IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Conflict and Reconciliation Processes between Affective/Social Robots and Humans

GUILLERMO ALVAREZ-PARDO[1,2] and ERNESTO FABREGAS[1].

[1]Departamento de Informática y Automática, Universidad Nacional de Educación a Distancia (UNED), Juan del Rosal 16, 28040 Madrid, España (e-mails: galvarez146@alumno.uned.es, Escuela Internacional de Doctorado; efabregas@dia.uned.es)

[2]Mathematics Department, Cuesta College, Highway 1, San Luis Obispo, 93495 California, USA (e-mail: guillermo_alvarezpar@cuesta.edu)

Corresponding author: Guillermo Alvarez-Pardo (galvarez146@alumno.uned.es, guillermo_alvarezpar@cuesta.edu).

ABSTRACT Most research on affective computing relates to recognizing and classifying emotions, usually through facial or body expressions, linguistics, electroencephalograms or other biosignals. A variety of authors have pointed out that for social and affective robots to establish effective, deep and durable bonds with humans, they must emulate human interactions as closely as possible; however, there are aspects of human behavior and interactions, like disputes and resolutions, that have been left aside from the design of such robots. This article introduces a non-intrusive, low-cost system that allows robots to recognize and *simulate* affections and personality on the basis of human-robot *actions*, while also allowing robots to recognize and shape the human's character and the nature of their relationship. It provides a system for robots to trigger and carry out *conflict and reconciliation processes* with humans.

INDEX TERMS Affective computing, social affective robots, automatic learning, artificial intelligence, conflict resolution

I. INTRODUCTION

AFFECTIVE Computing was originally defined as computing that relates to, arises from, or influences emotions [1]. Nowadays, it can be characterized as the study and development of systems and devices that can *recognize*, *interpret*, *process*, and *simulate* human affects [2]. In the current literature, abundant examples can be found for the first three of those capabilities, while simulation has received much less attention. As a simple quantitative analysis, Google Scholar throws 157 entries for "recognition of affects" and amounts also in the hundreds for each of the two subsequent terms, but only 26 entries for "simulation of affects". This article attempts to fill part of the void regarding simulation present in the current state of the art.

The system proposed here allows for the following:

1) Simulation of emotions on the robot's side, generated as a consequence of attention/neglect or met/broken expectations and accompanied by the robot's ability to launch and carry out conflict and reconciliation processes (CRP) with the human.
2) Characterization and classification of personalities and relationships, including the robot's ability to recognize the type of relationship the human is establishing and develop expectations, all of this based on actions.
3) Simulation of a personality in the robot.
4) Robot's ability to influence or shape the human's personality or behavior.

The rationale for endowing robots with CRP capabilities is that human-robot relationships (HRR) could be more profound if they include disputes and resolutions the way human-human relationships do. The remainder of this paper is structured as follows:

Section II discusses literature on affective/social computing/robots and conflict relevant to CRP. Section III comments on the steps and considerations followed during the development of the work presented in this paper. Section IV introduces a theory on CRP and Section V presents a model for CRP that adheres to that theory. Section VI includes additional considerations about the model. Section VII discusses the testing and simulations performed with a human for short-term analysis and with a model simulating a human for long-term analysis. Section VIII discusses the main contributions of this paper, its limitations, ethical concerns, possible applications, and directions for future research.

II. LITERATURE REVIEW

This section discusses concepts regarding affective computing and conflict and reconciliation processes that have similarities or connections with the work presented in the following sections.

A. AI AND AFFECTIVE COMPUTING. EXAMPLES, APPLICATIONS AND LIMITATIONS

In recent decades, affective and social robots have become popular for pedagogical, medical, and recreational purposes. [3] reviews a decade (January 2013 - May 2022) of literature on human-robot interactions for wellbeing. It identifies three challenges of affective robotics: first, understanding the fundamental mechanisms of human behavior; second, developing systems for robots to dynamically adapt to human behavior, meeting the needs of each individual and personalising their behavior accordingly; third, transitioning from affective computing to a robot in a real-world context. The review also points out that a common problem with the current state of the art is that most robots are not fully autonomous and "researchers usually program human-robot interactions as a one off experience, for a limited scope and very short interaction durations (usually no longer than 20 minutes)."

[4] focuses on continual learning (CL). It defends the importance of building systems with long-run memory, able to remember past interactions and personalise towards each user while also influencing the learning of novel expressions. On the other hand, it warns that a system with long-run memory "might require a lot of interactions before the model can successfully adapt, negatively impacting the initial user experience"; the authors point out that this could be ameliorated with *adversarial training*. It regrets that, in the existing approaches, few models focus on learning task-oriented behaviors.

[5] and [6] are two recent studies that show how affective robots make a difference on education and engagement. [6] involved 16 children with cochlear implants or hearing aids. Their performance was evaluated in a conventional setup, a tablet setup and a tablet+robot setup, with the robot being humanoid and with emotion recognition capabilities. Whereas the objective test metrics taken in the three scenarios did not show statistically significant differences, the test duration, engagement and attention to the test significantly increased with the tablet+robot setup compared to the others. [5] also shows how affective robots are effective at teaching children something as complex as coding. While this study did not count on various control groups, it did show how students improved their performance after working with the robot and reported high satisfaction and an overwhelming desire to take more lessons of this type.

[7] points out that common affective measures based on physiological and psychological responses usually require intrusive and expensive tools that are impractical in real settings. In response, the authors propose an emotion recognition system based on typing dynamics and mouse interactions.

[8] introduces a system for a robot to simulate types of personality characteristics, including *brave*, *steady*, *sincere*, *kind-hearted*, *self-confident*, *tenacity*, *forward-looking*, and *optimistic*. It relies heavily on a pre-existing tool called AIWAC smart box.

[9], [10], [11], [12] and [13] are patents from the last seven years that contain detailed descriptions and have connections with the work presented in this paper. [9], granted to Huawei in 2023, provides a method intended to allow a robot to simulate "sensibility" and "emotional needs in a manner similar to that of human beings, thereby gradually building trust and dependency." Huawei's method is similar to others in that it uses signs coming from voice, expression, body, skin, etc., while it is different from others in that it is able not only to determine the current emotional status, but also predict a future one. [10], granted to The George Washington University in 2020, uses "emotional dimensions include at least activation, valence, and dominance, and the at least three levels of emotional dimensions include a high state, a neutral state, and a low state." [13], granted to Microsoft in 2017, classifies emotion types based on dialog.

B. CONFLICT, PERSONALITY, AND TYPE OF RELATIONSHIP

A variety of authors have shown that the frequency and nature of conflict in personal relationships are significantly determined by the individual's character, the type of relationship with the partners and the type of environment. The types of relationship, closely tied to the types of environment, can be summarized in four categories: between parents and children, friends or pals, colleagues or peers, and romantic couples or spouses.

[14] explains how the financial character of each individual determines the origin of conflict in couple relationships. [15] shows how sex is tied to how often, with whom and in what form individuals tend to fall out and reconcile. [16] has shown that attachment styles among parents influence the type of conflicts that are more common in the child-parent and child-child relationships. [17] and [18] have shown how the individualist or collectivist nature of the individual or the environment determines the inclination to one or other conflict styles, especially in the workplace.

On the other hand, different models have been developed for personality analysis. Two of the most popular ones are the HEXACO [19] and the OCEAN [20] models.

C. FUNCTIONAL CONFLICT AND INTEGRATIVE CONFLICT

In a relationship, conflict can emerge naturally from misadjustments between the parts. Conflict works then as a bell that brings attention to a particular problem for the sole purpose of having it fixed, and vanishes when appropriate adjustments are made. This can be referred to as "functional conflict", [21], [22], [23]. These "misadjustments" are sometimes referred to as different forms of "violations". [24] defends that in the context of a relationship, conflicts arise from violations that can be classified into two categories: those that place an obstacle towards one's needs and those that break one's expectations.

Various studies have also focused on how conflict is fundamental, and not just accidental, in many relationships, how the generation and handling of conflict is in itself an emotional need for some individuals and how the process conflict-fight-reconciliation might be part of a ritual that is required in many relationships. This has to do with what [25] called "integrative functions of conflict". The following list shows some possible integrative functions of conflict.

1) Conflict as an escape from boredom in search of excitement

[26] explains that boredom has not received the attention it deserves from moralists and philosophers. People get into trouble, search for conflict or put themselves at risk to escape boredom and find excitement. This attitude might be caused by the fact that, for millions of years, humans got used to living with some amount of fighting for their lives, and, therefore, an excess of comfort goes against our nature. [27] documented how people's need for excitement impels them to practice a variety of violent activities, such as hunting or certain sports.

In the context of a relationship, a CRP may act as a simple valve to escape the boredom of the routine and find some excitement. This functionality could take place at an individual level, when a particular member of the relationship suffers the effects of boredom and seeks evasion, or at a collective level, when the relationship, as a complex organism, wants to bring excitement.

2) Conflict as a way of attracting attention and asserting one's self

It is well-known that children are able to cry or break things to attract the attention of their parents or caretakers. [28] points out that "The reason for the empirical lack of resolution in children's disputes is that a basic function of conflict is to achieve a concrete, particular social organization through the display of opposition and the constitution of accountable alignment structures", "contradicting a central tenet in functional theory that disputes aim toward resolution". This inclination, which is softened through education, does not always completely disappear in adults. Different authors have noticed that some individuals seem to obtain a sense of their own value and self-esteem by antagonizing others [29]. For them, quarreling and triggering CRPs are egotistic needs.

3) Conflict as a way of realizing the values and norms of the environment and one's place in it

[28] and [30] provide a sufficient illustration of how children and young adults use conflict to realize the norms of their community and their class and gender identities. "First, it helped the participants learn how to use the physical space of the park, and second it helped them learn about ideological norms emphasized by the community of practice" [30]. Similar usages of conflict have been documented among K-12 students [31].

4) Conflict as a bond with reality against separateness

[32] sustains that separateness is the root of all anxiety. This is because we, humans, are not part of nature the way other animals and plants are. Thus we seek rituals and practices that help us diminish or cope with that separateness. The author explains that some of those ways of handling separateness are more detrimental than beneficial and holds romantic love to be the most healthy and effective way. [33] explains that sentiments are tools of the subject as well as states of the subject and states of the organism. As tools, they are mostly used to tie the subject to objects around them, especially those convenient for their survival. These objects could be persons, animals, things, or parts of their inner life, such as thoughts, memories or other sentiments. They are also states of the subject because they modify it and bring symptoms and reactions that could turn into a syndrome that goes beyond the limits of the subject into what he calls the organism. The latter implies a more severe transformation and may occur for better or for worse. In the context of these ideas, conflict could very well be another way of bringing the sentiments that tie the individual to objects around them as a way of fighting separateness. [25] also pointed out that conflict can strengthen pre-existing bonds or create new ones.

5) Conflict as a test to the relationship and to the veracity of the sentiments

[33] also explains that the intimate character of emotional experiences, together with the non-specificity of their expression, carries as a major consequence that the subject must accept, willingly or by force, some uncertainty about the sentiments that are being expressed. This makes people prone to bringing or demanding proof of authenticity. The generation of conflict might work as a means for some subjects in their search for these proofs. If the subject who claims to love us spends time and resources attending to the conflict we have generated or exploited, this might be a tangible proof of authenticity. Also, if the relationship survives, it would prove the solidity of the bond and relationship. It is worth pointing out that these proofs would not come just from the conflict but from the entire CRP.

This functionality might also be tied to the idea that "What does not kill us makes us stronger". According to this, some CRPs not only prove the solidity of the relationship but also strengthen it.

6) Conflict as a way of bringing reason, consistency and creativity

In the context of Artificial Intelligence and education, some studies have suggested that individuals in sad moods tend to process information in a more systematic way, being more thorough in separating strong arguments from weak ones and using a more rational analysis of the situation; in contrast, a happy mood is associated with more inconsistent decision-making and less causal reasoning, [34], [35]. After a period of happiness and lightness, conflict might very well act as an attractor of that sad mood or seriousness that facilitates consistency and a more rational decision-making. [36] also suggests that conflict can bring a level of reasoning able to produce better decision-making than the usual rationality limited to "rule following". Einstein famously said: "Creativity comes from anxiety as the day comes from the dark night", "Without crisis, there are no challenges; without challenges, life is a routine, a slow agony."

III. VISION AND METHODOLOGY

The following steps led to the creation of the computational model presented in the following sections.

1) *Understanding CRP in humans*. This was done through an extensive study of the literature about it in scientific branches like psychology, psychiatry, social sciences, *et cetera*. Part of this study has been included in Section II-C. The study suggested that humans could indeed benefit from having CRP capabilities in some of the robots they interact with.
2) *Studying the state of the art in affective and social robots, especially with regard to CRP*. No example of a robot or computational model designed for CRP was found in the literature (journals, conferences, patents, etc.). This suggested that a first approach to that concept would be an effort worth doing.
3) *Modeling the appearance of conflict and its resolution*. It would be necessary to work with a system of metrics and thresholds that allows the robot to determine when to launch a CRP and when to consider it addressed and finished. This would require an understanding of the possible metrics and thresholds that regulate when humans do the same. The literature showed that a major reason for conflict in human relationships was *broken expectations*. This key reason would be incorporated into the computational model, so the robot would have some metrics representing expectations and some others representing the perceived human behavior and would trigger a CRP when those two are not in sufficient harmony.
4) *Modeling expectations, personality and the dimensions of a relationship*. Expectations depend on the type of relationship and the literature has studied four major types of relationship: among parents and children, friends or pals, colleagues or professional partners, and romantic partners or marriages. Therefore, a model that considers these dimensions and uses them to characterize a personality or relationship would be convenient. This was the inception of the Pat Palprolov model.
5) *A model tying human actions with personality or relationship type*. This was important because the ability to simulate emotions based on actions is one of the key differentiators of the model introduced here. This was achieved using the concept of *loom and fabric of actions*.
6) *Testing and simulations*. Once the computational model was completed, it was instrumented using Python. The model was then tested with a variety of human behaviors in the short run and with another model simulating a variety of humans in the long run. The key results of this stage are presented in Section VII.

IV. A BRIEF THEORY ABOUT CONFLICT AND RECONCILIATION PROCESSES

A. PAT PALPROLOV. A PERSONALITY AND RELATIONSHIP MODEL

Attending to the ideas presented in II-B, a novel structure is introduced here. It is based on four categories that are considered to be *dimensions* of a relationship and *dimensions* of a personality. These dimensions are here named *paternalistic*, *pals*, *professional* and *love*, shortened with the acronym Pat Palprolov or PPPL. The following are brief descriptions of the nature of each:

1) *Paternalistic Dimension*. This has to do with treating the other as someone whose education, security, well-being or happiness one is responsible for. It includes actions such as supporting, patronizing or expressing love. It is the only non-symmetric dimension, in the sense that one part plays the role of parent, while the other plays the role of child, which are very different from each other.
2) *Pals Dimension*. This relates to the usual interactions between friends. [37] provided a cross-cultural characterization of friendship which involves six significant rules: stand up for a friend in their absence, share important news, provide emotional support, trust and confide the other, volunteer help and make the other person happy.
3) *Professional Dimension*. This is tied to the specificities of the relationships between peers, coworkers or colleagues. More formal and distant interactions take place in this context. Expressions of affection are not expected. Shared activities usually involve problem-solving or working to accomplish a particular goal.
4) *Love Dimension*. This emphasizes the expression of affection. It is the farthest from the professional dimension and somewhere between the pals dimension and the paternalistic dimension.

They are considered to be *dimensions* of the personality and *dimensions* of a relationship because everybody has the ability to lean towards one or another dimension and combine several of them, and every relationship participates in each of them to some extent. For example, the relationship between two peers at their workplace will participate mostly in the professional dimension, with its characteristic formality and goal-oriented interactions; however, casual conversations or interactions, as if they were pals, will be likely to take place too; one part may very well feel inclined to protect the other in a variety of scenarios, which would relate to the paternalistic aspect; and expressions of personal appreciation or affection will not be totally discarded, which leaves room for the love dimension. Similarly, one could have a personal inclination to interact with people in a fashion more sided with the paternalistic, pals, professional or love dimensions. Clearly, these four dimensions are not mutually exclusive; within the limits of this relationship/personality model, they are exhaustive.

B. STRUCTURE OF A CRP

Different structures have been proposed for the analysis of a CRP. The Management Study Guide [38] considers five phases: Prelude, Triggering Event, Initiation, Differentiation and Resolution. To serve the system introduced in this article, the authors introduce the distinction between the following elements or moments in time. The upcoming definitions include references to the mathematical model that will be fully explained in Section V.

1) *Origin*. The origin is the element that one part perceives as a reason to initiate a CRP or the element that impels one part to initiate a CRP. The origin is often a combination of factors or influences. The origin will be tied to elements of the CRP system that will be called *strength function*, *dimensional weights*, and their respective *thresholds*.
2) *Trigger*, *actor* and *reactor*. The trigger is the action that one part takes to initiate a CRP. In human-human relationships, such action is sometimes taken in a more unconscious way, that is, without being part of a conscious CRP strategy. Such a distinction will not apply when a robot launches a trigger, as the concept of "unconsciousness" in robots is still a matter of debate, [39], [40], and a discussion about it falls outside the scope of this article. The actor is the part that launches the trigger and the reactor is the part that reacts to it.
   The system introduced in this article is limited to the case in which the robot is the actor and the human is the reactor. In other words, the system focuses on making the robot detect origins and launch appropriate triggers. These triggers are classified into three categories, *reproach*, *omission* and *action*, according to their intensity, and such intensity is determined with the help of the *stress* function and the respective *thresholds*.
3) *Realization*. This is the moment when the reactor realizes that the actor has initiated a CRP. For practical reasons (as the system works from the perspective of the robot), it can also be defined as the moment at which the actor realizes that the reactor has realized that a CRP has been initiated.
   In the current system, the robot will assume realization unless time passes and no significant reaction comes from the human. If the latter occurs, the robot will accumulate *stress* and may escalate to a more intense trigger.
4) *Negotiation*. The negotiation is the exchange of ideas or actions that takes place, usually initiated by the reactor, as part of an attempt to solve the conflict.
   In the present system, the negotiation is tied to the triggers that the robot is able to launch and the different sets of reactions that the robot can detect and recognize from the human. In particular, the negotiation starts when the robot launches its first trigger and consists of an exchange in which the human provides actions and the robot replies with different intensities for the same trigger, with different triggers or with a reconciliation signal.
5) *Reconciliation*. This takes place when both parts are satisfied, which puts an end to the CRP. The actor is usually the part that determines, with explicit agreement, the moment at which reconciliation occurs.
   In the system introduced in this article, reconciliation occurs when the strength function or dimensional weights return to the correct levels as determined by the thresholds (they had to be below those levels for the robot to launch a trigger in the first place). The robot explicitly expresses that reconciliation occurs by performing the reconciliation signal.

V. A MATHEMATICAL MODEL FOR CRP

A. THE *LOOM AND FABRIC OF ACTIONS*

The system works with a finite set of actions that the robot is able to recognize from the human. While the size and composition of that set can be changed, here are some guidelines for designing the set of actions and a particular set of actions as an example. These guidelines adhere to the concept of *loom and fabric of actions*, which refers to the structure in which the four dimensions play the role of a warp beam through which the actions are woven.

1) Some guidelines for the loom and fabric of actions

1) The actions should be as easy to perform, in such a way that only the human's inclinations determine what actions the human performs more often.
2) The actions should be tied to the four dimensions of the relationship in an even way, so that if the actions were performed randomly, the four dimensions would tend to have equal values. This also guarantees that the four dimensions reflect the human's preferences.
3) Different actions should participate in different dimensions to different extents. This allows the dimensions to evolve in one or another direction depending on how the human behaves.
4) $Exclusive$, $dual$, $triple$ and $general$ actions. A simple way of weaving all actions differently but evenly throughout the four dimensions is to consider four categories of actions:
   $Exclusive$ actions contribute towards only one dimension. We would need to have ${}_4C_1h = 4h$ actions like this, for any natural number h.

$Dual$ actions contribute towards only two dimensions. We would need to have ${}_4C_2r = 6r$ actions like this, for any natural number r.
$Triple$ actions contribute towards three dimensions. We would need ${}_4C_3s = 4s$ of these actions, for any s natural number.
$General$ actions contribute towards the four dimensions. We would need ${}_4C_4q = q$ of these actions, for any natural number q.

2) An example of fabric of actions

Figure 1 presents a fabric of actions that matches the guidelines stated above and uses $h = r = s = q = 1$. This gives a total of 15 actions. We can also consider a very special 16th action, namely the $null\ action$, i.e., $no\ action$. As shown in the table, each action can be identified with a vector in $\{0,1\}^4$ and each dimension is tied to eight actions. The headers include a brief name for each action, which refers to an example of the instrumentation of that particular action.

B. THE MEMORY MATRIX. NATURE AND STATE OF THE RELATIONSHIP

The memory matrix, *M*, is a matrix in $\mathcal{M}_{k\times4}(\mathbb{R})$, where *k* is the number of actions that the robot can *remember*. Each action realized by the human is stored in M starting from its bottom. When M is filled, the action stored in the first row is removed to leave room for the new action, stored in the k^{th} row. Each of the four dimensions of the relationship is stored in a different column. The concepts of $nature$ and $state$ of the relationship are defined on the basis of all the actions stored in M for the former and just the last actions for the latter.

The $nature$ of the relationship is identified with the dimensional weights, stored in vector W:

$$Action\ Load := \sum_{i=1}^{k}\sum_{j=1}^{4} M_{ij}$$

$$w_{Pat} := \frac{\sum_{i=1}^{k} M_{i1}}{Action\ Load}$$

$$w_{Pal} := \frac{\sum_{i=1}^{k} M_{i2}}{Action\ Load}$$

$$w_{Pro} := \frac{\sum_{i=1}^{k} M_{i3}}{Action\ Load}$$

$$w_{Lov} := \frac{\sum_{i=1}^{k} M_{i4}}{Action\ Load}$$

Clearly, $w_{pat} + w_{pal} + w_{pro} + w_{lov} = 1$

Let $d \in \{1, 2, ..., k\}$ be the $duration$ of a $state$, i.e., the number of actions that the robot will consider to define the $state$ of the relationship. d should be a small number compared to k, for example, $d \leq \frac{k}{5}$. Two reasonable numbers would be $k = 1000$ and $d = 200$.

The $state$ of the relationship is identified with the dimensional values, stored in the vector D. At any time, the values of the dimensions are defined as:

$$Pat := \frac{\sum_{i=k-d}^{k} M_{i1}}{d}$$

$$Pal := \frac{\sum_{i=k-d}^{k} M_{i2}}{d}$$

$$Pro := \frac{\sum_{i=k-d}^{k} M_{i3}}{d}$$

$$Lov := \frac{\sum_{i=k-d}^{k} M_{i4}}{d}$$

Notice that $Pat, Pal, Pro, Lov \in [0,1]$.

C. STRENGTH FUNCTION

The system uses a multi-attribute additive utility function, [41], that measures the strength of the human-robot relationship, understood as the consistency between the state of the relationship (represented with the dimensional values, D) and its nature (represented with the dimensional weights, W):

$$S(D,W) := D \cdot W =$$
$$w_{pat}Pat + w_{pal}Pal + w_{pro}Pro + w_{lov}Lov$$

Since the strength function is defined as the dot product of $nature$ and $state$, it will tend to throw high values when both are consistent with each other and low values when they are not. See Figure 2.

This resembles the concepts of additive utility function, system of weights, action of maximum utility and the personality model used in [42], [43], [44] and [45].

D. ORIGINS

This section describes the way in which the robot perceives and reacts to origins.

1) Hierarchy of needs and thresholds

The idea that needs adhere to a hierarchy, [46], and so do desires and expectations is subscribed here. Figure 3 shows the pyramid used in the system. When the primary needs are satisfied above a certain standard, the individual begins concerning about secondary ones. In this system, the strength function is the primary need: the robot will want to maintain its values above a minimum standard, designated Th_S, the threshold under which the robot perceives an origin. When the strength function is above an optimal standard, designated $Th_{S_{opt}}$, the robot will be concerned about its secondary needs, which are considered to be the dimensional weights.

The underlying rationale is as follows. The robot expects, before all, consistency from the human, measured through the strength function, S. However, the robot also wants the relationship to participate to a certain extent in each of the four dimensions. In other words, the robot does not want high values of S obtained on the basis of over-developing some

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	
	Cleaning	Secret Confession	Main goal	Caress	Hobby	Teach	Protection	Team Work	Play instrument or music.	Duty Assignment	Pats in the Back	Play	Scolding	Dancing	Talk	
	Exclusive				Dual						Triple				General	
Pat	1				1	1	1				1	1	1		1	**8**
Pal		1			1			1	1		1	1		1	1	**8**
Pro			1			1		1		1	1		1	1	1	**8**
Lov				1			1		1	1		1	1	1	1	**8**

FIGURE 1: *Loom and Fabric of Actions*. Example with 15 actions

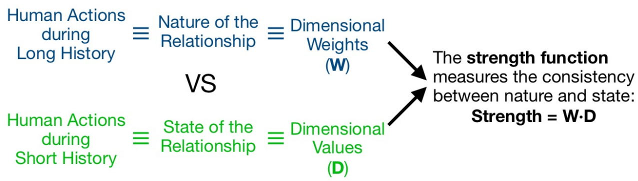

FIGURE 2: Nature · State = Strength

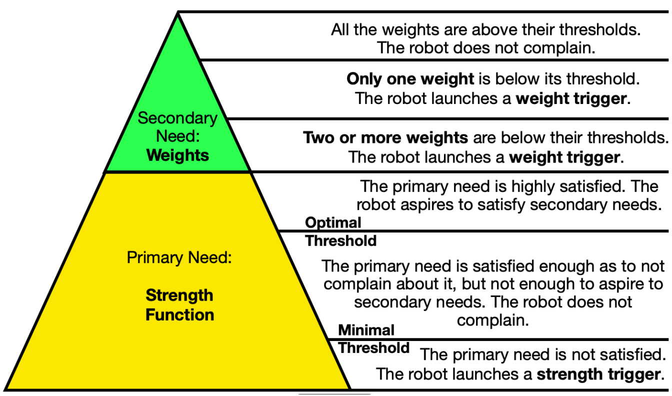

FIGURE 3: Pyramid of Needs

dimensions and neglecting others. This is formalized using thresholds for the dimensional weights: $Th_{w_{pat}}$, $Th_{w_{pal}}$, $Th_{w_{pro}}$, $Th_{w_{lov}}$. The robot will detect an origin if, while having $S > Th_{S_{opt}}$, any of the dimensional weights is below its respective threshold.

2) Stress

The system considers the concept of $stress$ as something that accumulates during an undesired situation. The stress function, St, will be fueled by any origin, will increase as the origin holds in time and will determine the moments at which the robot will escalate its complaint, i.e., it will determine the $intensity$ of the trigger. It is a monotonic increasing function, except when reset to 0. The latter occurs when there is a change in origin or when there is no longer an origin, i.e., when reconciliation has taken place. It adheres to the following guidelines:

1) Stress increases if the origin detected in the previous update is present in the new update.
2) Stress increases even more if the situation tied to the origin detected in the previous update has worsened. For example, if in the previous update there was a weight origin produced by $w_{lov} < Th_{w_{lov}}$ and in the new update the new w_{lov} is even less than the previous one, the stress will increase even more.

The following equations model these guidelines. Δ_{St} is the minimum increment in stress in each update if the origin detected in the previous update still holds. The subscript -1 refers to the value in the previous update. The equations for stress are the same for strength origins as for weight origins, so they can be shown at a time for any $E \in \{w_{pat}, w_{pat}, w_{pat}, w_{pat}, S\}$.

If $E \geq Th_E$, then $St_E = 0$
Else:
 if $E \geq E_{-1}$, then $St_E = St_E + \Delta_{St}$
 else, $St_E = St_E + \Delta_{St} + 100\frac{E_{-1}-E}{E_{-1}}$

The expression $100\frac{E_{-1}-E}{E_{-1}}$ turns the percentage worsening of the origin into percentage points that get added to the stress.

Figure 4 shows an example of the evolution of the stress function as the updates go by and the origin has not been cleared. It also shows how triggers are tied to stress, as explained in Section V-E.

E. TRIGGERS

This section introduces the classification of triggers and explains how the robot determines when and with what intensity to launch a trigger.

1) Classification of triggers according to origin and demanded action

When the origin comes from having $S < Th_S$, it is called a $strength\ origin$ and the robot will demand a triple action: the triple action that participates in the dimensions with the

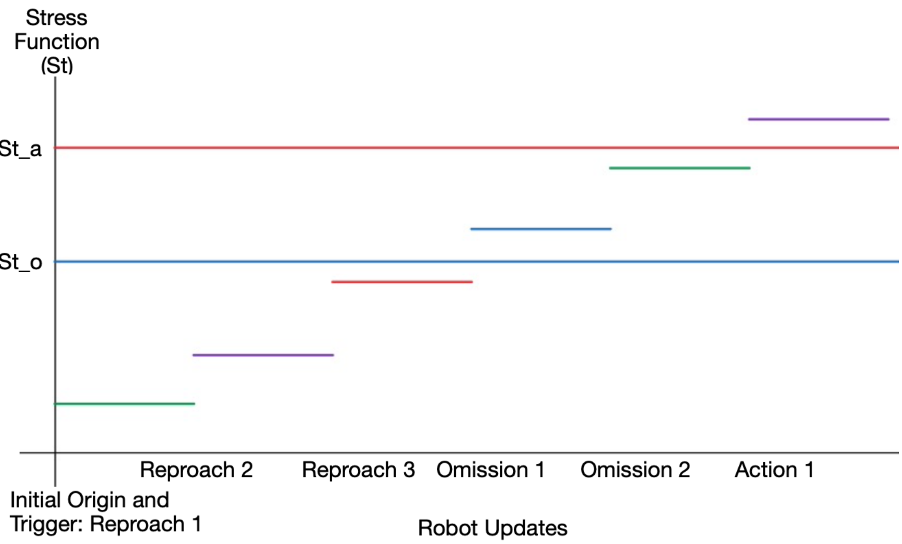

FIGURE 4: Example of Stress Evolution throughout Updates

three highest weights. This action is —besides the general action $Talking := (1,1,1,1)$— the one that contributes the most to increasing the strength function, i.e., the maximum utility action from a deterministic perspective. For example, if the three highest weights are w_{pat}, w_{pal}, w_{lov}, the action demanded by the robot will be $a_{12} := (1,1,0,1)$.

If the origin comes from having $S > Th_{S_{opt}}$ and one or more weights below their thresholds, it is called a $weight\ origin$. As in the case of $strength\ origins$, the demanded action will be the maximum utility action obtained deterministically:

1) If only one weight is below its threshold, the robot will demand the exclusive action associated to the dimension with the lowest weight. For example, if $w_{pat} < Th_{w_{pat}}$, the robot will demand action $a_1 := (1,0,0,0)$. This is the action that will contribute the most to increasing the low weight.
2) If two or more weights are below their threshold, the robot will demand the dual action (triple actions are reserved for strength origins) associated with the dimensions with the two lowest weights. For example, if the two lowest weights are w_{pat} and w_{lov}, the robot will demand action $a_7 := (1,0,0,1)$. Among dual actions, this is the one that contributes the most to increasing the lowest weights.

2) Classification of triggers according to intensity

According to their intensity, triggers are classified into three categories: $reproach$, $omission$ and $action$, corresponding to low, moderate and high intensity respectively. Figure 4 shows the following:

1) When the robot detects an origin, its first reaction will be launching a reproach trigger. This is instrumented in the form of a sentence that refers to the origin. For example, if there is a weight origin tied to only $w_{pat} < Th_{w_{pat}}$, the demanded action will be $a_1 := (1,0,0,0) := Cleaning$ and the robot will say: "It has been too long since last time you cleaned me". Once the reproach trigger is launched, the stress function begins to increase. It continues increasing if the human ignores the trigger or until the origin disappears. There are two thresholds for the stress function.
2) When the stress function exceeds Th_{St_o}, the robot will launch an omission trigger. In this kind of trigger, the robot uses a similar sentence, but also stops obeying the human's orders. It might say something like: "Do you want me just to help you out? I also need attention". The robot will begin questioning and declining to do some of the actions that are not the demanded action. This is the equivalent of a relationship that is malfunctioning because an unsolved conflict is going on.
3) When the stress function exceeds Th_{St_a}, the robot will launch an action trigger. In this kind of trigger, the robot will adopt a disruptive behavior. This could be simply instrumented with sounds or flashing lights. The robot will completely stop listening to orders. This is the equivalent of a serious argument between the human and the robot. Since the robot will stop listening to orders that are not the one it demanded in its trigger, the origin should disappear sooner or later.

F. NEGOTIATION AND RECONCILIATION

Reconciliation occurs when the origin is cleared. The interval from the appearance of the origin to reconciliation, during which the robot launches triggers and the human responds to them, is the equivalent of the negotiation. During this negotiation, the human is expected to adjust their behavior according to the robot's triggers. Some extent of those adjustments is expected to stay with the human after the reconciliation. In other words, the whole CRP is expected to have positive permanent effects on the human and their relationship with the robot.

VI. STRUCTURE OF THE ALGORITHM AND FURTHER MATHEMATICAL CONSIDERATIONS

This section discusses several considerations regarding the system. It also provides a simplified expression of the action flow of the system.

A. CONSIDERATIONS ABOUT TIME

The robot will detect and store in the memory matrix any action at the moment it is performed by the human. However, the robot will update its dimensional weights, dimensional values and strength function only occasionally. Each update represents an opportunity for the robot to detect an origin and, therefore, launch a trigger. The robot will not detect any origin until the memory matrix has been filled for the first time. This allows the robot to assess the nature and state of the relationship for the first time. From this moment on, the robot will use the following parameters:

1) C_1 is the minimum number of human actions between two updates. An example of arbitrary but reasonable criteria is $C_1 \approx \frac{d}{2}$. Before, it was suggested to use $d \approx \frac{k}{5}$, where d is the duration of the $state$ and k is the size of the robot's memory matrix. In this way, the human has a fair chance to reshape the state of the relationship before the next update, and can also alter the nature of the relationship to some extent, since $C_1 \approx \frac{k}{10}$
2) T_1 is the minimum number of days between two updates. While the appropriate value for this parameter will depend on the intended use of the robot, 7 or 14 days could be reasonable numbers for most scenarios.

After an update has occurred, the next one will take place as soon as both T_1 and C_1 have passed. This replicates human behavior in two ways. First, when a human involved in a relationship sees something they do not like, they usually do not complain immediately; they tend to wait for an appropriate moment. Second, if they have already complained, they usually do not expect an immediate solution; they tend to complain again, and perhaps escalate the complaint, if the situation has not been fixed after a reasonable period of time.

B. WEIGHTS, STRENGTH AND STRESS. ACTION FLOW

The following is a summary of the process and algorithm that the robot will use after having filled the memory matrix and each time that both T_1 and C_1 have passed from the previous update. The entire process is illustrated in Figure 5.

1) Strength test. If $Th_S \leq S \leq Th_{S_{opt}}$, the robot will not launch any trigger. This would put an end to the negotiation (and thus produce a reconciliation) if the robot had launched a trigger in its previous update; it will put the stress function back to 0. Otherwise, the robot will perform the following:
2) If $S < Th_S$, the robot will perceive a strength origin and will launch a trigger. The demanded action will be determined by the values of the dimensional weights and the intensity will be determined by the stress function according to the criteria described in Section V-E. Otherwise, we will have $S > Th_{S_{opt}}$ and the robot will run the weights test.
3) Weights test. If there is at least one $q \in \{Pat, Pal, Pro, Lov\}$ such that $w_q < Th_{w_q}$, the robot will perceive a weight origin and will launch a trigger. The demanded action will be determined by the values of the dimensional weights and the intensity will be determined by the stress function according to the criteria described in Section V-E. Otherwise, the robot will not launch any trigger. This would put an end to the negotiation (and thus produce a reconciliation) if the robot had launched a trigger in its previous update; it will put the stress function back to 0.

VII. SIMULATIONS

The model introduced in the previous sections was implemented in Python with the default parameters shown in Figure 6. The program would receive actions from $\{0, 1\}^4$ and would issue reactions expressed as a vector of the form $(O_S, O_w, O_w, I, A_1, A_2, A_3, A_4)$, where the components O express the origin (strength, one weight or two or more weights), the component I expresses the intensity of the trigger (1 for reproach, 2 for omission, and 3 for action), and the components A correspond to the demanded action.

Two different types of simulations were performed: one in which a human would provide actions to the program and another in which another program would simulate the human and would provide the actions.

A. SIMULATIONS WITH A HUMAN

This allowed us to try a variety of human behaviors, including those that aligned with the robot's demands and these which disregarded them. Several hundred iterations were done with both types. Thus, it was possible to study the behavior of the robot in the short run as a function of the human's actions and reactions. The following observations were made:

1) When the human tried to provide actions randomly, they indeed showed a preference for certain dimensions, which ultimately produced uneven weights, such as $(0.21, 0.28, 0.25, 0.26)$.
2) When the human provided actions aligned with the robot's triggers, they managed to keep the robot happy and to have few triggers, neither of them of intensity 3 and very few of intensity 2. In contrast, when the human disregarded the robot's triggers, the latter escalated to intensity 2 and 3, especially after the occasions in which the human changed the dimensional nature of their most frequent action or when the human began providing the null action too often.

These experiences show that the robot is good at interacting with a human and noticing whether the human is paying attention to and caring for the robot or not, and is able to change their attitude if not enough attention or care is given.

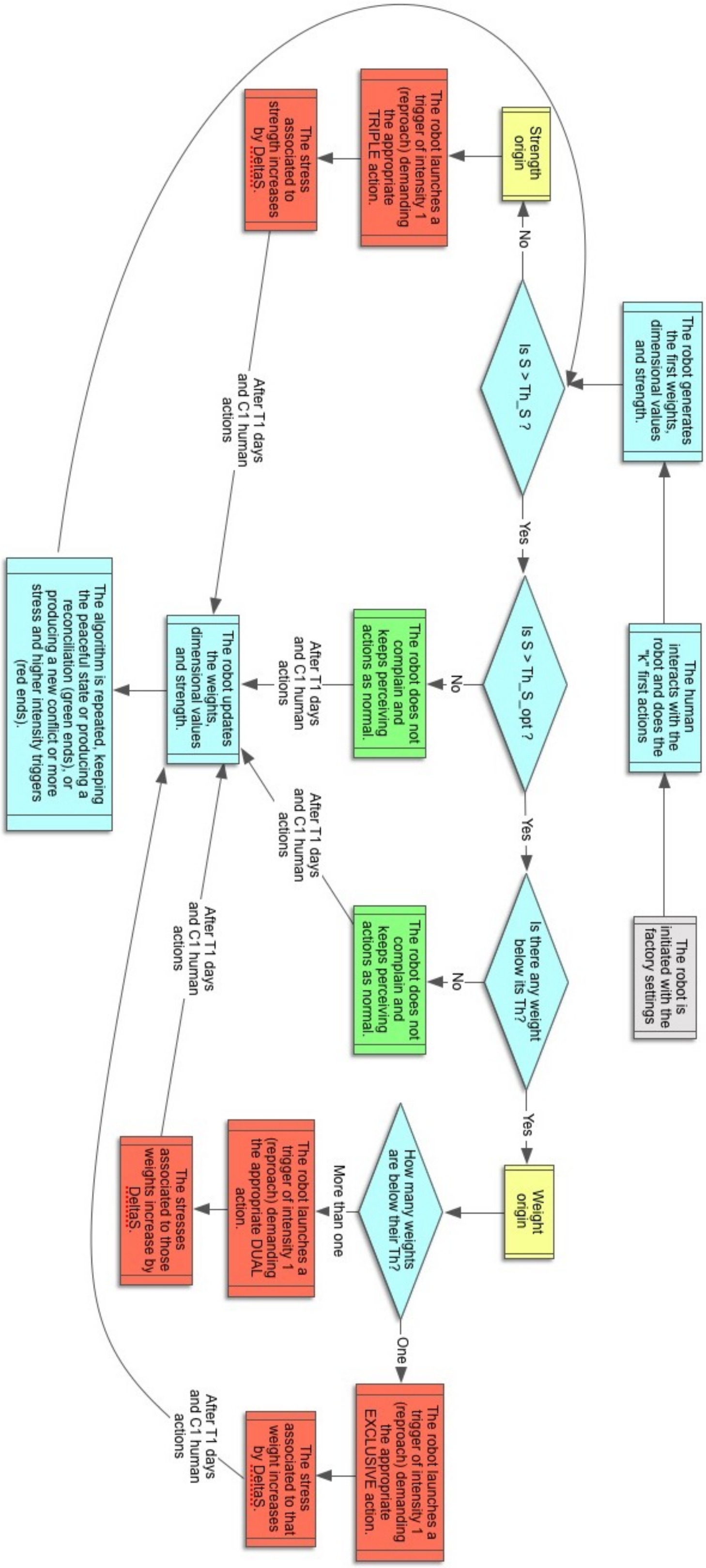

FIGURE 5: Algorithm Action Flow

Robot Parameters	k	d	Th_Sopt	Th_S	Th_w	Delta_St	Th_St_o	Th_St_a	C_1
	1000	200	0.5	0.4	0.2	10	50	75	100
Human Parameters	t	m	PPI						
	3	3	0.20						

FIGURE 6: Default Parameters

B. SIMULATIONS WITH A PROGRAM MODELING A HUMAN

To try thousands of iterations, we built a program capable of modeling the behavior of a human who pays reasonable attention to the robot's demands. This model identifies the human with a vector of $\mathbb{R}^{16}$ which is a probability distribution, $Human\ Probabilities := HP := (p_1, p_2, p_3, ..., p_{14}, p_{15}, p_{16})$. The sixteen components correspond to the probabilities that the human will do each action respectively. This comprises the fifteen actions included in our set of actions in Figure 1, plus the null action with probability p_{16}. These probabilities change when the robot launches a trigger, according to the trigger's demanded action and intensity, and evolve in time, in the absence of triggers, in a way consistent with observed human behavior. A *balanced* human would start with all actions and dimensions having similar probabilities, whereas an *unbalanced* human would have some actions and dimensions with substantially higher probabilities.

Three characteristics were studied and the following are the conclusions about them:

1) *Convergence*. Convergence was characterized as an equilibrium, i.e., as a prolonged absence of triggers, especially if the last ones took place further and further apart from each other. Convergence was achieved for both balanced and unbalanced humans, although more iterations and triggers were needed in the latter case. Figure 7 shows the distribution of triggers for 1,000 iterations with a very unbalanced human. It is visible how the triggers appear further apart from each other until they appear only incidentally.
2) *Sensitivity*. Once the equilibrium was achieved, there was the question of *How much can the human change without producing a trigger?* It turned out that the human can change their inclinations, i.e., the probabilities of their actions and dimensions, quite much and the robot would still not complain so far as the human keeps being as well-balanced as it was in the equilibrium. Triggers emerged only when the human became less well-balanced. Convergence was observed too after that. The extent of the human's *balanceness* was measured as a ratio between the probabilities of their most likely and least likely dimensions.
3) *Replicability*. Each series of iterations was repeated three times. This applies to both the series starting with the original *HP* and those starting with modifications in the *HP* after equilibrium was achieved. The results were essentially the same each of the three times.

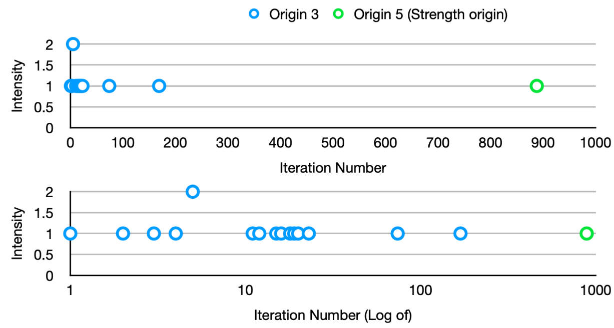

FIGURE 7: Distribution of triggers throughout 1000 iterations for a very unbalanced initial HP. Linear and logarithmic scales.

The results of these simulations can be summarized in the following points:

1) The system shapes the human towards balanceness, staying in equilibrium for as long as the latter lasts.
2) The system will tend to launch triggers of intensity 1 if the human pays reasonable attention to its demands, whereas it will tend to escalate to intensities 2 and 3 if the human disregards its complaints.
3) The system is consistent on replication.

VIII. CONCLUSIONS AND DISCUSSION

A. NOVEL ASPECTS AND GAPS IN THE LITERATURE THAT ARE FILLED BY THIS WORK

This paper has presented work on the novel idea of *conflict and reconciliation processes* (CRP) for human-robot interactions. It has introduced a novel personality and relationship characterization model, called *Pat Palprolov*, and has used it in combination with another novel concept, called the *loom and fabric of actions*, to quantify the *nature* and the *state* of a human-robot relationship on the basis of the *actions* performed by the human. Then it has defined a multi-additive utility function and a stress function, in combination with a system of thresholds and a hierarchy of needs, to determine when and how a robot would launch a CRP and when will the robot consider it finished. The interactions with a human for short-run analysis and with a model simulating a human for long-run analysis have suggested that the system tends to promote or appreciate human *balanceness* and consistency; the former defined as having no personality dimension over or underdeveloped and the latter defined as not breaking expectations.

It is remarkable that some of the modern developments closest in detail level (mechanisms, algorithms, equations, etc.) and spirit to the work presented here are not accessible through journals or conferences but through the public record of patents. This might be in part because companies or authors developing those specific devices prefer to keep them as industrial secrets or to get patents for them.

The system introduced in this paper resembles other recent developments in the state of the art while addressing challenges and filling gaps like those summarized in Section II-A. It can be said to have the following virtues:

1) It allows the robot to dynamically adapt to human behavior, meeting the needs of each individual and personalizing their relationship accordingly.
2) It is fully autonomous; potentially, with no human interaction needed once set up.
3) It is programmed to run for thousands of iterations in weeks or months, in a sort of long-run continuous learning.
4) It allows for shaping the user, especially their task-oriented behavior, with clear applications for education and therapeutic purposes.
5) It does not require intrusive or expensive tools, with the additional benefit of not depending on cloud services or network connections.
6) It does not rely on previous chips or smart boxes, making it inexpensive.
7) Its inputs are actions, which are easy to identify, rather than biosignals, much more difficult to capture.

B. ETHICAL CONCERNS

Some aspects of this work, including its foundational idea, could be considered provocative and raise ethical concerns. However, this is the case for most AI, especially when applied to social or affective robots. Dr. Breazeal, one of the leading figures in the field and director of the MIT Media Lab, in her interview for Soft Robotics Podcast, June 2022, speaks of how the closer the relationship among humans, the better the outcome usually is, and how she and her team are working on bringing those emotional elements into robots while being aware of the ethical and practical challenges of so doing [47]:

> We are trying to do it in ways that help people thrive and flourish, but now you are also talking about deeply persuasive systems; potentially, deeply manipulative systems, right? Systems that can engender not just trust like reliability or predictability, but trustworthiness, emotional bond and connection, you know... that could be extremely manipulative, right? So this is the provocative aspect of the work. It's, again, two sides of the coin.

C. LIMITATIONS, FUTURE RESEARCH AND POTENTIAL APPLICATIONS TO OTHER FIELDS

The main limitation of our study is the lack of experimentation with a variety of humans. Future research should conduct an experiment with a control group working with affective robots without CRP and an experimental group using affective robots with CRP. Future research could also consider the following aspects:

1) *Robot's personality*. If the programmers do not want the robot to help develop a balanced behavior in the human, it will be sufficient with changing the parameters. In particular, the thresholds for weights should be changed. For example, if the programmer wants the robot to inhibit the human's paternalistic dimension while developing the professional one, the following threshold vector can be used $TH := (0.16, 0.20, 0.25, 0.20)$. This would be equivalent to simulating a personality in the robot: a character more inclined to engage in certain dimensions than in others. Presumably, similar results would occur in terms of convergence, replicability and sensitivity.
2) *Demands without maximum utility*. When a dimensional weight is below its threshold, the robot will launch a trigger and demand the action of maximum utility in terms of increasing the low weight, which is obtained deterministically as the exclusive action associated to that dimension. For the robot to have a larger spectrum of reactions, future research could consider a criteria that also demands actions that are not of maximum utility.
3) *Keeping the flame alive*. If it was desired for the robot to launch CRP more often, perhaps as a way of keeping the flame of the relationship alive, it could be done in different ways. The simplest one, within the limits of the current design, would be to increase the thresholds for weights and for the minimum strength while lowering the threshold for optimal strength. A more elaborate option would be to introduce a new parameter that can measure $boredom$, understanding the latter as a prolonged state of quietness. With the help of this parameter, the robot could launch a CRP whenever too much time has passed from the last one.
4) *Actor, reactor and CRP among robots.* This study is limited to the case in which the robot is the actor and the human is the reactor. Future research could be done to switch the roles, allowing the robot to recognize triggers launched by the human and react to them. Also, a version of the current system could be used to allow robots to launch and carry out CRP with other robots.

A theory and model designed to incorporate CRP capabilities into affective robots could serve as a basis for wider research and models with applications to business, sociology, political sciences, international relationships and any other realm in which CRP could be used to improve the relationships between parties and to shape one or both of them. If a conflict can be handled and integrated as a part of a CRP which ultimately improves the bond between the parties while shaping one or both for the better, such a strategy might be of more benefit than other alternatives. Interestingly, the literature predominantly uses the concept of *reconciliation* for relationships between cultural, political or ethnic groups. But no mathematical model for *reconciliation dynamics* has yet been introduced.

Here are two examples from recent history that illustrate the potential of better understanding and handling CRP. In 2019 and 2020, a serious amount of social unrest took place in the USA in the context of racially loaded issues. The mainstream leaders in the country leaned towards one or another side, but it seems that they did not see the whole thing from a CRP perspective that could improve the relationships between the different social and racial sectors; therefore, they were not able to handle the situation with such a strategy in mind, which perhaps was a lost opportunity to have a more united country. In the same years, the Mexican president, in the context of a number of historical anniversaries, requested from the King of Spain an apology for the alleged abuses that the Spanish Empire would have inflicted upon Mexicans some centuries ago. The Spanish leaders predominantly disregarded the proposal, an attitude that counted with the approval of many intellectuals in the Hispanic world. It seems that neither of those leaders or intellectuals thought of such a proposal in light of a CRP perspective. A leader with these CRP ideas in mind could have taken the proposal as a trigger and approached the Mexican president to know more about what kind of pardon ceremony would he envision, with the idea that a well-designed and executed CRP could make more good to Spain, Mexico and the Hispanic block than a purely intellectual declination. It could be argued that an undue apology could also damage the countries and their relationships. It is with the help of models for CRP, of which the model presented in this paper is an example, that an optimal understanding, strategy and decision could be made.

ACKNOWLEDGMENT

The authors would like to thank David Rios Insua, mathematician, computer scientist, author, professor and AXA-ICMAT chair in Adversarial Risk Analysis, for providing the initial direction and supervision of this research; Leonardo Buceta Amo, physicist and engineer at Telefónica, for his invaluable work with the materialization of the system into a Python algorithm, with the simulations, and with other aspects of the editing of this article; and Jan Rychtar, Virginia Commonwealth University, for his review and feedback on this paper. The publication of this article as *open access* has been sponsored by *UNED funding for open access publishing*.

REFERENCES

[1] P. Rosalind, “Affective computing,” *M.I.T Media Laboratory Perceptual Computing Section Technical Report*, vol. 321, 1995.
[2] A. Banafa. (2016) What is affective computing?, last checked: 09/01/2023. [Online]. Available: https://www.bbvaopenmind.com/en/technology/digital-world/what-is-affective-computing/
[3] M. Spitale and H. Gunes, “Affective robotics for wellbeing: A scoping review,” in *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 2022, pp. 1–8.
[4] N. Churamani, S. Kalkan, and H. Gunes, “Continual learning for affective robotics: Why, what and how?” in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2020, pp. 425–431.
[5] C. Gena, C. Mattutino, G. Perosino, M. Trainito, C. Vaudano, and D. Cellie, “Design and development of a social, educational and affective robot,” in *2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*. IEEE, 2020, pp. 1–8.
[6] P. Uluer, H. Kose, E. Gumuslu, and D. E. Barkana, “Experience with an affective robot assistant for children with hearing disabilities,” *International Journal of Social Robotics*, vol. 15, no. 4, pp. 643–660, 2023.
[7] L.-E. Imbernon Cuadrado, A. Manjarres Riesco, and F. De La Paz Lopez, “Artie: An integrated environment for the development of affective robot tutors,” *Frontiers in computational neuroscience*, vol. 10, p. 77, 2016.
[8] M. Chen, J. Zhou, G. Tao, J. Yang, and L. Hu, “Wearable affective robot,” *IEEE Access*, vol. 6, pp. 64 766–64 776, 2018.
[9] M. DONG, “Method for predicting emotion status and robot,” Huawei Technologies, 2023, .US Patent 11,670,324 B2.
[10] C. H. PARK, “Emotional interaction apparatus,” George Washington University, 2020, .US Patent 10,593,349 B2.
[11] H. Wang, H. Wang, H. Chen, Y. Wang, Z. Li, P. Zhu, N. YAO, and J. Zhu, “Affective interaction systems, devices, and methods based on affective computing user interface,” Institute of Software of CAS, Shanghai Xiaoi Robot Technology, 2022, .US Patent 11,226,673 B2.
[12] C. Breazeal, F. Faridi, and O. Adalgeirsson, “Maintaining attention and conveying believability via expression and goal-oriented behavior with a social robot,” JIBO INC, Boston, 2018, .US Patent Publication US2018/029372A1.
[13] E. Un and M. Leung, “Emotion type classification for interactive dialog system,” Microsoft Technology Licensing LLC, 2017, .US Patent US 9,786,299 B2.
[14] S. L. Britt, E. J. Hill, A. LeBaron, D. R. Lawson, and R. A. Bean, “Tightwads and spenders: Predicting financial conflict in couple relationships,” *Journal of Financial Planning*, vol. 30, no. 5, pp. 36–42, 2017.
[15] R. Dunbar and A. Machin, “Sex differences in relationship conflict and reconciliation,” *Journal of evolutionary psychology*, vol. 12, no. 2-4, pp. 109–133, 2014.
[16] P. A. Cowan, C. P. Cowan, M. K. Pruett, and K. Pruett, “Fathers’ and mothers’ attachment styles, couple conflict, parenting quality, and children’s behavior problems: An intervention test of mediation,” *Attachment & human development*, vol. 21, no. 5, pp. 532–550, 2019.
[17] S. M. Elsayed-EkJiouly and R. Buda, “Organizational conflict: A comparative analysis of conflict styles across cultures,” *International Journal of Conflict Management*, vol. 7, no. 1, pp. 71–81, 1996.
[18] D. Cai and E. Fink, “Conflict style differences between individualists and collectivists,” *Communication Monographs*, vol. 69, no. 1, pp. 67–87, 2002.
[19] K. Lee and M. C. Ashton, “Psychometric properties of the hexaco personality inventory,” *Multivariate behavioral research*, vol. 39, no. 2, pp. 329–358, 2004.
[20] E. Tupes and R. Christal, “Recurrent personality factors based on trait ratings.(pp. 61-97): Usaf asd tech,” 1961.
[21] D. Skarmeas, “The role of functional conflict in international buyer–seller relationships: Implications for industrial exporters,” *Industrial Marketing Management*, vol. 35, no. 5, pp. 567–575, 2006.
[22] R. A. Cosier and M. Harvey, “The hidden strengths in family business: Functional conflict,” *Family Business Review*, vol. 11, no. 1, pp. 75–79, 1998.
[23] A. Menon, S. G. Bharadwaj, and R. Howell, “The quality and effectiveness of marketing strategy: Effects of functional and dysfunctional conflict in intraorganizational relationships,” *Journal of the Academy of Marketing Science*, vol. 24, pp. 299–313, 1996.
[24] H. Ren and B. Gray, “Repairing relationship conflict: How violation types and culture influence the effectiveness of restoration rituals,” *Academy of management review*, vol. 34, no. 1, pp. 105–126, 2009.
[25] R. C. North, H. E. Koch Jr, and D. A. Zinnes, “The integrative functions of conflict,” *Journal of Conflict Resolution*, vol. 4, no. 3, pp. 355–374, 1960.
[26] B. Russell, *The conquest of happiness*. George Allen Unwin Limited, 1930.
[27] N. Elias, E. Dunning *et al.*, *Quest for excitement. Sport and leisure in the civilizing process.* Basil Blackwell, 1986.

[28] D. W. Maynard, "On the functions of social conflict among children," *American sociological review*, pp. 207–223, 1985.
[29] H. T. W. Friends and I. People, "Dale carnegie," 1936.
[30] R. Petrone, ""You have to get hit a couple of times": The role of conflict in learning how to "be" a skateboarder," *Teaching and Teacher Education*, vol. 26, no. 1, pp. 119–127, 2010.
[31] V. Schmill Herrera, "Disciplina inteligente en la escuela. hacia una pedagogía de la no-violencia," 2008.
[32] E. Fromm, *The art of loving*. Harper Brothers, 1956.
[33] C. Castilla del Pino, *Teoría de los sentimientos*. Tusquets Editores, 2000.
[34] N. Schwarz and I. Skurnik, *Feeling and Thinking: Implications for Problem Solving*, 06 2003, pp. 263–290.
[35] G. Hertel, J. Neuhof, T. Theuer, and N. L. Kerr, "Mood effects on cooperation in small groups: Does positive mood simply lead to more cooperation?" *Cognition & emotion*, vol. 14, no. 4, pp. 441–472, 2000.
[36] B. A. Mellers, A. Schwartz, and A. D. Cooke, "Judgment and decision making," *Annual review of psychology*, vol. 49, no. 1, pp. 447–477, 1998.
[37] M. Argyle and M. Henderson, "The rules of friendship," *Journal of social and personal relationships*, vol. 1, no. 2, pp. 211–237, 1984.
[38] M. S. Guide. (2021) Understanding conflict - meaning and phases of conflict, last checked: 09/01/2023. [Online]. Available: https://www.managementstudyguide.com/understanding-conflict.htm
[39] R. Vimal, "Meanings attributed to the term'consciousness': an overview," *Journal of Consciousness Studies*, vol. 16, no. 5, pp. 9–27, 2009.
[40] S. Dehaene, H. Lau, and S. Kouider, "What is consciousness, and could machines have it?" *Science*, vol. 358, no. 6362, pp. 486–492, 2017.
[41] A. Mateos, A. Jiménez, and S. Rıos-Insua, "Solving dominance and potential optimality in imprecise multi-attribute additive problems," *Reliability Engineering & System Safety*, vol. 79, no. 2, pp. 253–262, 2003.
[42] J. G. Rázuri, P. G. Esteban, and D. R. Insua, "An adversarial risk analysis model for an autonomous imperfect decision agent," in *Decision Making and Imperfection*. Springer, 2013, pp. 163–187.
[43] P. G. Esteban and D. R. Insua, "A model for an affective non-expensive utility-based decision agent," *IEEE Transactions on Affective Computing*, vol. 10, no. 4, pp. 498–509, 2017.
[44] S. Liu and D. R. Insua, "An affective decision-making model with applications to social robotics," *EURO Journal on Decision Processes*, vol. 8, no. 1-2, pp. 13–39, 2020.
[45] S. Liu, "Toma de decisiones afectivas para agentes robóticos sociales." Ph.D. dissertation, Universidad Autónoma de Madrid, 2020.
[46] A. H. Maslow, "A theory of human motivation." vol. 50, pp. 370–396, 1943.
[47] C. Breazeal and M. ElDiwiny. (2022) Social robots, min 24:00, last checked: 9/16/2022. [Online]. Available: https://soundcloud.com/ieeeras-softrobotics/cynthia-breazeal-social-robots

GUILLERMO ALVAREZ PARDO completed a B.S. and a M.S. in Mathematics at the Complutense University of Madrid and The University of Manchester (2010), and an MBA at the Monterrey Institute of Technology and Higher Education in Mexico (2016). He is currently taking a PhD in Control and Systems Engineering at National University of Distance Education, Spain, while working as a professor and researcher at Cuesta College, California. His research lines focus on conflict and reconciliation processes in human-robot relationships and the development of math research opportunities for undergraduate students. At Cuesta College, he is the developer, coordinator and main professor of *Math 290: Introduction to Research*; the PI of NSF-awarded (2327792) IHBEM-MAVEN, which focuses on vaccination modeling for equity; the Math lead at NSF-awarded (2236402) SHIP-UR, and a collaborator of NSF-awarded (2110112) C6-LSAMP. His work or his students' work has been presented at conferences like MathFest, IUSE Summit, NCUR, AMATYC's Annual Conference, and IEEE's ISEC.

ERNESTO FABREGAS ACOSTA received the B.S. degree in automation control and the M.S. degree in digital systems from the Universidad Tecnológica de La Habana "José Antonio Echeverría", Cujae, Cuba, in 2004 and 2008, respectively, and the Ph.D. degree in computer science from the Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, in 2013. From 2015 to 2019, he was a Postdoctoral Fellow with UNED. Since 2019 to May 2022, he has been an Assistant Professor and from May 2022 he is Associate Professor with UNED. His current research interests include control of multi-agent systems, machine learning, images processing, mobile robot control, remote laboratories, and engineering education.