaction through an ecosystem of devices, or ii) using an embodied companion as a conversational interface. The use of the latter approach was investigated by van Ruiten et al. [41]. By conducting a controlled study, they confirmed that elderly users like to interact with social robots and to establish a social relation with them [5]. Successful results on how conversational agents and robots can be employed as an interface in this domain are reported also in several projects. For instance, projects *ROBOCARE* [6], *Nursebot* [36], *Care-o-bot* [16], *CompaniAble* (<http://www.companionable.net>), *KSERA* (<http://ksera.ieis.tue.nl>) and *GiraffiPlus* [7] aim at creating assistive intelligent environments in which robots offer support to the elderly at home, possibly also having a companion role. This success is due to the interaction between humans and machines having a fundamental social component [37]. Moreover, embodied agents have the potential to involve users in a human-like conversation using verbal and non-verbal signals for providing feedback, showing empathy and emotions in their behavior [10, 11, 21].

However, the social component of the interaction requires the development of user models that involve reasoning on both components (cognitive and affective) of the user's state of mind. In our opinion this is a key issue for enabling the adaptation of the robot's behavior to both physical and emotional user's needs, as in the case of the simulation of the empathic behavior. To have a baseline to start for developing such a behavior it was necessary to look at available definitions of empathy.

Empathy is seen as the ability to perceive, understand and experience what others are feeling, and to communicate such an understanding to them [35]. In particular, Baron-Cohen distinguishes between cognitive and affective empathy. Cognitive empathy refers to the understanding of how another feels, while affective empathy represents an active emotional reaction to another's affective state [2]. In the field of human-robot interaction, researchers have demonstrated the benefits of empathy in robot behavior design. There are many examples of robots that include empathy in their behaviors. Many of them address only the affective dimension of empathy. For instance, Paiva [33] defines empathic agents as "agents that

respond emotionally to situations that are more congruent with the user's or another agent's situation, or as agents that, by their design, lead users to emotionally respond to the situation that is more congruent with the agent's situation than the user's one". However, as recognized in [40], the cognitive component of empathy seems to be the most relevant one to attribute an empathic behavior to a robot. In [23], the possible role of empathy in socially assistive robotics is discussed. Tapus and Matarić propose that an empathic robot must be capable of recognizing the user's emotional state and that, on the other hand, it must express emotions. Leite et al. [22] propose a multimodal framework for modeling some of the user's affective states in order to personalize the learning environment by adapting a robot's empathic responses to the particular preferences of the child who is interacting with it. Cramer et al. showed that the iCat robot, when expressing empathic behaviors, was perceived as more trustworthy [8], and establishing a relation of trust is extremely important in the AAL domain.

Since we use the NAO robot, we looked at research work in which the robot was used for simulating empathic behaviors. In our study, NAO supports elderly people in their daily living, reasoning on their affective state and providing them an interface toward the technology that is present in the environment. NAO increases the acceptance and adoption of service robots in domestic environments. It can be connected to both indoor and outdoor environmental sensors and to vital signs measurement devices (<http://ksera.ieis.tue.nl>). According to the results of the pilot study described in [40], the robotic embodiment of NAO is not perceived, by itself, as empathic. However, when NAO is expressing affective behaviors and shows to understand the user's emotion, then the perception of empathy increases. Therefore, it is important to endow NAO with the capability of recognizing as precisely as possible the emotional state of the user, because a wrong recognition may compromise empathy. On the other hand, to ensure that the user understands the empathic goal of the communication, it is also important that the generated behaviors are accurately designed. To this aim, we developed the architecture of the robot's mind by including the affective reasoning at two different levels: i) at the recognition level, by perceiving and recognizing the user's affective state, and ii) at the feeling level, by endowing the robot with beliefs about its own emotions as a consequence of what has been recognized.

3 Simulating empathic behaviors in Nica

Our approach to the simulation of empathy is outlined in Fig. 1. NICA has to monitor the user's communicative behavior and, by analyzing his speech and facial expressions, it has to recognize the user's emotion. The recognized emotion and the user's communicative act trigger the robot's affective state that concerns its social emotions.

Both the emotional state of the user and that of the robot are then used to select the most appropriate empathic goal and to pursue such a goal through the execution of a behavior plan containing communicative acts.

Since the combination of verbal and non-verbal communication provides social cues that make the robot appear more intuitive and natural [28], a communicative act can be expressed using different modalities (voice, facial expressions, and body movements and gestures). The reasoning to decide which behavior should be performed is based on probabilistic models, since uncertainty is typical of this domain (e.g., when dealing with the evolution of the user's affective state over time).

We now describe in more details how the various components of the behavioral architecture were developed.

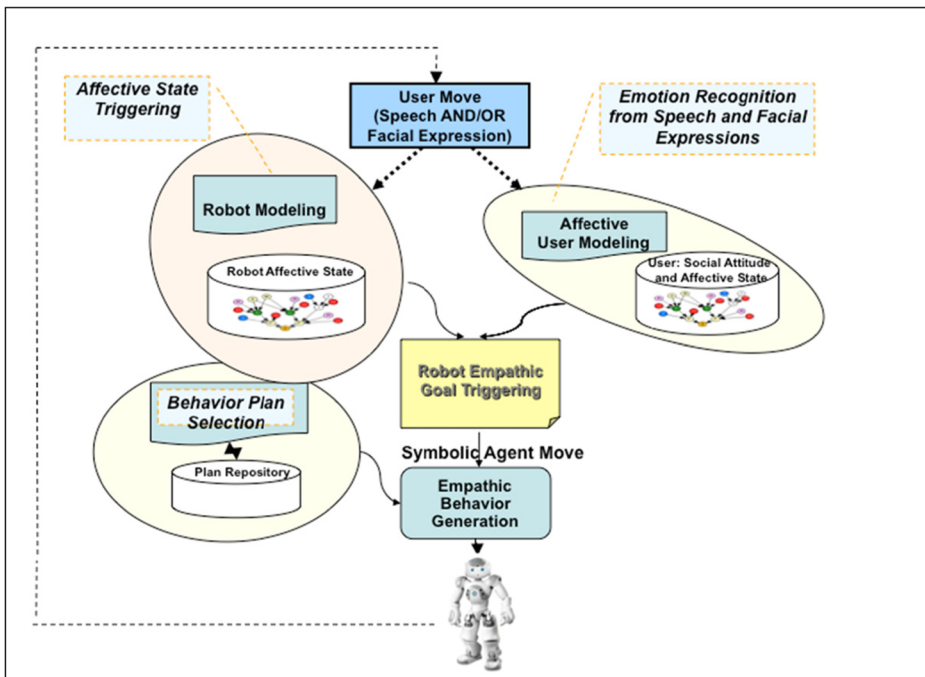


Fig. 1 Simulation of the empathic behavior in the robot

3.1 Recognizing the user's affective state

Humans express their emotions and perceive the emotions of others through multiple channels: speech intonation, face expressions, gestures and body postures. According to [24], most affective information is expressed through non-verbal communication. The vocal part and the facial expression contribute 38% and 55%, respectively, in conveying the emotional content. Since our approach towards simulating empathy with a socially assistive robot is based on the understanding of the emotions of the human users, to recognize them we analyze the combination of facial expressions with the speech prosody.

3.1.1 Facial expressions

Our expression detection system is able to identify Ekman's six basic emotions as defined by the Facial Action Coding System [14]: anger, disgust, fear, happiness, sadness, surprise. The facial expression recognition system processes the sequence of images captured by the robot camera and then applies a three-step procedure that consists of: Face Detection, Facial Features Extraction, and Facial Expression Classification.

Face Detection is a two-class problem that consists in determining whether a face is present in an image. Our system uses the well-known Viola and Jones face detector [42], that minimizes computational cost and achieves high detection performance. Then, the detected face is processed for facial features extraction.

The Staked Active Shape Model (STASM) approach [25] is used for locating 77 facial key points. STASM uses the Active Shape Model with a simplified form of SIFT descriptors and

with Multivariate Adaptive Regression Splines (MARS) [26] for descriptor matching. Given the facial key points, 32 geometrical features (linear, polygonal, elliptical and angular) are calculated. Linear features are defined by the Euclidean distance between two key points. The system uses 15 linear features to describe all parts of a face:

- 6 for the upper face: 3 for the left eyebrow, 2 for the left eye and 1 for the cheeks;
- 9 for the lower face: 1 for the nose and 8 for the mouth.

Polygonal features are determined by the area of irregular polygons constructed on 3 or more facial key points. In this case, 3 polygonal features are defined:

- 1 for the left eye;
- 1 between the corners of the left eye and the left corner of the mouth;
- 1 for the mouth.

Furthermore, 7 elliptical features are computed using the major and minor axes ratio:

- 1 for the left eyebrow;
- 3 for the left eye: eye, upper and lower eyelids;
- 3 for the mouth: mouth, upper and lower lips.

Finally, 7 angular features are defined by the angular coefficient of the straight line through two facial key points:

- 1 for the left eyebrow;
- 6 for the mouth corners.

The last step performed by the facial expression recognition system is classification. The facial expression classification module analyzes the geometrical facial features in order to predict the captured facial expression.

The facial expression recognition system adopted is fully automatic. It achieves a recognition rate of 95.46% using a Random Forest Classifier [34] and six-class expression classification.

3.1.2 *Speech prosody*

The analysis of the user's spoken utterance is made according to the dimensional approach, by using two classifiers to recognize the valence and arousal dimensions [39]. To this aim we developed a Web service called VOCE (VOIce Classifier of Emotions), that analyzes and classifies the voice prosody using an approach that is similar to the one described in [34]. In particular, the valence dimension is classified from positive to negative along a 4-point scale (from 1 = very negative to 4 = positive). Arousal is classified in a 3-point scale (high, medium, low).

The audio files of the spoken user's sentences in the corpus were analyzed using Praat functions [4] to extract features associated to:

- the variation in fundamental frequency (pitch min, mean, max and standard deviation, slope);
- the variation of energy and harmonicity (min, max and standard deviation);

Table 1 F1- measure for arousal classification

| Class | F1-measure C4.5 | F1-measure K-NN |
|---------|-----------------|-----------------|
| high | 0.89 | 0.90 |
| medium | 0.79 | 0.75 |
| low | 0.77 | 0.78 |
| average | 0.81 | 0.81 |

- the central spectral moment, standard deviation, center of mass, skewness and kurtosis; and
- the speech rate.

In developing VOCE we exploited several classification algorithms. An analysis of their performance showed that C4.5 and K-NN were the most accurate ones. Tables 1 and 2 report the F1 measure figures for arousal and valence, respectively (validated using a 10-Fold Cross Validation technique).

Based on these results, we set VOCE to use a K-NN classifier.

Looking at the F1-measure values, high arousal and very negative valence (corresponding to anger) are very well recognized, while intermediate values (neutral valence and medium arousal) have lower F1-measure and therefore their recognition is less accurate.

3.2 Affective user modeling

In developing the empathic reasoner of a social robot and defining feasible behaviors for it, we integrated two kinds of knowledge: knowledge coming from the experience of human caregivers and knowledge about the role of empathy in assisting elderly people [3, 9]. We had to purposely collect such a knowledge, since most existing datasets for AAL take into account mainly data that allows to recognize activities or routines and do not emphasize affective factors and affective reactions to other people's actions (i.e., the behavior of caregivers in response to the recognized affective state).

3.2.1 Collecting expert experience

In order to collect experience from human caregivers, we used the diary approach. In particular, two human caregivers recorded daily their experience during the assistance of two elderly women, both affected by a chronic disease, during a period of one month for 12

Table 2 F1- measure for valence classification

| Class | F1-measure C4.5 | F1-measure K-NN |
|---------------|-----------------|-----------------|
| positive | 0.81 | 0.84 |
| neutral | 0.78 | 0.81 |
| negative | 0.81 | 0.78 |
| very negative | 0.92 | 0.94 |
| average | 0.83 | 0.84 |

hours/day. Those women lived alone and had a son/daughter who could intervene only in case of need for solving relevant medical and logistic problems.

Data was collected using a paper-diary on which the caregiver had to annotate two categories of entries: the schedule of the daily tasks and the relevant events of the day, with particular regards to affective ones, using a schema like the one reported in Table 3.

In the diary schema, each row represents a relevant event along with the attributes for describing it and the action performed by the caregiver when this event occurred. For example, let's consider the last row. At 18:00 (*time*) Maria is sad (*event*). The caregiver inferred that Maria's state is sadness (*recognized affective state*) since she was moaning, saying "Oh my, oh my" (*signs*) with a sad voice and face. The caregiver did not know the *reason* of Maria's sadness, so she went toward Maria (*action*) trying to console her by saying "I'm sorry to see you are sad. Why are you so sad?" (*communicative action*). After her intervention, she noticed that Maria was less sad (*effect*).

Overall, the collected corpus included about 900 entries. We used this data to extract the knowledge needed to build the reasoning model of the agent, so as to make its behavior believable. Specifically, we tried to understand which events and context conditions are relevant to goal and action triggering, and when affective and social factors are important in real-life scenarios.

Albeit we collected the experience of only two human caregivers, the period they spent with elderly people and the amount of data collected were large enough for providing us the knowledge for defining situation-oriented action plans and dialogue strategies. Among other things, we discovered that the emotions considered important by the caregivers were a subset of Ekman's six basic emotions, and specifically: anger, sadness and happiness. Therefore, in developing the model for reasoning on the user's affective state, we considered only these three emotions, plus the neutral state. In particular, we determined the combinations of valence and arousal that allow us to distinguish anger (negative or very negative valence with a high arousal), sadness (negative valence with a medium-low arousal), and happiness (positive valence).

3.3 Reasoning on the user's affective state

As said, the robot's belief about the user's affective state is monitored with a model based on DBNs [18]. In fact, when modeling affective phenomena we must take into account the fact that the affective state smoothly evolves during the interaction, from one step to the subsequent one, and that the state at any time during the interaction depends on the state in the previous turn. Indeed, the DBN is a suitable formalism for representing situations that gradually evolve from a dialog step to the next one. Moreover, DBNs allow one to simulate probabilistic reasoning and to deal with uncertainty in the relationships among the variables involved in the inference process.

Our model, illustrated in Fig. 2, is exploited to infer which is the most probable emotional state the user is experiencing at every step of the interaction, by monitoring signals coming from speech and facial expressions. During the daily life in a home environment, each communicative act of the user is analyzed in terms of prosody, and facial expressions are detected. These sensed data become the evidences of the model. The values of the corresponding nodes are entered and propagated in the network to recognize the user's emotion and the overall polarity of her affective state. Since this component of the affective user model allows us to recognize the user's affective state, it can be exploited to implement the cognitive component of empathy.

Table 3 Some entries from the caregivers' paper-diary

| Time | Event | Signs | Reason | Action | Communicative action | Recognized affect | Effect |
|-------|------------------|--|---------------|---|---|-------------------|---|
| 9.00 | ... | ... | medical visit | I remind Maria about the appointment at the doctor at 10.00. | Remind "Maria, I would like to remind you that today you have an appointment at the doctor at 10.00 in the morning." | ... | ... |
| 9.30 | ... | ... | medical visit | I ask and help Maria to dress up according to the weather condition. | Ask for "Today is a wonderful summer day. You can put on your beautiful dress that you like so much!" | ... | Maria is dressed |
| 9.40 | ... | ... | medical visit | I send a reminder to Maria's daughter about the medical visit of her mother. | ... | ... | The daughter answered that she is coming. |
| 9.45 | Maria is worried | Sit down, Tapping fingers | medical visit | I go toward Maria and try to encourage her. | Encourage "Come on, don't worry! You will not have any problem, for sure." | anxious | Maria is less anxious |
| 18.00 | Maria is sad | Sit down, Moaning "Oh My, poor me..." Sad voice Sad face | Unknown | I go toward Maria and try to console her and to understand the reason of her affective state. | Console "I'm sorry to see you sad. Why are you sad?" ... | sadness | Maria is less sad |

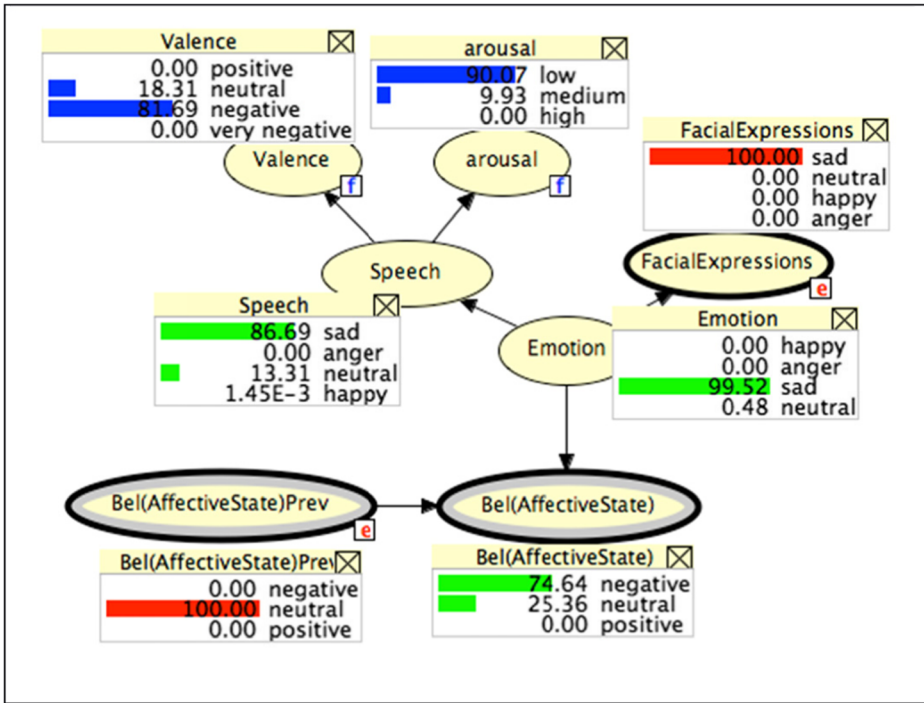


Fig. 2 The DBN model of the robot’s beliefs about the user affective state

The probability of the dynamic variable ($Bel(AffectiveState)$), representing the valence of user’s affective state, is exploited by the agent to check the whether there has been a change in the user’s affective state, thus causing the activation of the empathic goal. The dependence of this state, at any time, from the state at the previous time is expressed by a temporal link between the $Bel(AffectiveState)Prev$ and the $Bel(AffectiveState)$ variables, that represent the belief of the robot about the valence of the user’s affective state at times $t-1$ and t , respectively.

Figure 2 also reports an example of how the model is used. In this case the facial expression denotes sadness for sure (probability 1.0) and the speech turns out to be sad with a high probability (0.86) because the voice has a negative valence with a low arousal. These values are propagated in the model, determining the probability of the emotion node and then of the belief of the robot that the user is in a negative affective state with a probability of about 0.75.

3.4 Emotion triggering in the robot’s mind

In many systems, the empathic behavior of the agent is driven only by the emotion recognized in the user. However, in order to simulate a realistic empathic behavior and, in particular, its affective component, we aim at triggering an affective state in the robot, as a consequence of the affective state recognized in the user, for simulating the activation on empathic goals. In this context, we consider social emotions.

In particular, the adopted emotion modeling approach is based on event-driven emotions according to Ortony, Clore and Collin’s (OCC) theory [31]. In this theory, positive emotions (happy-for, hope, joy, etc.) are activated by desirable events while negative emotions (sorry-for, fear, distress, etc.) arise after undesirable events. In addition, we considered also Oatley and

Johnson-Laird’s theory in which positive and negative emotions are activated, respectively, by the belief that some goal will be achieved or will be threatened [30]. In modeling NICA’s emotions, we considered those concerning *Well-being* (joy, distress), *Attribution* (admiration, reproach), *FortuneOfOthers* (happy-for, sorry-for) and a combination of these categories (anger).

The model of emotion activation is represented using a DBN as well, since we need to reason about the consequences of the observed event on the monitored goals in successive time slices. As explained in more detailed in [12], the intensity of the emotion felt by the robot is computed as a function of the *uncertainty* of the robot’s beliefs that its goals will be achieved (or threatened) and of the *utility* assigned to achieving these goals. According to the utility theory, these two parameters are combined to measure *the variation in the intensity of an emotion* as a product of the change in the probability of achieving a given goal, times the utility that achieving this goal brings to the robot.

Let us consider, for instance, the triggering of emotion *sorry-for* in the robot’s mind represented in Fig. 3 (where *R* denotes the robot and *U* the user). This emotion concerns the goal of “preserving the other from bad”. The robot’s belief about the probability that this goal will be threatened ($Bel\ R\ (Thr\text{-}GoodOf\ U)$) is influenced by its belief that some undesirable event *E* occurred to the user ($Bel\ R\ (Occ\ E\ U)$). According to Elliott and Siegle [15], the main variables affecting this probability are the desirability of the event ($Bel\ R\ not(Desirable\ E)$) and the probability that the robot attaches to the occurrence of this event ($Bel\ R\ (Occ\ E\ U)$).

The user’s behavior is interpreted as an *observable* consequence of an occurred event, that activates emotions through a model of the impact of such an event on the robot’s beliefs and goals.

The probability concerning which is the emotion felt by the user ($Feel\ U(emotion)$) is inferred by the affective user modeling component (the probability of the emotion node in the network in Fig. 2); alternatively, the user may explicitly say that he feels

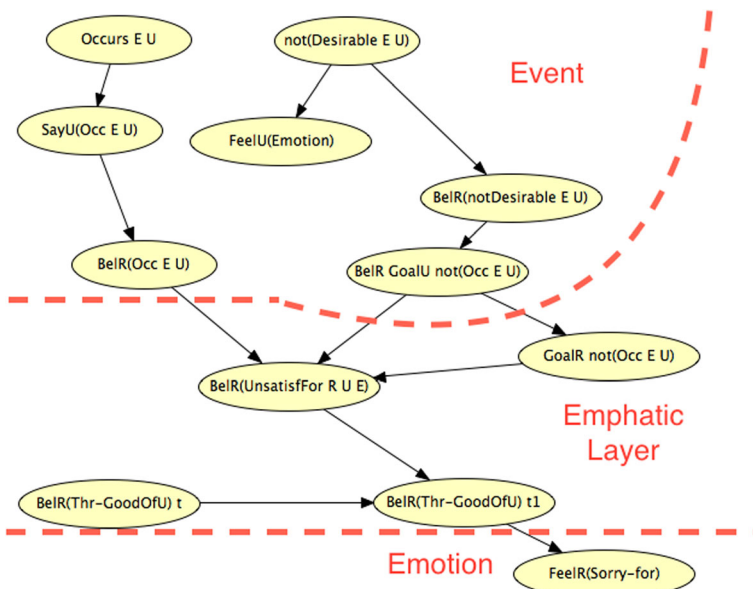


Fig. 3 A portion of the DBN representing the robot’s mental state for the triggering the feeling of Sorry-For

sadness or distress, which denotes that the event is undesirable. This affects R 's belief that U would not desire the event E to occur ($\text{Bel } R \text{ Goal } U \neg(\text{Occ } E \ U)$), and (since R is in an *empathy* relationship with U , R adopts U 's goals) also R 's own desire that E does not occur ($\text{Goal } R \neg(\text{Occ } E)$). This way, the probability values of these nodes concur to increase the probability that the robot's goal of *preserving others from bad* will be threatened. Variation in the probability of this goal activates in R the emotion *sorry-for*, thus simulating the feeling of the emotion in the robot. The intensity of this emotion is the product of this variation times the *weight* the robot associates to the mentioned goal. The strength of the link between the goal-achievement (or threatening) nodes at two contiguous time instants defines the way the emotion, associated with that goal, decays, in absence of any event affecting it.

By varying appropriately this strength, we simulate a faster or slower decay of emotion intensity. Different decays are associated to different emotion categories (positive vs. negative, *FortuneOfOthers* vs. *Wellbeing*, and so on) in which the *persistence* of emotions varies [12].

3.5 Triggering empathic behaviors

Deciding how to behave as a consequence of the triggering of an emotion in the robot's state of mind is a key issue in simulating an empathic behavior. The robot has to trigger an empathic goal and select a plan accordingly. The list of empathic goals of NICA is inspired by the indications that human caregivers provided during the data gathering phase at the beginning of the project, by the literature on empathy and pro-social behavior [13, 32], and by the results of another study on the influence of empathic behaviors on people's perceptions of a social robot [41].

Currently we consider the following empathic goals:

- *console* by making the user feel loved and cuddled;
- *encourage* by providing comments or motivations (e.g., “don't be sad, I know you can make it!”);
- *congratulate* by providing positive feedback on the user's behavior;
- *joke* by doing some humor in order to improve the user's attitude;
- *calm down* by providing comments and suggestions to make the user feel more relaxed.

For instance, in case the *sorry-for* emotion is felt by the robot, the *console* goal should be triggered.

Once a goal has been selected as the most appropriate to the emotion felt by the agent, the behavior planner module computes the robot's behavior using plans represented as context-adapted recipes. Each plan is described by a set of *preconditions*, that must be true to select the plan, the *effect* that the plan achieves, and the *body*, i.e. the conditional actions that make up the plan. After the execution of each action in the plan, the corresponding effect is used to update beliefs in the robot's mental state.

A sample portion of one of the possible plans for achieving the *console* goal is shown in Fig. 4. The tag `<Cond>` allows to select actions on the basis of the current situation. For instance, the action `<Act name="Express" to="U" var="Feel(R,R_emotion)"/>` is used to express the emotion the robot is feeling as a consequence of the emotion recognized in the user (in this case, the *sorry-for* of robot R that will be performed since the user feels sad). In the same way, action “Ask”

Fig. 4 A portion of the Console plan

```

<Plan name="Console">
<SelectCond><Cond R_emotion="sorry-for"/> </SelectCond>
<Body>
<Act name="Move" to="U"/>
<Cond var="Feel(U,Sad)">
    <Act name="Express" to="U" var="Feel(R,R_emotion)"/>
</Cond>
<Cond var="Know_Reason" value="0" >
    <Act name="Ask" to="U" var="Why(U,Feel(U,U_emotion))"/>
    <Act name="Wait" until="Answer(U, Reason)"/>
</Cond>
<Cond var="Know_Reason" value="1">
    <Act name="Inform" to="U" var="Understand(R,U)"/>
</Cond>
<Act name="Express" to="U" var="Console(R,U)"/>
... ..

```

about “Why the user is sad” will be performed only if the robot does not know already why the user is sad. Moreover, if the action is complex, then it can be specified in a sub plan describing elementary robot actions.

Each communicative act in the plan is then rendered using a simple template-based surface generation technique [38]. Templates are selected on the basis of the type of communicative act and its content, and are expressed in a meta-language that is interpreted and executed by the robot’s body. We also simulated emotions by activating specific color combinations for the LEDs in NAO’s eyes².

Plans and surface generation templates have been created and optimized by combining actions on the basis of pragmatic rules that were derived from the dataset.

For instance, as shown in Table 4, the communicative act Express(R,U,Feel(R,sorry-for)) is rendered using the sentence “I’m sorry to see you are sad” and by coloring in blue the robot’s eyes, as suggested by the study conducted by Johnson [20].

4 An example

Let us show an example of empathic behavior of the robot in a typical interaction scenario that we envisaged as a suitable one for testing our robot framework (see Fig. 5):

It’s morning and Nicola, a 73 y.o. man, is at home alone. He feels lonely and sad since it’s a long time since he last saw his grandchildren. Nicola is sitting on the bench in his living room, that is equipped with sensors, effectors and the NICA robot. After a while Nicola starts whispering and says: “Oh My ...oh poor me...”

The communication is perceived and interpreted by the robot, that activates the most appropriate behavior. In this scenario the voice classifier recognizes a *negative* valence with a *low* arousal from the prosody of the spoken utterance. Since the facial expression classifier cannot detect Nicola’s face and expression, due to his posture, this information will not be available to NICA’s emotion monitoring functionality.

² In this work, since the robot had to deal with European elderly people, we used color combinations that are suitable for the European culture. Of course, different cultures may require different colors for conveying the same feelings, but this does not change the basics of our approach (a different color can be easily used for each culture).

Table 4 An example of plan execution in the proposed scenario

| Plan Action | Snapshot of execution |
|-------------------------------------|--|
| MoveTo(N,NICOLA) |  <p>Oh My...oh poor me!</p> |
| Express(N,Sorry-for(N,NICOLA)) |  <p>I am sorry to see you sad</p> |
| Ask(N,NICOLA,Why(NICOLA,Sad)) |  <p>Why are you sad?</p> |
| Express(N,NICOLA,Console(N,NICOLA)) |  <p>Don't worry</p> <p>I can send them a message if you want</p> |

The evidences about the voice valence and arousal are then propagated in the DBN model and the belief about the user being in a negative affective state takes a high probability (0.74), as shown in Fig. 6a. Then, since the robot’s goal of keeping the user in a state of well-being behavior is threatened, the DBNs modeling the robot’s affective mind are executed to trigger the robot’s affective state. In this case, the robot is feeling *sorry-for* (Fig. 6b).

As described in the previous section, according to the social emotion felt by the robot (*sorry-for*), the goal to pursue in this situation is *console*. Then, the corresponding plan is

Fig. 5 A simulation of the scenario



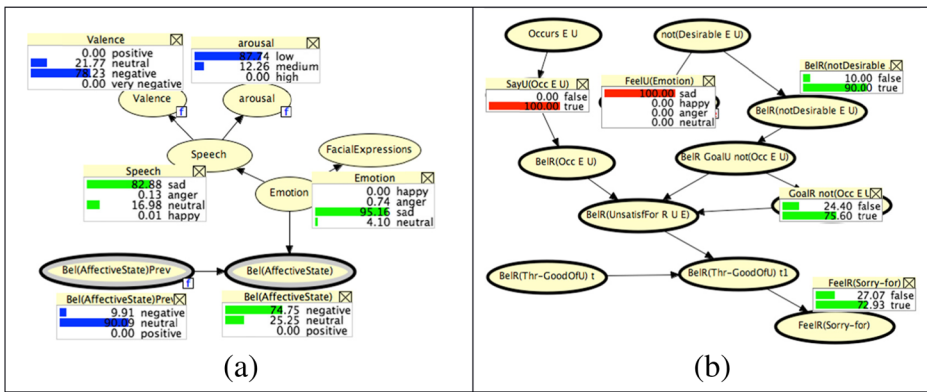


Fig. 6 a and b emotion monitoring and triggering in the described scenario

selected and the execution of its actions begins. The plan and some snapshots of its execution are shown in Table 4.

5 Evaluation

The evaluation was carried out with the main goal of testing whether the affective and cognitive components of empathy were perceived from the robot behavior. Since, at present, statistically significant on-field evaluation was not feasible, due to the lack of enough smart homes equipped with social robots for collecting statistically significant data, the NICA prototype was evaluated both by experts and through user study.

The former approach aimed at evaluating the robot behaviors for each of the empathic goals that NICA is able to fulfill. The latter approach was split into two different studies: one aimed at assessing the perception and interpretation of the empathic behavior of the robot by elderly users, and the other aimed at assessing how empathic the robot appeared in the interaction with a user.

5.1 Expert-based evaluation

Each behavior that NICA could use to interact empathically with elderly users was evaluated by three psychologists, experts in communication. They were asked to look at some videos showing interactions between an actor, playing the role of an elderly man, and the robot for each of the emotional situations that we are able to recognize (sadness, happiness, anger).

For each video they evaluated the behavior plan in terms of communicative acts and verbal or non-verbal signs used for each communicative act. In general, they reported that all plans used for achieving the emotional situation were appropriate. However, they proposed to improve some non-verbal signs for some of the robot communicative acts. For instance, for the communicative act *Express(R,U,sorry-for(R,U))* they suggested to add to the blue eyes also an inclination of the head. In total they proposed improvements on 5 of the 30 communicative acts that the robot is able to perform.

5.2 User-based evaluation

Our original goal in this work was to assess whether the behavior of the robot was recognized as empathic by elderly people representing the potential users of the system. The evaluation consisted of two studies:

- the former aimed at evaluating the perception of the robot's behaviors in terms of the expressed emotion and of the empathic goal the robot was supposed to achieve through that behavior;
- the latter aimed at understanding whether the behavior of the robot was recognized as empathic, both from the cognitive and affective points of view, by looking at the simulation of interactions between an actor (representing an elderly person) and the robot in situations requiring empathic answers.

Both studies were carried out with the same sample of eighteen participants (N=18), in a recreation center for elderly people. All participants were older than 65 and were equally distributed by gender. Their average age was 75.6 with a 10.5 standard deviation. The two studies were carried out in two different days with one month of distance between them.

5.2.1 First study

Setup A facilitator, which was a psychotherapist expert in elderly problems, explained to each subject the purpose of the experiment. Each of them was prompted to sit down on a chair in front of a table and to evaluate the behavior of NICA according to the two dimensions explained previously. The five empathic behaviors were played by NICA, and after the execution of each behavior the subject was asked to fill in a survey. At the end of the study subjects were fully debriefed by the facilitator.

Measures Following the interaction, each participant filled a survey regarding “**what the robot was expressing**” and “**which was the robot's goal**” by answering simple questions with marks on a scale from 1 (I disagree completely) to 5 (I agree completely). Table 5 reports the questions that the facilitator asked each subject after each behavior execution (#Behavior).

Results After performing the experiment, data was collected from the completed survey forms. Tables 6 and 7 report the experimental results. The columns indicate the behavior executed by the robot and the rows the possible answers. Each cell shows the average score for each behavior. In bold we highlighted values above the average.

We can conclude that in most cases the robot was able to convey the intended empathic goal, since it was clearly recognized by the subject.

As regards the specific empathic expression conveyed, it was not always clearly recognized. In most cases, subjects assigned the same score to two affective expressions. However, for the majority of behaviors, the recognized emotions have the same valence as the intended ones. For instance, the *console* behavior was conveying both *sadness* and being *sorry-for someone*.

These results may depend also by the lack of facial expressions in the NAO robot, since the global empathic goal was clearly recognized and only the affective expressions were sometimes mismatched.

Table 5 An example of the evaluation questionnaire for behavior #1

| #Behavior | In your opinion what was the robot expressing? | Value 1 to 5 (1=I disagree completely, 5=I agree completely) | In your opinion what do you think was the robot goal? | Value 1 to 5 (1=I disagree completely, 5=I agree completely) |
|-----------|--|--|---|--|
| #1 | Sorry-for someone | | Console someone | |
| | Happy-for someone | | Encourage someone | |
| | Encouragement | | Calm-down someone | |
| | Happyness | | Congratulate someone | |
| | Sadness | | Make someone happy | |
| | Relax | | | |
| | Congratulation | | Joke | |
| | I do not know | | I do not know | |
| | Other: | | Other: | |

5.2.2 Second study

Setup The second study was carried out with the same subjects as the previous one (details on the subjects are in the paragraph above), but at a different time (about one month later). The facilitator explained to each subject the purpose of the second study that is to evaluate whether the behavior of the robot was empathic towards the user by looking at the simulation of interactions between an actor and the robot in situations requiring empathic answers.

Each subject was prompted to sit down on a chair in front of a 27-inch display and to look at videos representing the interaction between an elderly user (a professional actor) and the robot for the five empathic goals implemented in NICA (one example is presented in Table 4). After each video the subject was asked to fill in a survey. At the end of the study subjects were fully debriefed.

Measures The survey included simple statements to be evaluated in terms of agreement or disagreement on a 5-points Likert scale. The statements were intended to test how empathic the robot appeared in the interaction with the user, and whether the robot expressed:

Table 6 Results of the first question in a scale from 1 to 5

| What was the robot expressing? | #1 Calmdown | #2 Congratulate | #3 Console | #4 Encourage | #5 Joke |
|--------------------------------|-------------|-----------------|------------|--------------|---------|
| Sorry-for someone | 3.17 | 1.00 | 4.61 | 1.00 | 1.00 |
| Happy-for someone | 1.00 | 4.56 | 1.00 | 3.67 | 1.00 |
| Encouragement | 1.00 | 2.44 | 1.28 | 4.33 | 1.00 |
| Happiness | 1.00 | 4.39 | 1.00 | 3.94 | 3.78 |
| Sadness | 1.89 | 1.00 | 4.61 | 1.00 | 1.00 |
| Relax | 3.22 | 2.06 | 1.00 | 1.00 | 2.33 |
| Congratulation | 1.00 | 5.00 | 1.00 | 2.72 | 1.00 |
| I do not know | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Other | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

Table 7 Results of the second question in a scale from 1 to 5

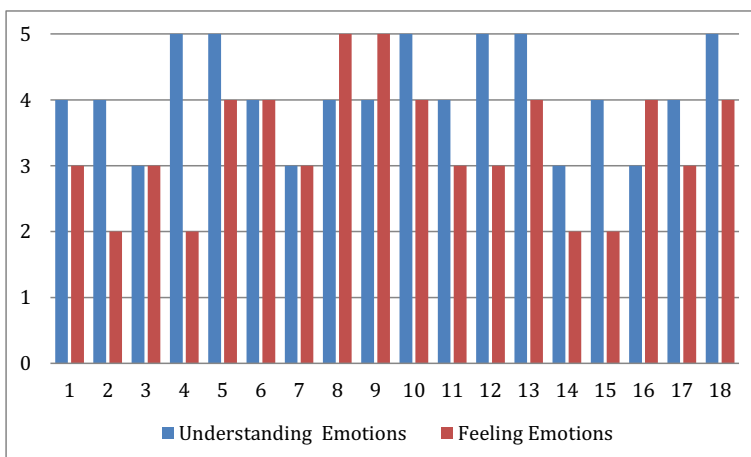
| Which was the robot's goal? | #1 Calm down | #2 Congratulate | #3 Console | #4 Encourage | #5 Joke |
|-----------------------------|--------------|-----------------|------------|--------------|---------|
| Console someone | 2.4 | 1 | 4.6 | 1 | 1 |
| Encourage someone | 1.5 | 2.9 | 2.1 | 4.5 | 1 |
| Calm-down someone | 5 | 1 | 3.2 | 1 | 1 |
| Congratulate someone | 1 | 4.6 | 1 | 3.3 | 1 |
| Make someone happy | 3 | 2.7 | 2 | 2.2 | 2.4 |
| Joke | 1 | 1 | 1 | 1 | 4.6 |
| I do not know | 1.2 | 1 | 1 | 1 | 1 |
| Other: | 1 | 1 | 1 | 1 | 1 |

- i) its feeling as an emotional reaction to the user's affective state (affective empathy) and
- ii) some understanding of the user's emotion (cognitive empathy).

In the survey two simple statements were used: “The robot can feel emotions.”, and “The robot can understand the emotion of the elderly person.”.

Results After performing the study, data was collected from the filled survey forms. Results are illustrated in Fig. 7. The results of the second study did not reveal differences due both to gender and age in participants. All the subjects perceived as empathic the behavior of NICA during the interaction with the actor.

The recognized capability of NICA of understanding emotions was higher, on average, than its ability to feel emotions (4.11 vs 3.35). This may be due to the fact that, in all the videos, NICA was using the verbal communication channel to express its understanding of the user's feeling, while for expressing its own feeling NICA was using non-verbal communicative signs (i.e., color of the eyes, postures, gestures).

**Fig. 7** Results of the evaluation of the perception of the cognitive and affective dimensions of empathy

6 Conclusions and future work

This paper motivated the importance of taking into account affective factors when modeling an elderly user in social interaction with a caring robot. The idea underlying our work is that endowing the robot with a social empathic behavior is fundamental in everyday life environments in which, besides assisting an elderly user in performing his tasks, the robot has to establish a social long-term relationship with him so as to enforce his trust and confidence. So, we illustrated how this capability has been designed and implemented in a caring assistant for elderly people. We also pointed out the complexity associated with the implementation of an empathic attitude in NICA, a general architecture for social assistive robots that has been embodied in the NAO robot. In particular, the robot is able to recognize the affective state of the user by analyzing communicative signals extracted from speech, facial expressions and gestures, and to trigger its own affective state accordingly. We used a combination of several classifiers, used to detect the features that come into play when deciding the affective state to be triggered, with Dynamic Belief Networks, used to reason on them and take a decision.

Since, at the moment, there are not enough smart homes equipped with social robots for collecting statistically significant data, we assessed whether the behaviors of the robot were appropriate for conveying the intended empathic response by performing an expert-based evaluation and a user study.

Results were quite satisfactory since in both experiments empathic behavior was perceived correctly. The experts evaluated positively the behavior of the robot for each of the defined empathic goal. They just suggested minor changes (adding some non-verbal communicative acts, or changing some words in the spoken sentences). As regards the user study we can conclude that empathy was clearly recognized, while its affective component was not always perceived as expected. In part, this may be due to the absence of facial expressions in the NAO robot. However, even in cases in which the robot expression was mismatched for another one, the recognized affective component had the same valence as the expected one. These results are a basis for studying the user response in AAL environments when the robot has to provide continuous support and establish a long term relation with the user.

An important issue to be addressed in our future work concerns the implementation of strategies for avoiding repetitions of the same agent's behavior in similar situations. Endowing the robot with a social memory and adopting a more emotion-oriented BDI architecture [19] might improve this aspect of the project.

Acknowledgments This work fulfills the research objectives of the PON02_00563_3489339 project "PUGLIA@SERVICE - funded by the Italian Ministry of University and Research (MIUR). Our thanks go to Isabella Poggi and Francesca D'Errico for their useful suggestions on the evaluation of the robot's behaviors.

References

1. Anderson C, Keltner D (2002) The role of empathy in the formation and maintenance of social bonds. *Behav Brain Sci* 25:21–22
2. Baron-Cohen S (2011) *The science of evil: on empathy and the origins of cruelty*. Basic Books, New York, NY
3. Beadle JN, Sheehan AH, Dahlben B, Gutchess AH (2013) Aging, empathy, and prosociality. *J Gerontol Ser B Psychol Sci Soc Sci*:1–10

4. Boersma P, Weenink D (2007) Praat: doing phonetics by computer (version 4.5.15) [computer program]. <http://www.praat.org/>. Retrieved 24.02.2007
5. Broekens J, Heerink M, Rosendal H (2009) Assistive social robots in elderly care: a review. *Gerontechnology* 8(2):94–103. doi:10.4017/gt.2009.08.02.002.00
6. Cesta A, Cortellessa G, Pecora F, Rasconi R (2007) Supporting interaction in the robocare intelligent assistive environment, AAAI Spring Symposium: Interaction Challenges for Intelligent Assistants, pp. 18–25
7. Coradeschi S, Cesta A, Cortellessa G, Coraci L, Galindo C, Gonzalez J, Karlsson L, Forsberg A, Frennert S, Furfari F, Loutfi A, Orlandini A, Palumbo F, Pecora F, von Rump S, Stimec A, Ullberg J, Östlund B (2014) GiraffPlus: a system for monitoring activities and physiological parameters and promoting social interaction for elderly in human-computer systems interaction: backgrounds and applications 3. *Adv Intell Syst Comput* 300:261–271. doi:10.1007/978-3-319-08491-6_22
8. Cramer H, Goddijn J, Wielinga B, Evers V (2010) Effects of (In) accurate empathy and situational valence on attitudes towards robots. *ACM/IEEE International Conference on Human Robot Interaction (HRI)*. 141-2. New York. March 2005
9. Davis M (1980) A multidimensional approach to individual differences in empathy. *Cat Sel Doc Psychol* 10: 85
10. De Carolis B, Ferilli S, Greco D (2013) Towards a caring home for assisted living. In: Proceedings of the workshop “the challenge of ageing society: technological roles and opportunities for artificial intelligence” co-located with the 13th Conference of the Italian Association for Artificial Intelligence (AI*IA 2013), Turin, 6 Dec 2013. <http://ceur-ws.org/Vol-1122/>
11. De Carolis B, Mazzotta I, Novielli N, Pizzutilo S (2013) User modeling in social interaction with a Caring Agent. In: *UMADR - User modeling and adaptation for daily routines*, Springer HCI Series. Springer
12. de Rosi F, De Carolis B, Carofiglio V, Pizzutilo S (2003) Shallow and inner forms of emotional intelligence in advisory dialog simulation. In: Prendinger H, Ishizuka M (eds) “Life-like characters. Tools, affective functions and applications. pp 271-294. Springer
13. Eisenberg N, Miller PA (1987) The relation of empathy to prosocial and related behaviors. *Psychol Bull* 101: 91–119
14. Ekman P, Oster H (1979) Facial expressions of emotion. *Annu Rev Psychol* 30:527–554
15. Elliott C, Siegle G (1993) Variables influencing the intensity of simulated affective states. In Proceedings of the AAAI Spring Symposium on Mental States. '93:58–67
16. Graf B, Hans M, Shaft RD (2004) Care-O-bot II – development of a next generation robotics home assistant. *Auton Robots* 16(2):193–205
17. Hoffman ML (1981) The development of empathy. In: Rushton J, Sorrentino R (eds) *Altruism and helping behavior: Social personality and developmental perspectives*. Erlbaum, Hillsdale, pp. 41–63
18. Jensen FV (2001) *Bayesian networks and decision graphs, statistics for engineering and information science*. Springer New York, New York, doi:10.1007/978-1-4757-3502-4
19. Jiang H, Vidal JM, Huhns MN (2007) EBDI: an architecture for emotional agents, Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems, May 14-18, Honolulu, Hawaii, ACM, p. 11
20. Johnson DO, Cuijpers RH, van der Pol D (2013) Imitating human emotions with artificial facial expressions. *Int J Soc Robot* 5(4):503–513
21. Klein J, Moon Y, Picard R (1999) This computer responds to user frustration. In Proceedings of the Conference on Human Factors in Computing Systems. Pittsburgh, USA, ACM Press, pp 242–243
22. Leite I, Martinho C, Paiva A (2013) Social robots for long-term interaction: a survey. *Int J Soc Robot*. doi:10.1007/s12369-013-0178-y
23. Mataric M, Tapus A, Feil-Seifer D (2007) *Personalized socially assistive robotics, workshop on intelligent systems for assisted cognition*, rochester, New York, USA
24. Mehrabian A (1968) Communication without words. *Psychol Today* 2:53–56
25. Milborrow S, Nicolls F (2008) Locating facial features with an extended active shape model. In *Computer Vision-ECCV 2008* (pp. 504-513). Springer Berlin Heidelberg
26. Milborrow S, Nicolls F (2014) Active shape models with SIFT descriptors and MARS, In Proceedings of the 9th International Conference on Computer Vision Theory and Applications, Lisbon, pp. 5–8

27. Nehmer J, Becker M, Karshmer A, Lamm R (2006) Living assistance systems: an ambient intelligence approach, in International Conference on Software Engineering, ACM, pp. 43–50
28. Niewiadomski R, Ochs M, Pelachaud C (2008) Expressions of empathy in ECAs, in Proceedings of the 8th Int.Conf. on IVA, LNAI, vol. 5208. Springer-Verlag, Berlin, Heidelberg, pp 37–44
29. Nijholt A, de Ruyter B, Heylen D, Privender S (2006) Social interfaces for ambient intelligence environments. Chapter 14. In: Aarts E, Encarnação J (eds) True visions: the emergence of ambient intelligence. Springer, New York, pp. 275–289
30. Oatley K, Johnson-Laird PN (1987) Towards a cognitive theory of emotions. *Cognit Emot* 29–50
31. Ortony A, Clore GL, Collins A (1988) The cognitive structure of emotions. Cambridge University Press
32. Paiva A, Dias J, Sobral D, Aylett R, Sobrepepe P, Woods S, Hall L (2004) Caring for agents and agents that care: building empathic relations with synthetic agents. In Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1 (pp. 194–201). IEEE Computer Society., ACM Press
33. Paiva A et al. (2011) Empathy in social agents. *Int J sVirtual Real* 10(1):65–68
34. Palestra G, Pettinicchio A, Del Coco M, Carcagni P, Leo M, Distanto C (2015) Improved performance in facial expression recognition using 32 geometric features. *ICIAP, lecture notes in computer science*, Springer, pp. 518–528
35. Picard R (1997) *Affective computing*. MIT Press, Cambridge, MA
36. Pineau J, Montemerlo M, Pollack M, Roy N, Thrun S (2003) Towards robotic assistants in nursing homes: challenges and results, *robotics and autonomous systems* 42(3–4):271–281
37. Reeves B, Nass C (1996) *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press, New York, NY
38. Reiter E, Dale R (2000) *Building natural language generation systems*. Studies in natural language processing. Cambridge University Press, Cambridge, United Kingdom. ISBN 0-521-62036-8
39. Russell J (1980) A circumplex model of affect. *J Pers Soc Psychol* 39:1161–1178
40. Speigel C, Greczek J, Mataric M (2016) Creating a baseline for empathy in socially assistive robotics. <http://robotics.usc.edu/~cspiegel/naoempathy4.pdf>
41. van Ruiten AM, Haitas D, Bingley P, Hoonhout HCM, Meerbeek BW, Terken JMB (2007) Attitude of elderly towards a robotic game-and-train- buddy: evaluation of empathy and objective control. In: Cowie R, de Rosi F (eds) Proceedings of the Doctoral consortium, in the scope of ACII2007 Conference
42. Viola PA, Jones MJ (2004) Robust real-time face detection. *Int J Comput Vis* 57(2):137–154
43. Vogt T, Andrié E, Bee N (2008) EmoVoice - a framework for online recognition of emotions from voice. In Proceedings of the 4th IEEE tutorial and research workshop on Perception and Interactive Technologies for Speech-Based Systems: Perception in Multimodal Dialogue Systems (PIT '08), Springer-Verlag, Berlin, 188–199



Berardina De Carolis is Assistant Professor at the Department of Computer Science at the University of Bari, Italy. Her research interests are mainly in User Modeling and Intelligent Interfaces.



Stefano Ferilli is Associate Professor at the Department of Computer Science at the University of Bari, Italy. His research interests are mainly in Machine Learning and Intelligent Systems.



Giuseppe Palestra is a Ph.D student at the Department of Computer Science at the University of Bari, Italy. His research interests are mainly in Social Assistive Robotics and Computational Vision.