SUBMITTED TO IJSR. PLEASE DO NOT REDISTRIBUTE.

Noname manuscript No.
(will be inserted by the editor)

Toward Improving Human-Robot Collaboration with Emotional Awareness

Mahni Shayganfar · Charles Rich · Candace L. Sidner

Received: date / Accepted: date

Abstract Current computational theories used for human-robot collaboration specify the structure of collaborative activities, but are weak on the underlying processes that generate and maintain these structures. We argue that emotions are crucial to these underlying processes and have developed a new computational theory, called Affective Motivational Collaboration Theory, that combines emotion-based processes, such as appraisal and coping, with collaboration processes, such as planning, in a single unified framework. To illustrate the application of this new theory, we present detailed computational walkthroughs contrasting the behavior of an emotionally aware robot with an emotionally ignorant robot in the same situations. These walkthroughs are the starting point for our implementation of the theory.

Keywords Human-Robot Collaboration · Emotional Awareness · Affective Motivational Collaboration Theory

1 Introduction

A key aspect of the sociability of robots is their ability to collaborate with humans in the same environment. Collaboration is a coordinated activity in which the participants work jointly to satisfy a shared goal [23]. There are many challenges in achieving a successful collaboration between robots and humans. To meet these challenges, it is crucial to understand what makes a collaboration not only successful, but also efficient. Existing computational

Mahni Shayganfar · Charles Rich · Candace L. Sidner
100 Institute Road, Worcester, MA, USA 01609-2280
Tel.: +1 508-831-5357
Fax: +1 508-831-5776
E-mail: mshayganfar@wpi.edu
E-mail: rich@wpi.edu
E-mail: sidner@wpi.edu

models of collaboration explain some of the important concepts underlying collaboration; such as the presence of a reason for collaborators' commitment, and the necessity of communicating about mental states in order to maintain progress over the course of a collaboration. The most prominent collaboration theories are based on plans and intentions [11] [23] [32], and are derived from Bratman's BDI architecture [7]. Two theories, Joint Intentions [11] and SharedPlans [20, 21, 23], have been used to support teamwork and collaboration between humans and robots or virtual agents [9] [40] [59] [68]. However, these theories explain only the structure of a collaboration. For instance, in SharedPlans theory collaborators build a shared plan containing a collection of beliefs and intentions about the actions in the plan. Collaborators communicate these beliefs and intentions via utterances about actions that contribute to the shared plan. This communication leads to the incremental construction of a shared plan, and ultimately successful completion of the collaboration. In contrast, in Joint Intentions theory, the notion of joint intention is viewed as a persistent commitment of the team members to a shared goal. In this theory, once an agent enters into a joint commitment with other agents, it should communicate its private beliefs to other team members.

Although existing collaboration theories explain the important elements of a collaboration structure, the underlying processes required to dynamically create, use, and maintain the elements of this structure are largely unexplained. For instance, a general mechanism has yet to be developed that allows an agent to effectively integrate the influence of its collaborator's perceived or anticipated emotions into its own cognitive mechanisms to prevent shared task failures while maintaining collaborative behavior. Therefore, a process view of collaboration must include certain key elements. It should inherently involve social interactions since all collaborations occur between social agents, and it should essentially constitute a means of modifying the content of social interaction as the collaboration unfolds. The underlying processes of emotions possess these two properties, and social functions of emotions explain some aspects of the underlying processes in collaboration. This paper makes the case for a process model of emotions and demonstrates how it furthers collaboration between humans and robots.

Humans are emotional and social beings; emotions are involved in many different social contexts including collaboration. Although there are purely personal emotions, most emotions are experienced in a social context and acquire their significance in relation to this context [36]. For instance, humans are influenced by the emotions of those around them. They also have emotions about the actions of people around them. They have emotions about the events that occur in the other people's lives. Also, humans' concern about their relationships with others elicits emotion. They can feel emotion about their personal successes and failures and those of others. Moreover, socially shared and regulated emotions can provide social meanings to events happening in the environment [67].

There is also a communicative aspect of emotions. For instance, emotions are often intended to convey information to others [15]. Emotions are also involved in verbal behaviors. For instance, an utterance can include both content and relational meaning. An emotion might appear to be elicited by the content of the utterance, but in fact be an individual's response to the relational meaning [45]. The interpretation of these relational meanings are handled by the appraisal of events. Appraisal processes give us a way to view emotion as social [63]. Meaning is created by an individual's social relationships and experiences in the social world, and individuals communicate these meanings through utterances. Consequently, the meaning of these utterances and the emotional communication change the dynamic of social interactions. A successful and effective emotional communication necessitates ongoing reciprocal adjustments between interactants that can happen by interpreting each other's behaviors [36]. This adjustment procedure requires a baseline and an assessment procedure. While the components of the collaboration structure, e.g., shared plan, provide the baseline, emotion-related processes provide the assessment procedure.

Since collaboration is a type of social context, the social functions of emotions are required for an agent to perform adequately in such an environment. In this paper, we present two pairs of hypothetical interaction scenarios. Each pair contrasts an emotionally-aware with an emotionally ignorant robot interacting with a human in the same situation. These scenarios highlight the necessity of giving robots the capacity to understand and regulate emotions, as well as to provide emotion-driven responses. We then briefly introduce Affective Motivational Collaboration Theory which explains the underlying processes of emotions and collaboration. The emotion-aware examples show how the mechanisms of this theory are involved in agreeing on a shared goal with a robot (Sections 3.3 and **??**), and delegating a new task to the robot (Sections 3.5 and **??**). The emotion ignorance examples are the same, except that the robot ignores the human's verbally or nonverbally expressed emotions. The same four examples in Sections 3.3 to 3.6 are revisited in more detail in our computational walkthroughs in Section **??**. In this section, we show how the mechanisms of Affective Motivational Collaboration Theory operate to produce the robot behaviors in Section 3.

As we discussed above, there are certain types of emotion-regulated mechanisms with which a collaborative robot can modify and maintain a collaboration structure (e.g., shared plan). We explain these mechanisms and their corresponding operations in Affective Motivational Collaboration Theory. In this paper, we briefly describe some parts of this theory that are required to discuss our examples. We have also implemented all the rules associated with each mechanism using JESS (JAVA Expert System Shell), in order to generate the same type of collaborative behaviors as appear in our examples. In the future, we are going to use these rules and the processes involved in each mechanism to develop collaborative behaviors in an interactive robot.

2 Related Work

The prominent collaboration theories are mostly based on plans and joint intentions [11,23], and they were derived from the BDI paradigm developed by Bratman [7] which is fundamentally reliant on folk psychology [47]. The two theories, Joint Intentions [11] and SharedPlans [23], have been extensively used to examine and describe teamwork and collaboration. The SharedPlans theory is a general theory of collaborative planning which accommodates multi-level action decomposition hierarchies, and allows the process of generating complete plans. The SharedPlans theory shows how a group of collaborators can incrementally form and execute a shared plan, and describes how a shared plan coordinates their activities towards achieving a shared goal. Furthermore, SharedPlans theory emphasizes that collaborative plans are an interleaving of collaborators' mutual beliefs and intentions about the actions in the plan [20, 21,23]. In contrast, the Joint Intentions theory as another formal theory of collaboration is based on the idea of individual and joint intentions to act as a team member. In this theory, a joint intention is a shared commitment to perform an action while in a group mental state. Joint Intentions theory describes how team members can jointly act together by sharing mental states about their actions while an intention is viewed as a commitment to perform an action [11].

There are many research focusing on different aspects of collaboration based on different collaboration theories, i.e., SharedPlans, Joint Intentions, and hybrid theories of collaboration, e.g., STEAM [61]. Some of these works present algorithms and computational models in a teamwork environment based on the underlying structure of the SharedPlans theory [34,35,68,69], and Joint Intentions theory [9,41]. The hybrid teamwork model, STEAM [61], has also been successfully applied to a variety of domains [26,29,37,53]. All of the works presented in this section lack a systematic integration of collaboration theories with some theories capable of describing underlying collaboration processes. Therefore, they either do not explain the structure and the underlying processes of collaboration, or their approach in either or both of these views is application oriented. The collaboration structure of Affective Motivational Collaboration Theory is based on the SharedPlans theory [20,21, 23], and it focuses on the processes that generate, maintain and update this structure based on mental states. COLLAGEN [48,49] incorporates certain algorithms for discourse generation and interpretation, and is able to maintain a segmented interaction history, which facilitates the discourse between a human and a robot [50]. We use its latest incarnation, i.e., Disco, for our implementation.

Furthermore, there are some works focusing on the concepts of robot assistants [10], or teamwork and its challenges in cognitive and behavioral levels [42, 54]. Some researchers have an overall look at a collaboration concept at the architectural level [13,14,60]. There are other concepts such as joint actions and commitments [19], dynamics of intentions during collaboration [31] providing more depth in the context of collaboration. Some of these works emphasize the applicability of emotions in their architectures, and some others emphasize the collaborative aspect of their robots. The applications of different prominent collaboration theories show the importance and the applicability of these theories in robots and collaborative systems. The following examples briefly review some of the applications of artificial emotions and appraisal theory of emotions in robots and autonomous agents.

Applications of Artificial Emotions – There are many research areas, including robotics and autonomous agents, that employ the structure and/or functions of emotions in their work with a variety of motivations behind modeling emotions [66]. Some of these works are inspired by specific psychological theories, some are freely using the concept of emotion without using the theoretical background in social sciences [62], and some are using a combination of concepts from the psychological theories [28]. We can also see the application of emotion theories in designing companion robots, robots capable of expressing emotions and social behaviors, as well as robots which can convey certain types of emotion products, e.g., empathy [8,44,58]. Robots also use emotions theories for some other purposes such as automatic affect recognition using different modalities [70], and behavior adaptation [33].

Furthermore, emotions have different intra/interpersonal functions. Motivation is one of the crucial functions of emotions, since it can initiate, direct and maintain goal-directed behaviors. The motivation mechanism in our work is inspired by Murray's theory as well as Bach's approach on Dörner's theory [3–6]. It is focused on the role of emotion-driven motives in cognitive processes, e.g., intention formation, during collaboration.

Applications of Appraisal Theory – Appraisal theories of emotion were first formulated by Arnold [2] and Lazarus [30] and then were actively developed in the early 80s by Ellsworth and Scherer and their students [51, 52,55–57]. Computational appraisal models have been applied to a variety of uses including contributions to psychology, robotics, AI, and HCI. For instance, Marsella and Gratch have used EMA [38] to generate specific predictions about how human subjects will appraise and cope with emotional situations and argue that empirical tests of these predictions have implications for psychological appraisal theory [18]. However, EMA does not focus on the dynamics of collaborative contexts. There are several examples in artificial intelligence and robotics of applying appraisal theory [1,27,38]. In robotics, appraisal theory has been used to establish and maintain a better interaction between a robot and a human [27,52,65]. Appraisal theory has also been used in robots' decision making [16], or in their cognitive systems [24,25]. In the virtual agents community, empathy and affective decision-making is a research topic that has received much attention in the last decade [39,43,46,64].

3 Example Scenarios

3.1 Backstory

The scenarios transpire in a lunar facility using collaborative robots to work with astronauts. The mission is to finish installing the solar panels required to provide energy for the operation of the facility. Most of the panels have already been installed. However, the facility is now faced with a low batteries situation, which forces the team to be cautious about consuming energy. A female astronaut is inspecting the working conditions in the field and planning the installation of the remaining panels in collaboration with the robot.

3.2 Astronaut-Robot Interaction

The Robot and the Astronaut will collaborate with each other to achieve their shared goal, which is to install two solar panels. They will face various difficulties, ranging from the task being unpleasant and challenging to conflicts between their private and/or shared goals because of a blocked or a protracted sub-task. The Robot and the Astronaut will go through a series of assessment processes to figure out a) how did the current blockage happen? b) why is the current task is blocked? and c) what is the next action they are going to take? The Robot uses its cognitive abilities and communication skills to overcome these problems, to motivate the Astronaut, and to propose alternative tasks.

3.3 Agreeing on Shared Goal (Emotion Awareness)

This hypothetical interaction and the emotion ignorant version in the next section demonstrate how the process of agreeing on a shared goal is improved by the Robot's awareness of its collaborator's emotions (here, frustration).

A1. Astronaut: Oh no! Finishing the quality check of our installation with this measurement problem is so frustrating. I think we should stop now!

A2. Robot: <u>I see. This is frustrating.</u> But, I can help you with the measurement tool and we can finish the task as originally planned.

A3. Astronaut: Can you fix the measurement tool?

A4. Robot: The next task is fixing the panel and it requires you to prepare and attach the welding rod to your welding tool. To save our time, I will fetch another measurement tool while you are preparing your welding tool.

A5. Astronaut: That would be great!

The Astronaut's first turn (A1), shows her verbally conveying her frustration with respect to a malfunctioning measurement tool. In reply, the Robot's first turn (A2) shows the Robot perceiving the Astronaut's frustration and acknowledging it verbally. The underlined section of the Robot's utterances (in turn A2) shows the influence of using emotion-driven processes which lead to acknowledgement of the Astronaut's emotion. Notice the absence of these utterances as the consequence of ignoring the Astronaut's emotions in the corresponding turn (B2) in the next example. In Section **??**, we will show how the computational mechanisms discussed in Section 4.1 are involved in this process, specifically how these emotion-driven goal-directed mechanisms work together and lead to the Robot's behavior of acknowledging the perceived emotion of the Astronaut properly, thereby avoiding unsuccessful termination of the collaboration.

Continuing in turn A3, the Astronaut's utterance shows a change of underlying belief from termination of the collaboration to the possibility of seeking instrumental support by asking the Robot whether it is possible to fix the measurement tool. Notice that the proper acknowledgement of the Astronaut's emotion helps to change her emotion from frustration to neutral. Now that the Astronaut is no longer expressing a negative emotion (i.e., frustration), and is instead asking for help, the Robot can provide an alternative task as a potential solution (A4). This example illustrates an advantage of emotional awareness: although the Robot, according to the shared plan, could have provided the same alternative task as a solution to the Astronaut immediately, it delayed instead, providing the potential solution after first dealing with the Astronaut's negative emotional state. Finally, since agreeing on a shared goal is a collaborative negotiation process, emotion awareness plays a crucial role in providing a fair offer. As a result, the Astronaut's response in the last turn (A5) shows acceptance and contentment with the Robot's potential solution to continue collaboration and agreement on the shared goal.

3.4 Agreeing on Shared Goal (Emotion Ignorance)

This example shows the beginning of the same process of agreeing on a shared goal as in the previous section, except that it diverges from reaching an agreement. It begins with the same utterance (B1) as the previous example (A1).

B1. Astronaut: Oh no! Finishing the quality check of our installation with this measurement problem is so frustrating. I think we should stop now!

B2. Robot: I can help you with the measurement tool, or we can terminate this task. <u>What do you want me to do?</u>

B3. Astronaut: As I said the measurement tool does not work properly. We can not continue!

B4. Robot: Okay. Do you want me to fix this problem or terminate the task?

B5. Astronaut: Can you fix my measurement tool?

B6. Robot: I cannot fix your measurement tool, but I can fetch another one for you if you want?

B7. Astronaut: No, I don't want another measurement tool! We don't have time for that!

B8. Robot: Okay. You want me to terminate this task. Terminating this task can influence the quality of installation of this solar panel which can cause the mission to fail. Or, do you want us to work on another task? This can help us to install the panel using your welding tool, but I do not know whether the quality of our installation will be acceptable.

B9. Astronaut: I told you we have this problem and we should terminate the mission! We cannot continue without the measurement tool!

As mentioned earlier, emotional awareness is beneficial in collaboration by channeling the collaboration process towards the shared goal. Without emotional awareness a collaborative robot will try to maintain the status of the shared goal and protect it from failure without considering its collaborator's negative emotion. In this example, the emotionally ignorant Robot does not acknowledge the Astronaut's frustration (compare B2 with A2 above), since it does not perceive that emotion. Then, while negotiating the shared goal, the Robot fails to offer a potential solution with respect to the Astronaut's emotional state. As a result, it causes the failure of the negotiation process during collaboration.

The Robot in this example does not perceive the Astronaut's emotion, and therefore does not include the Astronaut's emotion (frustration) as an influential factor in its computational mechanisms (see details in Section 4). Hence, in the Robot's first response (B2), it does not acknowledge the Astronaut's emotion, and instead immediately conveys two available alternative actions according to the existing shared plan, and asks the Astronaut to select between them. As shown in the Astronaut's response (B3), this immediate proposal does not result in any progress in collaboration. As a result, the Astronaut repeats herself about the task status while still expressing frustration. The Astronaut's response does not change the Robot's mental state and this causes the Robot to try to repeat its own question while still missing the Astronaut's frustration (B4). The Robot's utterance creates an ambiguous assumption for the Astronaut about whether the Robot can fix the broken measurement tool for her. This ambiguity makes the Astronaut even more frustrated and causes her to ask a question to remove the ambiguity of the Robot's proposal (B5). In return, the Robot not only misses the Astronaut's intensified frustration, but also nullifies the Astronaut's assumption about fixing the malfunctioning measurement tool and proposes the potential solution of replacing the tool, and asks whether the Astronaut agrees on that (B6). As we shall see, the

Robot's reasoning is different in B6 because its assessment of the Astronaut's cognitive state and its strategies for motivating the Astronaut are different.

In B7, the Astronaut modifies its assumption and announces the shortage of time as justification for expressing her anger. At this point, the Robot's response becomes more crucial since its wrong method of interaction and emotionally ignorant behavior shifted the Astronaut's emotional and mental states into a noncollaborative status. Consequently, the Robot again attempts to revive the collaboration process; it provides more information about the repercussions of terminating the collaboration process, to see whether the Astronaut can pursue another task (B8). Notice the underlined section of the Robot's turn B8 indicates its reasoning about the problem dissociated from the Astronaut's mental state. Finally, the poor interaction of the Robot caused by its emotionally ignorant behavior leads to an unsuccessful termination of their collaboration (B9).

3.5 Task Delegation (Emotion Awareness)

In this and the next section, a different collaborative behavior, task delegation, is used to illustrate how collaboration critically depends on understanding how worried the other collaborator is. This example shows that when the Robot is aware of the Astronaut's worry, it can use its own Motivation mechanism driven by emotions to come up with a way to alleviate her worry. Its solution is to postpone all questions as long as possible.

C1. Astronaut: I still have some problems with attaching the first panel! We do not have enough time. You should begin to install the second panel.

C2. Robot: Okay. Don't worry. I can handle that.

C3. Astronaut: I will try to fix it ASAP.

C4. Robot: I might need to ask some questions while I am installing the second panel.

C5. Astronaut: That's fine. Just let me know.

At the beginning of this example the Astronaut (C1) is worried because of the lack of time to achieve the shared goal (finishing installation of solar panels). She proposes that the Robot begin installing the second panel, since the first one still has some problems. The Robot in its first turn (C2), perceives the Astronaut's emotion (i.e., worry) and, using the same cognitive mechanisms (see Section 4.1), acknowledges the Astronaut's emotion just as it did in Section 3.3. The underlined utterance in the Robot's turn C2, shows the Robot's awareness of the Astronaut's emotion. Also, because of perceiving the Astronaut's worry the Robot does not ask her if it is okay to leave the current task (which was helping the Astronaut to install the first panel). The reason is that the Robot knows redirecting the Astronaut's attention away from the object of worry will create frustration, as the function of worry is to resolve the object of worry.

After acknowledging the Astronaut's emotion (C2), the Robot infers that it needs to postpone asking questions about the missing parts of the shared plan, since installing a panel is a collaborative task and some of the primitive tasks need to be done by the Astronaut. Then, the Astronaut perceives the Robot's response as a proper acceptance of the task delegation and tries to communicate the status of her own task, even though she is still worried about finishing the overall task (C3). Now, the Robot perceives the Astronaut's worry and without asking detailed questions about the delegated task, declares the possibility of asking some follow-up questions later (C4). Thus, the Robot not only prevents overwhelming the Astronaut with many questions, but its utterance implicitly reveals its knowledge of the tasks' requirements to the Astronaut. As a result, the Robot's approach to acceptance of the delegated task mitigates the Astronaut's negative emotion (i.e., worry) which makes her to respond positively to the Robot's proposal.

3.6 Task Delegation (Emotion Ignorance)

This last example illustrates how ignoring the collaborator's emotions in a task delegation situation can negatively impact the progress of a collaboration. In this example, the emotionally ignorant Robot is planning in its most efficient manner (because time is short) and therefore asking a lot of questions so that it can work out the plan. But asking questions exacerbates the Astronaut's worry which leads to an unsuccessful collaboration due to the lack of time.

D1. Astronaut: I still have some problems with attaching the first panel! We do not have enough time. You should begin to install the second panel.

D2. Robot: Okay. I can handle that. <u>Is it okay if I leave the current task though?</u>

D3. Astronaut: Yes, that's fine. I can finish it on my own.

D4. Robot: Okay. But according to the plan, I need your help when we need to weld the panel. Will you be able to help me?

D5. Astronaut: Yes, I will.

D6. Robot: We also need your welding tool and supplies as well as the measuring tool to install the second panel.

D7. Astronaut: Just let me know whenever you need something!

D8. Robot: Thanks. I will. How about the conflict that might occur while I am trying to fix the second panel? You are going to need my help at the same time.

D9. Astronaut: Robot, I really don't understand what you are talking about!

D10. Robot: Do you want me to provide some examples?

D11. Astronaut: We don't have time for this anymore!

The very first Astronaut's utterance above (D1) is the same as the first utterance in the previous example (C1). The Astronaut is worried and expresses her worry. However, the Robot does not perceive and consequently does not acknowledge the Astronaut's emotion. As a result, the Robot responds to the Astronaut by trying to determine a proper solution for an action selection problem. The action selection problem is primarily caused by the shift in the Robot's focus of attention from an unfinished ongoing task (unsatisfied postconditions) to a new partially known nonprimitive task (i.e., installing the second panel). Therefore, the Robot immediately tries to confirm leaving the current unfinished task (D2). Notice the absence of acknowledgment of the Astronaut's emotion by the Robot in this turn (compare C2 above and D2 here).

This absence of emotion awareness is the beginning of the failure of the task delegation process. As we can see, the Robot's response does not mitigate the Astronaut's worry about the future of the collaboration. The underlined section in D2 shows the Robot's need for confirmation of leaving an unfinished task. Next, the Astronaut tries to help the Robot select the proper action by responding positively about the Robot leaving the current task (D3). Now, the Robot shifts its focus of attention to the new task and starts to ask about required information such as task dependencies, existing preconditions and required resources (D4). Although this type of interactive behavior is crucial in many collaborative contexts, here it is counter-productive. Thus, the Astronaut curtly responds to the Robot's question (D5). The Robot then asks another question about the required inputs for the task (D6). At this point, since the Astronaut believes that the Robot's questions are unnecessary, she becomes frustrated and impatiently answers the Robot's question (D7). However, once again, not only does the Robot miss the Astronaut's emotion, but it also wants to prevent failure of a task in the future (D8). Notice that the underlined section in D8 is the result of the Robot's inference about the possibility of a future problem. Also, note that while the Robot is capable of operating based on a partial plan, instead, it continues to attempt to develop a complete plan due to ignorance of the Astronaut's frustration. Then, the Astronaut does not understand the event referenced by the Robot and since she is frustrated, she does not even try to remove the ambiguity of the existing issue (D9). Once again, the Robot misses the Astronaut's frustration and tries to see whether the Astronaut wants the Robot to clarify the issue for her by providing her some examples (D10). The underlined utterance in D10 indicates another situation in which the Robot misses the Astronaut's emotion. At last, the Astronaut terminates the collaboration task because of the lack of time (D11).

4 Computational Framework

In this section, we briefly describe Affective Motivational Collaboration Theory and the five underlying emotion-regulated mechanisms in this theory. Each mechanism constitutes one or more processes which are involved in generating collaborative behaviors for the Robot. We also explain different types of mental states in our computational framework. Notice in Fig. 1 there are two components, Perception and Action, which are not part of Affective Motivational Collaboration Theory. These components only provide required input and output to our framework which can differ based on the capabilities of a particular sociable robot.

4.1 Affective Motivational Collaboration Theory

Affective Motivational Collaboration Theory (see Fig. 1) is about the interpretation and prediction of the observable behaviors in a dyadic collaborative interaction. The collaboration structure of Affective Motivational Collaboration Theory is based on the SharedPlans theory of collaboration [20,21,23]. Affective Motivational Collaboration Theory focuses on the processes that generate, maintain and update this structure based on mental states. The collaboration structure is important because social robots ultimately need to co-exist with humans, and therefore need to consider humans' mental states as well as their own internal states and operational goals. The processes involved in collaboration are important because they explain how the collaboration structure is formed and dynamically evolved based on the collaborators' interaction.

Affective Motivational Collaboration Theory focuses on the processes regulated by emotional states. It aims to explain both rapid emotional reactions to events as well as slower, more deliberative responses. These observable behaviors represent the outcome of reactive and deliberative processes related to the interpretation of the Robot's relationship to the collaborative environment. These reactive and deliberative processes are triggered by two types of events: *external* events, such as the human's utterances and primitive actions, and *internal* events, comprising changes in the Robot's mental state, such as belief formation and emotional changes. Affective Motivational Collaboration Theory explains how emotions regulate the underlying processes in the occurrence of these events during collaboration.

Emotion-regulated processes operate based on the Robot's mental state, which also includes the anticipated mental state of the human, generated according to the Robot's model of the human. These mental states include beliefs, intentions, goals, motives and emotion instances. Each of these mental states possess multiple attributes impacting the relation between cognition and behavior or perception.

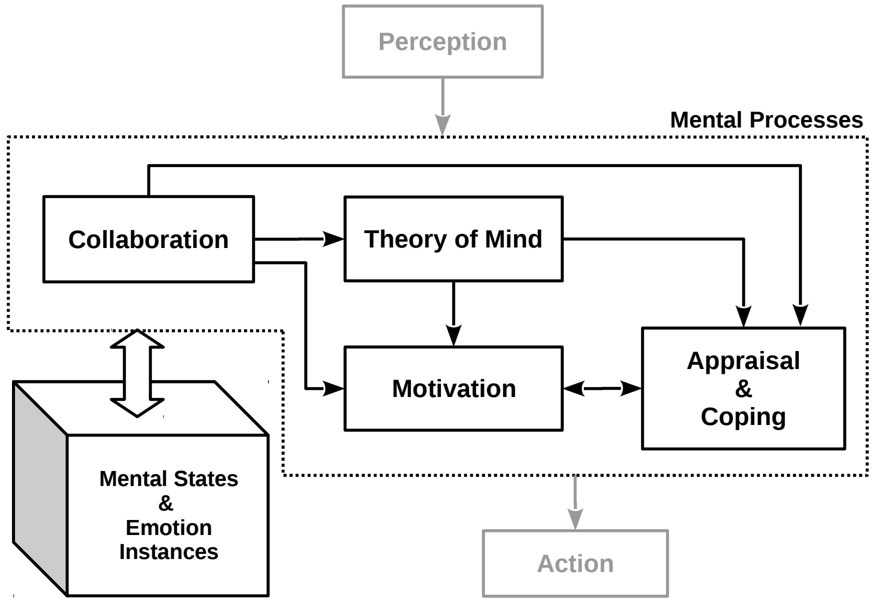

Fig. 1 Computational framework based on *Affective Motivational Collaboration Theory* (arrows indicate primary influences between mechanisms).

In summary, Affective Motivational Collaboration Theory consists of five mechanisms all of which store and retrieve data in the Mental States. We will describe each mechanism and their influences on each other briefly below.

4.2 Collaboration Mechanism

The *Collaboration* mechanism (see Fig. 1) constructs a hierarchy of tasks and also manages and maintains the constraints and other required details of the collaboration specified by the plan. These constraints on task states and on the ordering of tasks include the inputs and outputs of individual tasks, the preconditions specifying whether it is appropriate to perform a task (which can be used as an indication of an impasse), and the postconditions specifying whether a just-completed task was successful (or failed). The Collaboration mechanism includes processes to update and monitor the shared plan. It also keeps track of the focus of attention, which specifies the salient objects, properties and relations at each point of the collaboration. These processes depend on the operation of other mechanisms. For instance, the Appraisal mechanism is required to evaluate the current mental state with respect to the current status of the collaboration. Also, the Appraisal and Motivation mechanisms provide interpretation of task failure and the formation of a new mental state (e.g. an intention) respectively.

4.3 Appraisal & Coping Mechanisms

Appraisal is a subjective evaluation mechanism based on individual processes each of which computes the value of the appraisal variables. The Appraisal mechanism is responsible for evaluating changes in the Robot's mental state, the anticipated mental state of the human, and the state of the collaboration environment. Collaboration requires the evaluative function of the Appraisal mechanism for various reasons. The course of a collaboration is based on a full or a partial plan [21, 22] which needs to be updated as time passes and collaborators achieve, fail at or abandon a task assigned to them. The failure of a task should not destroy the entire collaboration. Appraising the environment and the current event helps the Robot to update the collaboration plan in response to changes in the environment and avoid further critical failures during collaboration. Appraisal also helps the Robot to have a better understanding of the human's actions by making inferences based on appraisal variables (see Section **??** for some examples) [38] [57]. Furthermore, in order to collaborate successfully, a collaborator cannot simply use the plan and reach to the shared goal; there should be an adaptation mechanism not only for updating the plan but also the underlying mental state. The output of Appraisal can directly and indirectly impact other mechanisms. For instance, the Motivation mechanism uses this data to generate, compare and monitor motives based on the current internal appraisal of the Robot as well as the appraisal of the environment.

The Coping mechanism is responsible for adopting the appropriate behavior (action) with respect to interpretation of the ongoing internal and external changes. The Coping mechanism provides the Robot with different coping strategies associated with changes in the Robot's mental state with respect to the state of the collaboration. In other words, the Coping mechanism produces cognitive responses based on the appraisal patterns.

4.4 Motivation Mechanism

The *Motivation* mechanism operates whenever the Robot a) requires a new motive to overcome an internal impasse in an ongoing task, or b) wants to provide an external motive to the human when the human faces a problem in a task. In both cases, the Motivation mechanism uses the Appraisal mechanism to compute attributes of the competing motives. The purpose of Motivation mechanism in Affective Motivational Collaboration Theory is to generate new emotion-driven goal-directed motives considered as "potential" intentions. These motives are generated based on what the Robot believes about the environment including the Robot and the other collaborator and the corresponding appraisals. The Robot uses these motives to reach to a private or shared goal according to new conditions caused by changes in the environment. The Motivation mechanism consists of an arrangement of three distinct processes. First, several motives are generated with respect to the current mental state. Only one of these competing motives is most likely to become a new intention. Therefore, a comparison process decides which motive is more likely to be consistent with the current state based on the values of the motive attributes (e.g., motive insistence and motive urgency). Finally, the new motive will be used to form a new intention. As a result, the Robot can take an action based on the new intention to sustain the collaboration progress. Furthermore, the Motivation mechanism can serve the Theory of Mind mechanism by helping the Robot to infer the motive behind the human's current action.

4.5 Theory of Mind Mechanism

The *Theory of Mind* mechanism is the mechanism for inferring a model of the human's anticipated mental state. The Robot uses the Theory of Mind mechanism to infer and attribute beliefs, intentions, motives and goals to its collaborator based on the user model it creates and maintains during collaboration. The Robot progressively updates this model during the collaboration. The refinement of this model helps the Robot to anticipate the human's mental state more accurately, which ultimately impacts the quality of the collaboration and the achievement of the shared goal. Furthermore, the Robot can make inferences about the motive (or intention) behind the human's actions using the Motivation mechanism. This inference helps the Robot to update its own beliefs about the human's mental state. In the reverse appraisal process [12], the Robot also applies the Appraisal mechanism together with updated beliefs about the human's Mental States to infer the human's current mental state based on the human's emotional expression. Finally, the Collaboration mechanism provides the collaboration structure, including status of the shared plan with respect to the shared goal and the mutual beliefs to the Theory of Mind mechanism. Consequently, any change to the Robot's model of the human will update the Robot's mental state.

4.6 Perception & Action

Perception is outside of our theory and is responsible for producing the sensory information used by the mechanisms in our framework; it is only a source of data to the computational framework (see Fig. 1). Thus, our computational framework starts with high-level semantic representation of events (including utterances). The output of the Perception component provides a unified perception representation across all of the mechanisms.

The Action component in Fig. 1, which is also outside of our theory, functions whenever the Robot needs to show a proper behavior according to the result of the internal processes of the collaboration procedure; it is only a sink of data in our computational framework. The only input to the Action component is provided by the Coping mechanism. This input will cause the Action component to execute an appropriate behavior of the Robot. This input to Action has the same level of abstraction as the output of the Perception mechanism, i.e., it includes the Robot's utterances, primitive actions and emotional expressions.

4.7 Mental States & Emotion Instances

The Mental States shown in Fig. 1 comprise the knowledge base required for all the mechanisms in the overall framework.

4.7.1 Beliefs

Beliefs are a crucial part of the Mental States. We have two different perspectives on categorization of beliefs. In one perspective, we categorize beliefs based on whether or not they are shared between the collaborators. The SharedPlans [23] theory is the foundation of this categorization in which for any given proposition the Robot may have: a) private beliefs (the Robot believes the human does not know these), b) the inferred beliefs of the human (the Robot believes the human collaborator has these beliefs), and c) mutual beliefs (the Robot believes both the Robot and the human have these same beliefs and both of them believe that). From another perspective, we categorize beliefs based on who or what they are about. In this categorization, beliefs can be about the Robot, the human, or the environment. Beliefs about the environment can be about internal events, such as outcomes of a new appraisal or a new motive, or external events such as the human's offer, question or request, and general beliefs about the environment in which the Robot is situated. Beliefs can be created and updated by different processes. They also affect how these processes function as time passes.

4.7.2 Intentions

Intentions are mental constructs directed at future actions. They play an essential role in: a) taking actions according to the collaboration plan, b) coordination of actions with the human collaborator, c) formation of beliefs about the Robot and anticipated beliefs about the human, and d) behavior selection in the Coping mechanism. First, taking actions means that the Robot will intend to take an action for primitive tasks that have gained the focus of attention, possess active motives, and have satisfied preconditions for which required temporal predecessors have been successfully achieved. Second, intentions are involved in action coordinations in which the human's behavior guides the Robot to infer an anticipated behavior of the human. Third, intentions play a role in belief formation, mainly as a result of the permanence and commitment inherent to intentions in subsequent processes, e.g., appraisal of the human's reaction to the current action and self-regulation. Lastly, intentions are involved in selecting intention-related strategies, e.g., planning, seeking instrumental support and procrastination, which is an essential category of the strategies in the Coping mechanism [38]. Intentions possess a set of attributes, e.g. *involvement, certainty, ambivalence* which moderate the consistency between intention and behavior. The issue of consistency between the intentions (in collaboration) and the behaviors (as a result of the Coping mechanism in the appraisal cycle) is important because neither of these two mechanisms alone provides solution for this concern.

4.7.3 Motives

Motives are emotion-driven goal-directed mental constructs which can initiate, direct and maintain goal-directed behaviors. They are created by the emotion-regulated Motivation mechanism. Motives can cause the formation of a new intention for the Robot according to: a) its own emotional states (how the Robot feels about something), b) its own private goal (how an action helps the Robot to make progress), c) the collaboration goal (how an action helps to achieve the shared goal), and d) the human's anticipated beliefs (how an action helps the human). Motives also possess a set of attributes, e.g., *insistence* or *failure disruptiveness*. These attributes are involved in the comparison of newly generated motives based on the current state of the collaboration. Ultimately, the Robot forms or updates an intention about the winning motive in the Mental States.

4.7.4 Goals

Goals help the Robot to create and update the structure of the collaboration plan. Goals direct the formation of intentions to take appropriate corresponding actions during collaboration. Goals also drive the Motivation mechanism to generate required motive(s) in uncertain or ambiguous situations, e.g., to minimize the risk of impasse or to reprioritize goals. Goals have three attributes. The *specificity* of goals has two functions for the Robot. First, it defines the performance standard for evaluating the progress and quality of the collaboration. Second, it serves the Robot to infer the winner of competing motives. The *proximity* of goals distinguishes goals according to how "far" they are from the ongoing task. Proximal (or short-term) goals are achievable more quickly, and result in higher motivation and better self-regulation than more temporally distant (or long-term) goals. Goals can influence the *strength* of beliefs, which is an important attribute for regulating the elicitation of social emotions. The *Difficulty* of goals impacts collaborative events and decisions in the appraisal, reverse appraisal, motive generation and intention formation processes. For instance, overly easy goals do not motivate; neither are humans motivated to attempt what they believe are impossible goals.

4.7.5 Emotions

Emotions in Mental States are emotion instances that are elicited by the Appraisal mechanism. These emotion instances include the Robot's own emotions as well as the anticipated emotions of the human which are created with the help of the processes in the Theory of Mind mechanism.

5 Collaboration

The Collaboration mechanism constructs a hierarchy of goals associated with tasks in the form of a hierarchical task network (see Figure 2), and also manages and maintains the constraints and other required details of the collaboration including the inputs and outputs of individual tasks, the *preconditions* (specifying whether it is appropriate to perform a task), and the *postconditions* (specifying whether a just-completed task was successful). Collaboration also keeps track of the focus of attention, which determines the salient objects, properties and relations at each point, and shifts the focus of attention during the interaction.

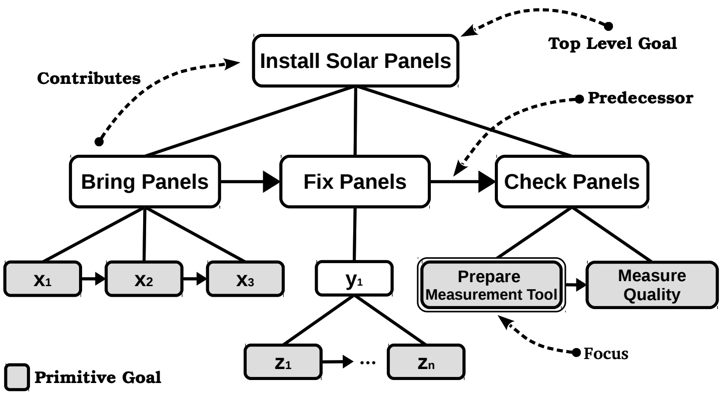

Fig. 2 Collaboration structure (shared plan).

Here, we briefly describe the methods which retrieve information about the collaboration structure, and are used in our algorithms to compute the values of appraisal variables. In these methods, ε_t is the event corresponding to time t, and g_t is a given goal at time t.

- *recognizeGoal(ε_t)* returns the unique goal to which the given event (action, utterance, or emotional expression) directly contributes, or *ambiguous* if this method does not recognize a goal in the plan.
- *topLevelGoalStatus(g_t)* returns the status of the top level goal whether it is ACHIEVED, FAILED, BLOCKED, INAPPLICABLE, PENDING, or IN PROGRESS. In our example, "Install Solar Panels" is the top level goal.

- *currGoalStatus(g_t)* returns the current goal status whether it is ACHIEVED, FAILED, BLOCKED, INAPPLICABLE, PENDING, or IN PROGRESS. In our example, "Prepare Measurement Tool" is the current (focused) goal.
- *precondStatus(g_t)* returns the status of the precondition for the given goal whether it is SATISFIED, UNSATISFIED or UNKNOWN. For instance, the precondition for fixing a panel is whether the panel is appropriately located on its frame.
- *isLive(g_t)* returns *true* if all the predecessors of the given goal are ACHIEVED and all the preconditions of the goal are SATISFIED; otherwise returns *false*.
- *isFocusShift(g_t)* returns *true* if the given goal is not the previous focus (top of the stack); otherwise returns *false*.
- *isNecessaryFocusShift(g_t)* returns *true* if the status of the previous focus was ACHIEVED; otherwise returns *false* [**?**].
- *isPath(g_1, g_2)* returns *true* if there is a path between g_1 and g_2 in a plan tree structure; otherwise returns *false*.
- *doesContribute(g_t)* returns whether the given goal contributes to another goal in the higher level of the plan hierarchy. For instance, an abstract (nonprimitive) goal of "Bring Panels" contributes to the higher level goal of "Install Solar Panels".
- *extractContributingGoals(g_t)* returns all the contributing goals of the given goal. For instance, "Prepare Measurement Tool" and "Measure Quality" are two goals contributing to the "Check Panels" nonprimitive goal.
- *extractPredecessors(g_t)* returns the predecessors of the given goal. For instance, the "Fix Panels" goal is the predecessor of another goal called "Check Panels".
- *extractInputs(g_t)* returns all the required inputs for the given goal. For example, the goal "Fix Panels" requires inputs such as the *welding tool* and the *panel*.
- *isAvailable(g_t)* returns whether the given input is available. For instance, if the *welding tool* is required for the goal "Fix Panels", is it available now?
- *isAchieved(g_t)* returns whether the given goal is achieved, i.e., whether all the postconditions of the given goal are SATISFIED.
- *isFocused(g_t)* returns whether the focus is on given goal now. In this example, the focus is on the goal "Prepare Measurement Tool". The focused goal is the goal that the robot is currently pursuing.
- *getResponsible(g_t)* returns responsible agents of the given goal. In a dyadic collaboration, both of the agents can be partly responsible for a nonprimitive goal, while each is responsible for one or more primitive goals. For instance, both the robot and the astronaut are responsible for the nonprimitive goal of "Install Solar Panels", whereas it is only the astronaut who is responsible for the primitive goal of "Prepare Measurement Tool".

6 Appraisal Processes

We consider four appraisal variables to be the most important appraisal variables in a collaboration context, i.e., *Relevance* (Algorithm 1), *Desirability* (Algorithm 2), *Expectedness* (Algorithm 3), and *Controllability* (Algorithm 4). There are other appraisal variables introduced in psychological [57] and computational literature [17]. We believe most of these variables can be straightforwardly added to our appraisal mechanism later. All of the algorithms in this section use mental states of the robot (discussed in Section **??**) which are formed based on the collaboration structure. These algorithms use the corresponding recognized goal of the most recent event at each turn.

6.1 Relevance

Relevance as an appraisal variable measures the significance of an event for the robot. An event can be evaluated to be relevant if it has a positive utility or it can causally impact a state with a positive utility [38]. Relevance is an important appraisal variable since the other appraisal variables are more meaningful only for relevant events.

Algorithm 1 determines the relevance of the given event with respect to the current mental state. The relevance of the event depends on the significance of the event with respect to the current collaboration status. The significance of an event is determined based on the utility of the event as it is also presented in [17,38]. We believe although the utility of the event represents the significance of the event, the other collaborator's expressed emotion also plays a role by influencing the significance of the utility through a threshold value. As a result, evaluating the relevance of the events can cause a collaborative robot to respond effectively to the events which can positively impact the status of the shared goal, without dedicating all resources to every single event. The relevance process also benefits from the information that the collaboration structure contains, e.g., shared goal.

Algorithm 1 (Relevance)

```
function IsEventRelevant(Event ε_t)

    g_t ← recognizeGoal(ε_t)

    U ← getEventUtility(g_t)
    τ_t ← getEmotionalThreshold(g_t)

    if (U ≥ τ_t) then
        return RELEVANT
    else
        return IRRELEVANT
```

After perceiving an event, it is the belief about that event which represents the event in the robot's mental state. Also, *recognizeGoal* returns the goal (g_t) to which the current event contributes, unless it is *ambiguous*; g_t represents the shared goal at time (turn) t within the shared plan. We compute the utility ($0 \leq \mathcal{U} \leq 1$) of the event based on the values of the attributes associated with the existing beliefs in the mental state, as well as the attributes of the motive associated with the recognized goal. We use three of the belief attributes discussed in Section **??** to compute the belief related part of the utility:

– *Strength*: The extent to which the pre and postconditions of a goal and its predecessors and/or contributing goals are SATISFIED or UNSATISFIED makes a belief about the goal stronger. Respectively, an UNKNOWN pre and postcondition status of a goal and its predecessors and/or contributing goals forms beliefs with lower strength.
– *Saliency*: Beliefs related to the goal at the top of the focus stack are more salient than beliefs related to any other goal in the plan, whether those goals are already ACHIEVED or FAILED, or they will be pursued in the future.
– *Persistence*: The recurrence of a belief over the passage of time (turns) increases the persistence of the belief. Beliefs occurring only in one turn have the lowest value of persistence.

We also use two of the motive attributes discussed in Section **??** to compute the motive related part of the utility ($\mathcal{U}$):

– *Urgency*: There are two factors impacting the urgency of a motive: a) whether the goal directing the given motive is the predecessor of another goal for which the other collaborator is responsible, and b) whether achieving the goal directing the given motive can mitigate the other collaborator's negative valenced emotion.
– *Importance*: A motive is important if failure of the directing goal causes an impasse in the shared plan (i.e., no further goal is available to achieve), or achievement of the directing goal removes an existing impasse.

We compute the utility of an event based on these five attributes. The value of each attribute is between 0 and 1, and we consider the same weight for each attribute. These weights can be learned or modified when our framework is fully implemented. The value of the overall utility is computed using a simple weighted averaging function which results in an overall value between 0 and 1.

The significance of an event in a collaborative environment is based not only on the utility of the event, but it is also influenced by the perceived emotion of the human collaborator. The human's emotion influences the decision about the utility of the event in the form of a threshold value τ_t (see Algorithm 1). For instance, a positively expressed emotion of the human reduces the threshold value which consequently makes the robot find an event relevant with even a slightly positive utility. This threshold value (τ_t) is currently determined based on whether the valence of the human's perceived emotion is positive (e.g., happiness) or negative (e.g., anger). Consequently, an event can be considered

IRRELEVANT even though the utility has a relatively positive value, because relevance is influenced by the human's perceived emotional state.

6.2 Desirability

Desirability characterizes the value of an event to the robot in terms of whether the event facilitates or thwarts the collaboration goal. Desirability captures the valence of an event with respect to the robot's preferences [17]. In a collaborative robot, preferences are biased towards those events facilitating progress in the collaboration. Desirability plays an important role in the overall architecture; it makes the processes involved in the other mechanisms (e.g., Motivation and Theory of Mind), and consequently the robot's mental state, congruent with the collaboration status which is a collaborative robot's desire. Therefore, it causes the robot to dismiss events causing inconsistencies in the robot's collaborative behavior. Moreover, desirability is also crucial from the collaboration's point of view.

Algorithm 2 provides a process in which the desirability of an event is computed with regard to the status of the shared goal; i.e., it operates based on whether and how the event changes the status of the current shared goal. It distinguishes between the top level goal and the current goal because the top level goal's change of status attains a higher positive or negative value of desirability. For instance, failure of the top level goal (e.g., installing solar panel) is more undesirable than failure of a primitive goal (e.g., measuring the quality of the installed panel).

An AMBIGUOUS goal is a goal associated with the current event (ε_t) which is not recognized in the robot's plan; therefore it is UNDESIRABLE for a collaborative robot. A top level goal' status must be ACHIEVED (i.e., SATISFIED postcondition) to consider the event MOST-DESIRABLE. When the goal's status is FAILED (i.e., UNSATISFIED postcondition) or BLOCKED, the associated event has the MOST-UNDESIRABLE or UNDESIRABLE values respectively. A goal is BLOCKED if any of the required goals or goals recursively through the parent goal are not ACHIEVED. An INAPPLICABLE goal is also considered as UNDESIRABLE. A goal is INAPPLICABLE if any of its predecessors are not ACHIEVED, and/or its preconditions are not SATISFIED. For PENDING and INPROGRESS top level goals, the status of the current goal associated with the top level goal determines the status of the event ε_t. Only a non-primitive goal can have INPROGRESS status, if it has been started but is not yet completed. A goal can be PENDING if it is live, or if it is a non-primitive goal that has not been started yet. ACHIEVED current goals mark an event (ε_t) as DESIRABLE, while FAILED or BLOCKED current goals render the event associated with them as MOST-UNDESIRABLE and UNDESIRABLE respectively. PENDING or INPROGRESS current goals mark their associated events as NEUTRAL.

Algorithm 2 (Desirability)

```
function IsEventDesirable(Event ε_t)

    g_t ← recognizeGoal(ε_t)

    if (g_t = AMBIGUOUS) then
        return UNDESIRABLE

    if (topLevelGoalStatus(g_t) = ACHIEVED) then
        return MOST-DESIRABLE
    else if (topLevelGoalStatus(g_t) = FAILED) then
        return MOST-UNDESIRABLE
    else if (topLevelGoalStatus(g_t) = BLOCKED)
    (topLevelGoalStatus(g_t) = INAPPLICABLE) then
        return UNDESIRABLE
    else if (topLevelGoalStatus(g_t) = PENDING)
    (topLevelGoalStatus(g_t) = INPROGRESS) then

        if (currGoalStatus(g_t) = ACHIEVED) then
            return DESIRABLE
        else if (currGoalStatus(g_t) = FAILED) then
            return MOST-UNDESIRABLE
        else if (currGoalStatus(g_t) = BLOCKED)
        (topLevelGoalStatus(g_t) = INAPPLICABLE) then
            return UNDESIRABLE
        else if (topLevelGoalStatus(g_t) = PENDING)
        (currGoalStatus(g_t) = INPROGRESS) then
            return NEUTRAL
```

6.3 Expectedness

Expectedness is the extent to which the truth value of a state could have been predicted from causal interpretation of an event [38]. In the collaboration context the expectedness of an event evaluates the congruency of the event with respect to the existing knowledge about the shared goal. Thus, expectedness underlies a collaborative robot's attention. Congruent beliefs in a robot's mental state will lead to more consistent and effective outcomes of the processes in the overall architecture. The collaboration mechanism uses expectedness to maintain the robot's attention and subsequently its mental state with respect to the shared goal. Reciprocally, the appraisal mechanism uses the underlying information of the collaboration structure to evaluate the expectedness of an event. Therefore, a collaborative robot uses expectedness to maintain its own mental state towards the shared goal. The robot will also be able to respond to unexpected but relevant events.

Algorithm 3 (Expectedness)

```
function IsEventExpected(Event ε_t)

    g_t ← recognizeGoal(ε_t)
    g_top ← getTopLevelGoal(g_t)

    if (isLive(g_t)) then
        if (¬isFocusShift(g_t)
          isNeccessaryFocusShift(g_t)) then
            return MOST-EXPECTED
        else
            return EXPECTED
    else
        if (isPath(g_t, g_top)) then
            return UNEXPECTED
        else
            return MOST-UNEXPECTED
```

In Algorithm 3 we provide the process of computing the expectedness based on the shared plan and status of the shared goal. The key point in this algorithm is the status of the current shared goal (g_t) that is associated with the event ε_t and its relationship with the top level goal (g_{top}).

The intuition captured here is that one expects the current goal to be finished before undertaking another activity, but the goals that are the next focus of attention are also to be expected [**?**]. Therefore, if the goal is live, the algorithm checks whether the goal has not changed, or the interpretation of the last event results in a necessary focus shift. Shifting the focus to a new goal is necessary when the former goal is achieved and a new goal is required. Consequently the new event is the MOST-EXPECTED one. However, even if the focus shift is not necessary, the new event can be considered as EXPECTED, since the corresponding goal is already live. For goals that have not yet been started (that is, are not live), the algorithm must determine how unexpected it would be to pursue one now; if the goal is at least in the plan, i.e., on the path to the top level goal, it is just UNEXPECTED while any others are MOST-UNEXPECTED.

6.4 Controllability

Controllability is the extent to which an event can be influenced, and it is associated with a robot's ability to cope with an appraised event [17]. Thus, a robot can determine whether the outcome of an event can be altered by some actions under either of the collaborators' control. In other words, controllability is a measure of a robot's ability to maintain or change a particular state as a consequence of an event.

Algorithm 4 (Controllability)

1: **function** IsEventControllable(Event ε_t)

2: $\quad \alpha \leftarrow$ GetAgencyRatio(ε_t)
3: $\quad \beta \leftarrow$ GetAutonomyRatio(ε_t)

4: $\quad \lambda \leftarrow$ GetSucPredecessorsRatio(ε_t)
5: $\quad \mu \leftarrow$ GetAvailableInput(ε_t)

6: $\quad \mathcal{U} \leftarrow \frac{\omega_0 \cdot \alpha + \omega_1 \cdot \beta + \omega_2 \cdot \lambda + \omega_3 \cdot \mu}{\omega_0 + \omega_1 + \omega_2 + \omega_3}$

7: $\quad \tau_t \leftarrow$ getEmotionalThreshold()

8: $\quad$ **if** $(\mathcal{U} \geq \tau_t)$ **then**
9: $\qquad$ **return** controllable
10: $\quad$ **else**
11: $\qquad$ **return** uncontrollable

Controllability is also important for the overall architecture. For instance, the robot can choose to ask or negotiate about a collaborative task which is not controllable, or the robot can interpret or predict the other's emotional state (e.g., anger if the task is blocked, i.e., uncontrollable for the other), or form a new motive to establish an alternative goal for the current uncontrollable event. In general, other mechanisms in the architecture use the appraisal process of controllability in their decision making processes; meanwhile controllability uses the information from the collaboration structure, e.g., successful predecessors of a goal.

An important determinant of one's emotional response is the sense of control over the events occurring. This sense of subjective control is based on one's reasoning about self's power. For instance, the robustness of one's plan for executing actions can increase one's sense of power and subsequently the sense of control. In the collaboration context, we have translated the sense of control into a combination of four different factors including a) *agency* and b) *autonomy* of the robot, as well as the ratios of c) *successful predecessors*, and d) the *available inputs* of a given goal (i.e., g_t) in the shared plan.

In Algorithm 4, we compute the controllability of an event based on these four factors (lines 2 to 5). We use weighted averaging over these four factors to compute the utility of an event in terms of controllability of the event. The value of all these weights are set to 1.0 for the purpose of simplicity at this stage of the project. We will adjust these weights after further investigating the influence of these factors, and implementing other mechanisms in the overall architecture. After computing the value of the utility, we compare this value to an emotional threshold similar to what we discussed in Algorithm 1. This comparison leads to our decision about the controllability of an event (lines 8 to 11 in Algorithm 4).

Agency is the capacity of an individual to act independently in any given environment. In a collaborative environment collaborators are sometimes required to act independently of each other. Hence, they need to have some internal motives that are formed based on their own mental states rather than motives that are reinforced by the other collaborator. These internal motives will lead the collaborators to acquire new intentions towards new goals whenever it is required. We extract the motive associated with the current goal in the mental state. We consider a maximum agency value denoted as α in Algorithm 4 (i.e., $\alpha = 1.0$) if the robot's mental state possesses an internal motive towards the recognized goal; otherwise we consider the minimum agency value (i.e., $\alpha = 0.0$) for no motives or external motives only. Note that the process of forming new internal motives is beyond scope of this paper.

Autonomy is the ability to make decisions without the influence of others. Autonomy implies acting on one's own and being responsible for that. In a collaborative environment, tasks are delegated to the collaborators based on their capabilities. Therefore, each collaborator is responsible for the delegated task and the corresponding goal. In Algorithm 4, β denotes the value of autonomy with regard to the event (ε_t). This value is the ratio of the number of the goals contributing to g_t for which the robot is responsible over the total number of contributing goals to g_t. If the goal associated with the current event corresponds to a nonprimitive goal, the algorithm checks the responsible agent for each primitive goal contributing to the nonprimitive one and returns a value of which ($0 \leq \beta \leq 1$). However, if the associated goal of the current event corresponds to a primitive goal the value of β would be 0 or 1. In general, higher autonomy leads to a more positive value of controllability.

The structure of a shared plan accommodates the order of the required ***predecessors*** of a goal. Predecessors of a goal, g, are other goals that the collaborators should achieve before trying to achieve goal g. We use the ratio of successfully achieved predecessors of the recognized goal (g_t) associated with the current event over the total number of predecessors of the same goal. This ratio (denoted as λ in Algorithm 4) is the third factor used to compute the controllability of an event. If all of the predecessors of the given goal are already achieved, then $\lambda = 1$ which is the maximum value for λ. On the contrary, failure of all of the predecessors will lead to $\lambda = 0$. Therefore, a higher λ value positively impacts the value of controllability for the current event.

Finally, ***inputs*** of a task are the required elements that the collaborators use to achieve the specified goal of the task. These inputs are also part of the structure of a shared plan. We extract the required inputs of the associated goal with the current event, and check whether all the required inputs are available for the goal g_t. The outcome will be the ratio of the available required inputs over the total required inputs of the goal associated with the current event. This value (denoted as μ in Algorithm 4) will be bound to 0 and 1. Similar to the other factors in the controllability process, the closer the value of μ gets to 1, the more positive impact it has on the overall controllability value of the event.

In summary, the output of these four appraisal processes serves as critical input for the other mechanisms of the Affective Motivational Collaboration Framework, shown in Figure **??**. By providing adequate interpretation of events in the collaborative environment, the appraisal mechanism enables the robot to carry out proper collaborative behaviors.

7 Goal Management

In this work, we focus on a small part of a larger framework based on our *Affective Motivational Collaboration Theory* [**?**]. We introduce our goal management process based on a cost function including the influence of affective appraisal and reverse appraisal processes. Goal management is a crucial part of our investigation of the reciprocal influence of appraisal on a collaboration structure (see Fig. 3).

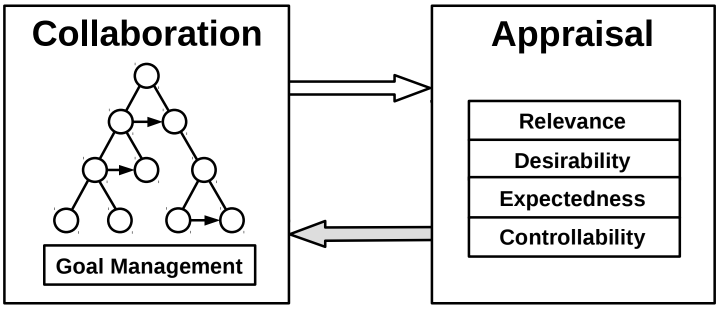

Fig. 3 Reciprocal influence of Collaboration and Appraisal (mechanisms in our framework).

Appraisals are separable antecedents of emotion with which the robot evaluates the environment. Our appraisal variables included: a) *relevance* to measure the significance of an event for the robot, b) *desirability* to characterize the value of an event to the robot in terms of whether the event facilitates or thwarts the collaboration goal, c) *expectedness* and d) *controllability* (these last two appraisal variables are beyond the scope of this paper). The outcome of each appraisal process is a specific value for the corresponding appraisal variable. The vector containing these appraisal variables can be mapped to a particular emotion instance at each point in time. However, it is not the actual emotion instance that is important for us. In fact, it is a) the functions of emotions in a social setting, i.e., goal management, and b) the meaning of the collaborator's perceived emotion in collaboration context.

A collaboration structure provides a hierarchy and constraints of the shared goals in the form of a shared plan (Fig. 5) which contains both the robot and the human collaborator's goals. The robot pursues the goals for which the robot is responsible in the shared plan. However, there can be several live goals available for the robot to pursue at each point in time during collaboration. A goal is *live* if all of its *predecessors* are achieved and all of its *preconditions* are satisfied. Therefore, a collaborative robot requires a mechanism to choose between a set of live goals. We believe appraisal processes are crucial to choose between the available live goals, since the appraisals are the immediate outcome of the robot's assessment of the collaboration environment.

For instance, Fig. 5 shows a non-primitive "Prepare Panels" goal decomposed into three primitive goals. Therefore, if "Prepare Panels" is live, its primitive goals can be pursued by the responsible agent. In our example, the astronaut is responsible for the "Check Connector" goal; the robot is responsible for the remaining two primitive goals. According to the collaboration mechanism in our overall framework, "Check Connector" is in focus, with the astronaut pursuing this goal. Suddenly, however the astronaut tells the robot that she can not find the connector and she is *worried* about failure of this goal. The robot's response to this situation will be explored below as we discuss details of our cost function.

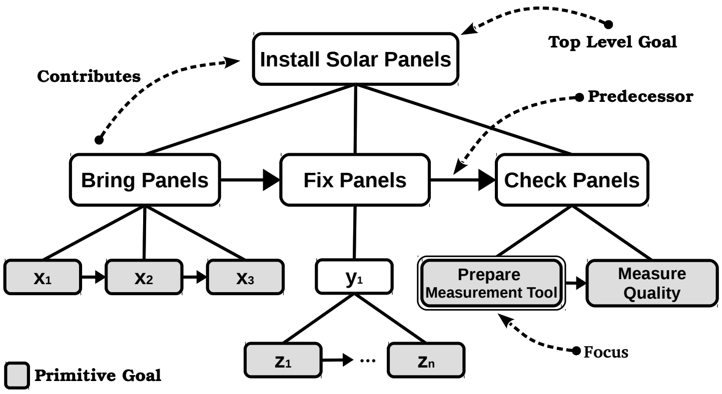

Fig. 4 Astronaut-robot collaboration structure (shared plan).

Equation 1 shows the function to calculate the cost of each live goal. The base in the equation calculates the cost of pursuing any given goal. The three functions used to calculate the cost are: *proximity* $P(g)$, *difficulty* $D(g)$, and *specificity* $S(g)$ (see equations 3 to 5).

$$Cost(g) = \Big(\omega_0.P(g) + \omega_1.D(g) + \omega_2.\frac{1}{S(g)+1}\Big)^{\Gamma} \tag{1}$$

For simplicity, we assume equal values for the weights: ω_i=1.

$$\Gamma = -C[(R_r + 1)D_r + \alpha(R_h + 1)D_h] \tag{2}$$

The exponent part of our cost function (Equation 2) captures a) the influence of the human's perceived emotional instance, and b) the influence of self appraisal of the given goal. $R_h \in [0, 1]$ and $D_h \in [-1, 1]$ are the relevance and desirability values respectively, which are based on the *reverse* appraisal of the human's perceived emotion. For instance, if the astronaut is *worried*, D_h is negative, e.g., -0.8 (depending on how undesirable the event is according to reverse appraisal), and R_h will be 1 for the active goal and its value descends to 0 for other live goals depending on their distance to the active goal in the shared plan (e.g., 0.1).

$R_r \in [0, 1]$ and $D_r \in [-1, 1]$ are relevance and desirability values, provided by the *self* appraisal functions for all of the live goals. For instance, for the active goal for which the astronaut was *worried*, D_r can be positive, e.g., 0.8 (depending on the self's desirability appraisal function); R_r can be 1, since the active goal is relevant for the robot. These values will change for the other live goals depending on how relevant they are with respect to the collaboration status (e.g., 0.9 and 0.8). Finally, $C \in [1, \infty)$ is a constant (e.g., 2) used to control the influence of affect on cost value. It is negative since undesirability (negative values) should increase the cost. $\alpha \in [1, \infty)$ is another constant (e.g., 3) used to control the importance of reverse appraisal relative to self appraisal.

The *proximity* of a goal indicates how far the goal is from the current active goal in the shared plan. It is calculated by the distance function (Equation 3) which returns the number of edges between the current active goal $g_{_{act}}$, and the given goal g in the shared plan. In our example, $P(g)$ is 2 for both "Check Impedance" and "Connect Adaptor" goals.

$$P(g) = max\{1, distance(g_{_{act}}, g)\} \tag{3}$$

The *difficulty* of a goal is a function of three parameters (Equation 4) which consider the difficulty based on a) topology of the shared plan tree (domain independent), and b) the amount of effort required to pursue a given goal (domain dependent). The $\sum pred_e(g)$ is the sum of efforts that all the *predecessors* of a given goal g require. The $\sum desc_e(g)$ is the sum of efforts that all the *descendants* of a given goal g require. The effort values represent the amount of effort for the goals with respect to the domain. In our example, we assume the values of all the goal efforts are 1 for simplicity. The $H(g)$ is the height of the given goal g. The heights of all primitives under "Prepare Panel" goal are 0 in our example.

$$D(g) = \Big(H(g) + 1\Big) \times \left[\sum_{m=0}^{M} pred_e(g) + \sum_{n=0}^{N} desc_e(g)\right] \tag{4}$$

The *specificity* of a goal is the function of *depth* (distance from the root) and *degree* (number of children in the graph) of a given goal g. The first non-primitive goal (root) is the least specific goal, and the primitives (leaves) are

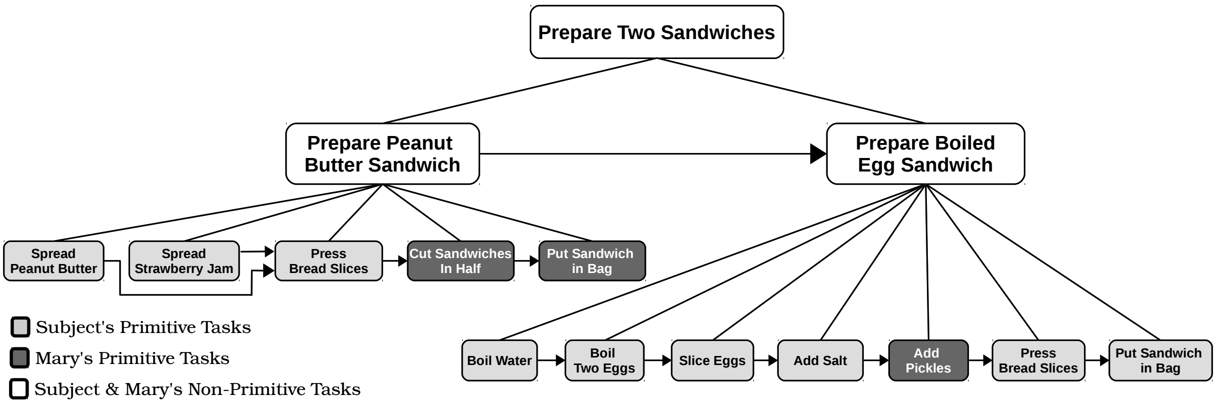

Fig. 5 Collaboration Task Model for the Evaluation.

the most specific goals. As calculated based on Fig. 5, the values of $S(g)$ for the three primitives under the "Prepare Panels" are 2.

$$S(g) = \frac{depth(g)}{degree(g) + 1} \tag{5}$$

The tuples below the goals in Fig. 5 indicate the cost value of each goal. The first number in each tuple is the normalized cost value without the influence of the affective part of the cost function, i.e., the exponent is equal to 1 in Equation 1. The second number of each tuple indicates the normalized value of the cost including the influence of affective appraisal and the astronaut's perceived emotion.

Based on our cost function, the cost of completing the primitive goal "Check Connector" is 0.82 (see Fig. 5). As shown, when affect is not considered the cost is 0.26; the negative emotion of the astronaut (worry) significantly increases the cost of the current goal, and also impacts the other two primitive live goals under the same parent. Therefore, instead of insisting on pursuing the same blocked goal which has caused the astronaut's negative emotion, the robot can mitigate the astronaut's emotions by adapting to her worry. The robot shifts the focus of attention to "Check Impedance" to maintain progress and prevent failure of the collaboration. The details about the robot's behavior is beyond the scope of this paper.

8 Evaluation

We developed our user study to test our hypothesis that humans will provide similar answers as our algorithms to questions related to different factors used to compute four appraisal variables. We conducted a between subject user study using an online crowdsourcing website – CrowdFlower[1]. We had one

[1] http://www.crowdflower.com

group of subjects for each questionnaire corresponding to an appraisal variable. There were 12 questions (including 2 test questions) in the controllability and expectedness questionnaires, 14 questions (including 2 test questions) in the desirability questionnaire, and 22 questions (including 3 test questions) in the relevance questionnaire. Each group originally had 40 subjects. To increase the quality of our subjects' answers, we limited the visibility of our questionnaires to a few English speaking countries, i.e., United States, Britain, and Australia. We also limited our subject pools to those that have acquired the highest confidence level on the crowdsourcing website. Our questionnaires included 2 or 3 test questions (depending on the length) to check the sanity of the answers. We eliminated subjects providing wrong answers to our sanity questions. We also eliminated subjects with an answering time less than 2 minutes. The final number of accepted subjects in each group is provided in Table 1.

Table 1 Evaluation Results

appraisal variables	# of subjects	mean	stdev	p-value
Relevance	29	0.713	0.107	¡0.001
Desirability	35	0.778	0.150	¡0.001
Expectedness	33	0.785	0.120	¡0.001
Controllability	33	0.743	0.158	¡0.001

To minimize the background knowledge necessary for our test subjects, we used a simple domestic example of preparing a peanut butter and jelly sandwich, and a hard boiled egg sandwich for a hiking trip. We provided clear textual and graphical instructions for all four questionnaires. The instructions presented a sequence of hypothetical collaborative tasks to be carried out by the test subject and an imaginary friend, Mary, in order to accomplish their goal of preparing two sandwiches. Figure 5 shows the corresponding task model for these instructions. Test questions introduced specific situations related to the shared plan; these situations included, among others, blocked tasks, and failure or achievement of a shared goal provided in the instruction. Each question provided three possible answers (which were counterbalanced in the questionnaire). One option provided a distinct alternative; another option was used to provide a dichotomy with the first alternative, and a third option was used to check whether the subjects perceived the other two options as equal. We also provided a brief description as well as a simple example for each appraisal variable, e.g., *relevance*, at the end of the corresponding instructions. Using this approach, we prepared four different online questionnaires for the appraisal variables: *relevance, desirability, expectedness* and *controllability.* Note that the collaboration structure and the instructions were the same for all four questionnaires.

Each question was designed based on different factors that we use in our algorithms (see Section **??**). Here, we present three example questions from the expectedness, controllability, and desirability questionnaires, and describe how each question relates to a specific factor within the corresponding algorithm. The input for our algorithms was the task model depicted in Figure 5.

Imagine you have pressed the two slices of bread (one covered with strawberry jam and one covered with peanut butter) together and passed it to Mary. Which of the following two actions is **more expected**?

A. Mary puts the given sandwich into a zip lock bag after cutting it in half.

B. Mary puts some pickles on another slice of bread.

C. Equally expected.

Fig. 6 Example Expectedness Question.

Figure 6 shows the example question from the expectedness questionnaire. In this example, with respect to Algorithm 3 (line 6), option A is more expected because the task related to this option provides the next available task in the focus stack (see the task model in Figure 5). Although the task in option B is part of the existing task model, it is considered as unexpected by our algorithm, since it is not live in the plan. We provided option C to determine whether the human subjects will similarly differentiate between these two options. This question was presented to the human subjects to determine whether their decision for the expectedness of this event is similar to the output of the expectedness algorithm. For this question, the human decision was 97% similar to the algorithm's output. Average results for the expectedness questionnaire are presented in Table 1.

Imagine you want to make a peanut butter sandwich. Which of the following two actions is **more controllable**?

A. You can spread the peanut butter on one slice of bread and you need Mary to spread strawberry jam on the second slice of bread.

B. You can spread the peanut butter on one slice of bread and strawberry jam on the second slice of bread.

C. Equally controllable.

Fig. 7 Example Controllability Question.

Figure 7 shows an example question from the controllability questionnaire. The algorithm's output is option B, and is determined by Algorithm 4 (line 3), similarly to the expectedness example above. In this example, option B is more controllable than option A, because the self over total ratio of the responsibility of the predecessors of the given task (see *Autonomy* in Section 6.4) is higher than the ratio in option A; i.e., self is responsible to spread peanut butter on one slice of bread and strawberry jam on another slice of bread. In this question, the humans decision was 90% in agreement with the algorithm's output.

Figure 8 shows an example question from the desirability questionnaire. The output based on the Algorithm 2 (line 14) is option C, since in both option A and option B, the focus goal has been achieved successfully. Therefore, in

Which of the following two actions is **more desirable**?

A. Imagine you pressed two slices of bread together with peanut butter and strawberry jam on them, and passed them to Mary. Mary cuts the peanut butter sandwich in half and puts them in the zip lock bag.

B. Imagine you want to make the egg sandwich. You have sliced the eggs, put them on one slice of bread, salted them, and waiting for Mary to put some pickles on your eggs. Mary puts some pickles on your eggs.

C. Equally desirable.

Fig. 8 Example Desirability Question.

this example, both options A and B are desirable. The humans decision was 77% in agreement with the algorithm's output in this question.

We conducted the user study to compare the results with the implemented algorithms discussed in Section **??**. As we mentioned, each question had 3 answers. Therefore, a totally random distribution would result in 33% agreement with our algorithms results. However, the average ratio indicating similarity between human subjects decisions and the output of our algorithms is significantly higher than 33%. The total number of subjects' answers similar to a) the *relevance* algorithm (n=29) averaged 71.3% (s=10.7%), b) the *desirability* algorithm (n=35) averaged 77.8% (s=15.0%), c) the *expectedness* algorithm (n=33) averaged 78.5% (s=12.0%), and d) the *controllability* algorithm (n=33) averaged 74.3% (s=15.8%). It is worth noting that the human subjects agreed 100% on some questions, while one some other questions there was a much lower level of agreement.

The results indicate that our algorithms provide appraisal variable outputs sufficiently similar to the decisions of human's appraisal. Our hypothesis in our evaluation was that our algorithms would correctly predict the judgements of humans on doing these tasks. Our results indicate that people largely performed as our hypothesis predicted. The p-values obtained based on a one-tailed z-test (see Table 1) show the probability of human subjects' data being generated from a random set. The very small p-values indicate that the data set is not random; in fact, the high percentage of similarity shows that the four appraisal algorithms predicted the human judgments.

9 Conclusion and Future Work

There is a correspondence between what a collaboration needs and the social functions of emotions. In this paper, we presented a theory explaining the processes underlying in collaboration using social emotions. We provided four hypothetical examples in two pairs, each dealing with an important collaborative behavior. The first pair was about agreeing on a shared goal; the second pair was about delegation of a new task. Each pair of examples contrasted a successful collaboration, due to the Robot's awareness of the Astronaut's emotion, with a failure in collaboration, due to the Robot's ignorance of the

Astronaut's emotions. These examples illustrated the importance of emotional awareness to attain successful collaborative behavior.

We then introduced the main components of Affective Motivational Collaboration Theory, our computational framework which integrates emotion-regulated mechanisms, such as appraisal and coping, with collaboration processes, such as planning, in a single unified framework. This framework let us describe the same examples in more computational detail.

We have started to implement the rules associated with these computational walkthroughs using JESS (Java Expert System Shell) which is a rule engine for the Java platform. In our current implementation we have categorized the rules into different modules associated with the mechanisms and the processes in Affective Motivational Collaboration Theory. In our future work, we will implement complete algorithms for each mechanism and process, thereby automating the computational walkthroughs. Our ultimate goal is a general software platform based on the collaboration structure of SharedPlans theory [22], and employing emotion-driven processes such as appraisal [38] to enable a robot to employ emotion-regulated collaborative behaviors in its interactions with humans.

References

1. C. Adam and E. Lorini. A BDI emotional reasoning engine for an artificial companion. In *Workshop on Agents and multi-agent Systems for AAL and e-HEALTH (PAAMS)*, volume 430, pages 66–78. Springer, 2014.
2. M. B. Arnold. *Emotion and personality.* Cassell Co., 1960.
3. J. Bach. The MicroPsi Agent Architecture. *In Proceeding of ICCM-5*, pages 15–20, 2003.
4. J. Bach. *Principles of Synthetic Intelligence PSI: An Architecture of Motivated Cognition.* Oxford University Press, Inc., 2009.
5. J. Bach. A motivational system for cognitive ai. In J. Schmidhuber, K. R. Thórisson, and M. Looks, editors, *Artificial General Intelligence*, volume 6830 of *Lecture Notes in Computer Science*, pages 232–242. Springer Berlin Heidelberg, 2011.
6. J. Bach. Micropsi 2: The next generation of the micropsi framework. In *Proceedings of the 5th International Conference on Artificial General Intelligence*, AGI'12, pages 11–20. Springer-Verlag, 2012.
7. M. E. Bratman. *Intention, Plans, and Practical Reason.* Cambridge, Mass.: Harvard University Press, 1987.
8. C. Breazeal. Role of expressive behaviour for robots that learn from people. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1535):3527–38, 2009.
9. C. Breazeal, A. Brooks, J. Gray, G. Hoffman, C. Kidd, H. Lee, J. Lieberman, A. Lockerd, and D. Mulanda. Humanoid robots as cooperative partners for people. *Journal of Humanoid Robots*, 1(2):1–34, 2004.
10. W. J. Clancey. Roles for agent assistants in field science: Understanding personal projects and collaboration. *IEEE Transactions on Systems, Man and Cybernetics, special issue on Human-Robot Interaction*, 34(2):125–137, 2004.
11. P. Cohen and H. J. Levesque. *Teamwork.* SRI International, 1991.
12. C. M. de Melo, J. Gratch, P. Carnevale, and S. J. Read. Reverse appraisal: The importance of appraisals for the effect of emotion displays on people's decision-making in social dilemma. In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society (CogSci)*, 2012.

13. N. Esau, L. Kleinjohann, and B. Kleinjohann. Integrating emotional competence into man-machine collaboration. In *Biologically-Inspired Collaborative Computing, September 8-9, Milano, Italy*, pages 187–198, 2008.
14. O. García, J. Favela, G. Licea, and R. Machorro. Extending a collaborative architecture to support emotional awareness. In *Emotion Based Agent Architectures (ebaa99*, pages 46–52, 1999.
15. E. Goffman. *The Presentation of Self in Everyday Life*. Anchor, 1959.
16. A. C. Gonzalez, M. Malfaz, and M. A. Salichs. An autonomous social robot in fear. *IEEE Transactions Autonomous Mental Development*, 5(2):135–151, 2013.
17. J. Gratch and S. C.Marsella. A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4):269–306, 2004.
18. J. Gratch, S. Marsella, N. Wang, and B. Stankovic. Assessing the validity of appraisal-based models of emotion. In *International Conference on Affective Computing and Intelligent Interaction*. IEEE, 2009.
19. B. J. Grosz and L. Hunsberger. The dynamics of intention in collaborative activity. *Cognitive Systems Research*, 7(2-3):259–272, 2007.
20. B. J. Grosz, L. Hunsberger, and S. Kraus. Planning and acting together. *AI Magazine*, 20(4):23–34, 1999.
21. B. J. Grosz and S. Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.
22. B. J. Grosz and C. L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, July 1986.
23. B. J. Grosz and C. L. Sidner. Plans for discourse. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 417–444. MIT Press, Cambridge, MA, 1990.
24. E. Hudlicka. Reasons for emotinos: Modeling emotinos in integrated cognitive systems. In W. D. Gary, editor, *Integrated Models of Cognitive Systems*, volume 59, pages 1–37. New York: Oxford University Press, 2007.
25. R. P. M. III and J. E. Laird. Emotion-driven reinforcement learning. In *CogSci 2008*, 2008.
26. A. Kabil, C. D. Keukelaere, and P. Chavaillier. Coordination mechanisms in human-robot collaboration. In *Proceeding of the 7th International Conference on Advances in Computer-Human Interactions*, pages 389–394, 2014.
27. H.-R. Kim and D.-S. Kwon. Computational model of emotion generation for human-robot interaction based on the cognitive appraisal theory. *Journal of Intelligent and Robotic Systems*, 60(2):263–283, 2010.
28. K. Kiryazov, R. Lowe, C. Becker-Asano, and T. Ziemke. Modelling embodied appraisal in humanoids : Grounding pad space for augmented autonomy. In *Proceedings of the Workshop on Standards in Emotion Modeling*, 2011.
29. H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, E. Osawai, and H. Matsubara. Robocup: A challenge problem for AI. *AI Magazine*, 18(1):73–85, 1997.
30. R. S. Lazarus. *Emotion and Adaptation*. OXFORD University Press, 1991.
31. H. J. Levesque, P. R. Cohen, and J. H. T. Nunes. On acting together. In *AAAI*, pages 94–99. AAAI Press / The MIT Press, 1990.
32. D. J. Litman and J. F. Allen. Discourse processing and commonsense plans. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 365–388. MIT Press, Cambridge, MA, 1990.
33. C. Liu and N. Sarkar. Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE TRANSACTIONS ON ROBOTICS*, 24(4):883–896, 2008.
34. K. E. Lochbaum. A collaborative planning model of intentional structure. *Computational Linguistics*, 24(4):525–572, 1998.
35. K. E. Lochbaum, B. J. Grosz, and C. L. Sidner. Models of plans to support communication: An initial report. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 485–490. AAAI Press, 1990.
36. C. Marinetti, P. Moore, P. Lucas, and B. Parkinson. Emotions in social interactions: Unfolding emotional experience. In *Emotion-Oriented Systems, Cognitive Technologies*, pages 31–46. Springer Berlin Heidelberg, 2011.

37. S. Marsella, J. Adibi, Y. Al-Onaizan, A. Erdem, R. Hill, G. A. Kaminka, Z. Qiu, and M. Tambe. Using an explicit teamwork model and learning in robocup: An extended abstract. In *RoboCup-98: Robot Soccer World Cup II*, volume 1604, pages 237–245. Springer Berlin Heidelberg, 1999.
38. S. C. Marsella and J. Gratch. EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1):70–90, March 2009.
39. S. W. McQuiggan and J. C. Lester. Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies*, 65(4):348–360, 2007.
40. V. Montreuil, A. Clodic, M. Ransan, and R. Alami. Planning human centered robot activities. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pages 2618–2623, 2007.
41. B. Mutlu, A. Terrell, and C.-M. Huang. Coordination mechanisms in human-robot collaboration. In *Proceedings of the HRI 2013 Workshop on Collaborative Manipulation*, 2013.
42. S. Nikolaidis, P. A. Lasota, G. F. Rossano, C. Martinez, T. A. Fuhlbrigge, and J. A. Shah. Human-robot collaboration in manufacturing: Quantitative evaluation of predictable, convergent joint action. In *ISR*, pages 1–6, 2013.
43. A. Paiva, J. Dias, D. Sobral, R. Aylett, P. Sobreperez, S. Woods, C. Zoll, and L. Hall. 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*, pages 194–201, 2004.
44. A. Paiva, I. Leite, and T. Ribeiro. Emotion modeling for sociable robots. In J. G. A. K. Rafael A. Calvo, Sidney D'Mello, editor, *Handbook of Affective Computing*, pages 296–308. Oxford University Press, 2014.
45. S. Planalp. *Communicating Emotion: Social, Moral, and Cultural Processes.* Cambridge University Press, 1999.
46. M. Pontier and J. F. Hoorn. How women think robots perceive them - as if robots were men. In *International Conference on Agents and Artificial Intelligence (ICAART-2)*, pages 496–504, 2013.
47. I. Ravenscroft. *Folk Psychology as a Theory.* Stanford Encyclopedia of Philosophy, 2004.
48. C. Rich and C. L. Sidner. COLLAGEN: A collaboration manager for software interface agents. *User Modeling User-Adapted Interaction*, 8(3-4):315–350, 1998.
49. C. Rich, C. L. Sidner, and N. Lesh. COLLAGEN: Applying collaborative discourse theory to human-computer interaction. *AI Magazine*, 22(4):15–26, 2001.
50. J. Rickel, N. Lesh, C. Rich, C. L. Sidner, and A. Gertner. Collaborative discourse theory as a foundation for tutorial dialogue. In *Proceedings Sixth International Conference on Intelligent Tutoring Systems*, 2002.
51. I. J. Roseman and C. A. Smith. Appraisal theory:overview, assumptions, varieties, controversies. In K. R. Scherer, A. Schorr, and T. Johnstone, editors, *Appraisal process in emotion*, pages 3–34. NY: Oxford University Press, 2001.
52. D. Sander, D. Grandjean, and K. R. Scherer. A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18(4):317–352, 2005.
53. P. Scerri, D. Pynadath, L. Johnson, P. Rosenbloom, M. Si, N. Schurr, and M. Tambe. A prototype infrastructure for distributed robot-agent-person teams. In *The Second International Joint Conference on Autonomous Agents and Multiagent Systems*, 2003.
54. P. Scerri, D. Pynadath, L. Johnson, P. Rosenbloom, M. Si, N. Schurr, and M. Tambe. A prototype infrastructure for distributed robot-agent-person teams. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '03, pages 433–440, New York, NY, USA, 2003. ACM.
55. K. R. Scherer. On the nature and function of emotion: A component process approach. In K. R. Scherer and P. Ekman, editors, *Approaches To Emotion*, pages 293–317. Lawrence Erlbaum, Hillsdale, NJ, 1984.
56. K. R. Scherer. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3459–3474, 2009.
57. K. R. Scherer, A. Schorr, and T. Johnstone. *Appraisal Processes in Emotion: Theory, Methods, Research.* Oxford University Press, 2001.

58. M. Shayganfar, C. Rich, and C. L. Sidner. A design methodology for expressing emotion on robot faces. In *IROS*, pages 4577–4583. IEEE, 2012.
59. C. L. Sidner and M. Dzikovska. A first experiment in engagement for human-robot interaction in hosting activities. In *Advances in Natural Multimodal Dialogue Systems*, volume 30 of *Cognitive Technologies*, pages 55–76. Springer Netherlands, 2005.
60. D. Sofge, M. D. Bugajska, J. G. Trafton, D. Perzanowski, S. Thomas, M. Skubic, S. Blisard, N. Cassimatis, D. P. Brock, W. Adams, and A. C. Schultz. Collaborating with humanoid robots in space. *International Journal of Humanoid Robotics*, 2(2):181–201, 2005.
61. M. Tambe. Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:83–124, 1997.
62. C. Urban. Pecs: A reference model for human-like agents. In *Deformable Avatars*. Netherlands: Kluwer Academic Publishers, 2001.
63. S. van Hooft. Scheler on sharing emotions. *Philosophy Today*, 38(1):18–28, 1994.
64. J. D. Velàsquez. Modeling emotions and other motivations in synthetic agents. In *Proceedings of the 14th Nnational Conference on Artificial Intelligence AAAI-97*, pages 10–15, 1997.
65. D. Vogiatzis, C. Spyropoulos, V. Karkaletsis, Z. Kasap, C. Matheson, and O. Deroo. An affective robot guide to museums. In *Proceedings of the 4th International Workshop on Human-Computer Conversation*, 2008.
66. T. Wehrle. Motivations behind modeling emotional agents: Whose emotion does your robot have?, 1998.
67. A. K. Wisecup, D. T. Robinson, and L. Smith-Lovin. *The Sociology of Emotions*. SAGE Publications, Inc., 2007.
68. J. Yen, J. Yin, T. R. Ioerger, M. S. Miller, D. Xu, and R. A. Volz. Cast: Collaborative agents for simulating teamwork. In *Proceedings of IJCAI2001*, pages 1135–1142, 2001.
69. J. Yin, M. S. Miller, T. R. Ioerger, J. Yen, and R. A. Volz. A knowledge-based approach for designing intelligent team training systems. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 427–434. ACM, 2000.
70. Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. A survey of affect recognition methods: Audio, visual and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):39–58, 2009.