Toward Improving Human-Robot Collaboration with Emotional Awareness

Mahni Shayganfar · Charles Rich ·
Candace L. Sidner

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.

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 [22]. 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] [22] [31], and are derived from Bratman's BDI architecture [7]. Two theories, Joint Intentions [11] and SharedPlans [19, 20, 22], have been used to support teamwork and collaboration between humans and robots or virtual agents [9] [39] [58] [67]. 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 [35]. 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 [66].

There is also a communicative aspect of emotions. For instance, emotions are often intended to convey information to others [15]. Emotions are also in-

volved 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 [44]. The interpretation of these relational meanings are handled by the appraisal of events. Appraisal processes give us a way to view emotion as social [62]. 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 [35]. 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 2.3 and 4.1), and delegating a new task to the robot (Sections 2.5 and 4.3). 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 2.3 to 2.6 are revisited in more detail in our computational walkthroughs in Section 4. In this section, we show how the mechanisms of Affective Motivational Collaboration Theory operate to produce the robot behaviors in Section 2.

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 Example Scenarios

2.1 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.

2.2 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.

2.3 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:** I see. This is frustrating. 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 4.1, we will show how the computational mechanisms discussed in Section 3.1 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.

2.4 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. What do you want me to do?

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 3). 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).

2.5 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 3.1), acknowledges the Astronaut's emotion just as it did in Section 2.3. 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.

2.6 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. <u>How about the conflict that might occur while I am trying to fix the second panel?</u> 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:** <u>Do you want me to provide some examples?</u>

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 indi-

cates 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).

3 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.

3.1 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 [19, 20, 22]. 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

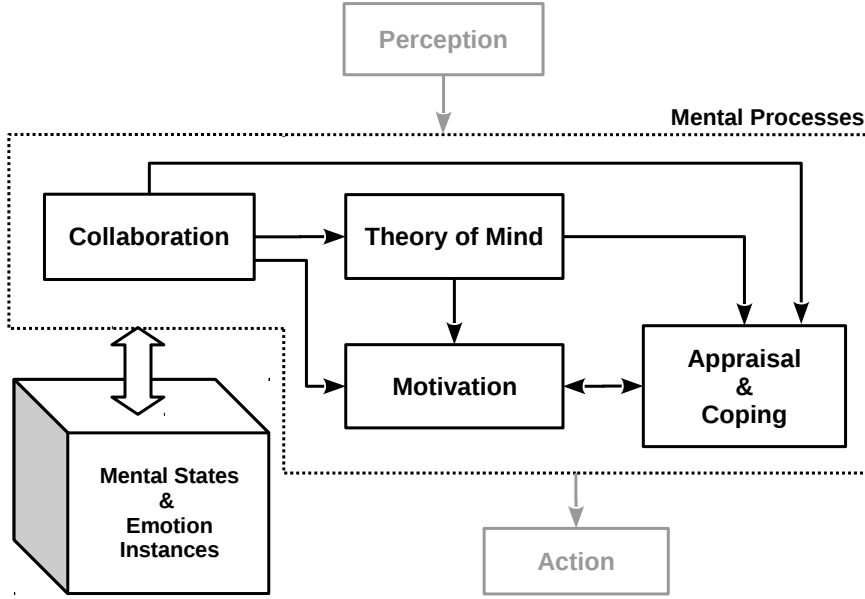states possess multiple attributes impacting the relation between cognition and behavior or perception.

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.

3.2 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.

3.3 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 [20,21] 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 4 for some examples) [37] [56]. 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.

3.4 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.

3.5 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.

3.6 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.

3.7 Mental States & Emotion Instances

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

3.7.1 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 [22] 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.

3.7.2 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 [37]. 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.

3.7.3 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.

3.7.4 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.

3.7.5 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.

4 Computational Walkthroughs

In this section, we explain in detail how the individual computational mechanisms described in Section 3 generate the Robot's behaviors in each example in Section 2. Since our walkthrough explanation of underlying processes is based on the collaborators' utterances, we use verbal expression of emotions. However, nonverbal emotional expression (e.g., facial expressions) can provide the same impact during collaboration. The following four walkthroughs are in the same order as the four examples in Section 2. The names of the mechanisms written in parentheses and bold below indicate which mechanism is involved in each step. There are also specific processes indicated after the mechanism, if appropriate.

4.1 Agreeing on the Shared Goal (Emotion Awareness)

This section provides a step-by-step walkthrough of the example presented in Section 2.3, beginning with the Astronaut's utterance below:

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

(Perception) The Robot perceives the Astronaut's utterances and emotion.

First, the Robot perceives the Astronaut's utterances as well as her emotion in the first turn (A1). The output of Perception is beliefs about the task in the Astronaut's focus of attention, and also the Astronaut's emotion which she has expressed verbally. The beliefs formed about the task (i.e., installing the panel) include:

- the Astronaut's proposal of *stopping* the task,
- which is a *future* event,
- and is *caused by* the measurement tool problem.

Also, beliefs formed about the Astronaut's emotion (i.e., frustration) include:

- the existence of a *negative-valenced* emotion,
- and is verbally conveyed as *frustration*.

(**Collaboration:** *Monitoring & Focus Shifting*) Based on these perceptions, the Robot uses the Collaboration mechanism, and forms new beliefs about the collaboration status. These new beliefs concern:

- the *unsatisfied* precondition of the Astronaut's current *task*,
- the *blocked* status of the Astronaut's current *task*,
- and consequently the *blocked* status of the *shared goal*,
- which causes the change in the Robot's *focus of attention* to the Astronaut's task.

(**Theory of Mind:** *Reverse Appraisal & User Modeling*) The Robot uses reverse appraisal to understand the meaning of the Astronaut's frustration according to the collaborative task status (e.g., precondition and shared goal status). The Robot updates the Astronaut's user model accordingly.

The reverse appraisal process forms beliefs about the anticipated appraisals of the Astronaut with respect to the current task's status, based on the Astronaut's utterances and emotion in A1, and the output of the Collaboration mechanism. Some of these anticipated appraisal values indicate that the event is interpreted as *relevant*, *undesirable*, *uncontrollable*, *urgent*, and *unexpected* by the Astronaut. Furthermore, the user modeling process updates the Astronaut's user model based on the output of the reverse appraisal process and the Collaboration mechanism; this user modeling process forms beliefs that:

- the Astronaut has *low autonomy*,
- the Astronaut is a *highly communicative* collaborator.

(**Appraisal**) The Robot appraises the Astronaut's utterances and emotion.

The Appraisal mechanism uses distinct processes to compute values for individual appraisal variables. Some of these processes use beliefs formed based on user modeling process for the appraisals, e.g., controllability. The output of these processes provides a vector of values describing the Robot's interpretation of the current event (A1). In this example, all the beliefs listed above including the negative-valenced emotion provide a vector of values that would be interpreted as worry. The Action component has the task of expressing this emotion, if required.

(**Motivation:** *Motive & Intention Formation*) The Robot forms new motives according to the result of:

a) *appraisal* with respect to the shared goal,
b) *reverse appraisal* of the Astronaut's emotion,
c) and the *user model* of the Astronaut.

Then, the motive comparison process compares current available motives and sorts them based on their distance to the Astronaut's emotional state and the achievement of the shared goal. Here, the distance function depends on

a) the Astronaut's emotional state (i.e., frustration) as an approximation of her mental state, and b) how taking an action based on the corresponding intention of a particular motive improves the possibility of the collaborators reaching a mutually accepted shared goal. The Robot chooses the first motive on the sorted list to form an intention to postpone asking questions about the alternative solution because of the Astronaut's negative emotion, the second motive to form an intention for acknowledging the Astronaut's perceived emotion, and the third motive to form another intention for resolving the current blocking in the collaboration. After this whole process, the Robot uses the Coping mechanism to take an action based on the current intention.

(Coping) With respect to the current intentions and based on different beliefs, the Robot selects distinct coping strategies. Therefore, based on the Astronaut's perceived negative emotion, the Robot chooses the *restraint* coping strategy (an emotion-focused strategy) which causes the Robot to postpone asking about an alternative solution. Then, because of the Astronaut's anticipated appraisals (reverse appraisal), e.g., undesirability of the event, the Robot chooses an *active* coping strategy (a problem-focused strategy) to acknowledge the Astronaut's negative emotion, which is an active step to remove the Astronaut's stressor within the collaboration. Finally, based on the current mental state, e.g., the Astronaut's blocked task, and the Astronaut's user model, the Robot chooses a *planning* coping strategy (another problem-focused strategy) to provide an alternative solution to the Astronaut. Notice that there are three coping strategies being used here.

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

The Astronaut's next utterance (A3) provides the Robot with a question about whether the Robot can fix the measurement tool.

> *A3. Astronaut:* Can you fix the measurement tool?

Using the same processes as for A2, the following beliefs hold:

- the precondition associated with the Astronaut's current *task* is still *unsatisfied*,
- the status of the Astronaut's current *task* is still *blocked*,
- and similarly the status of the *shared goal* is still *blocked*,
- however, the Astronaut's question changes the Robot's *focus of attention* to the measurement tool,
- also, the Astronaut's *emotion* has changed to *neutral*.
- but, her user model *stays the same*, i.e., having *low-autonomy* and being *highly communicative*.

Notice that the Collaboration and Appraisal mechanisms in the next two steps operate to form required beliefs for the Motivation mechanism. The new

motives will help the Robot to remove the current impasse by negotiating an alternative solution.

(Collaboration) The change in the focus of attention to the measurement tool causes the Robot to check the availability of a recipe to fix or replace the malfunctioning measurement tool. The Robot finds a recipe to replace the measurement tool.

(Appraisal) The Robot appraises the possibility of replacing the measurement tool with respect to: a) the status of the shared goal, and b) the Astronaut's user model. For instance, the Astronaut's *low-autonomy* in her current task lowers the possibility of *changeability* of the current state of the world for the Robot, and the Astronaut's *highly communicative* behavior impacts the Robot's decision on whether the offer that the Robot is about to make is *relevant* and *desirable* for the Astronaut. Robot finds the replacement of the measurement tool *relevant*, *desirable*, and *controllable*.

(Motivation: *Motive & Intention Formations*) The Robot forms new motives based on:

- belief about a positive change in the Astronaut's emotion (i.e., from frustration to neutral),
- belief about the lack of time,
- belief about result of the appraisal of the replacing the measurement tool,
- and belief about the availability of a recipe to replace the measurement tool.

The Robot forms the corresponding intentions with respect to the new motives. Once again, the Robot uses the Coping mechanism to take actions based on the recent intentions.

(Coping) With respect to the Robot's intention to alleviate the pressure of the shortage of time, the Robot chooses an active coping strategy to remove the current stressor (i.e., lack of time). Also, with respect to the Robot's intention to maintain the recent improvement in the Astronaut's emotional state, and the intention to fetch another measurement tool, the Robot chooses a planning strategy, to explain the next task to the Astronaut, and to negotiate and offer an alternative action.

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.

At this point, the Astronaut is content with the way the Robot outlined the shared goal and responds correspondingly (A5). The Robot perceives and interprets the Astronaut's response as an agreement on their new shared goal as discussed above.

> **A5. Astronaut:** That would be great!

4.2 Agreeing on Shared Goal (Emotion Ignorance)

This walkthrough begins with the same utterance as the previous one, and corresponds to the emotional ignorance example in Section 2.4. In emotional ignorance examples, we assume the Robot always perceives the Astronaut as expressing neutral emotion.

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

(Perception) The Robot perceives only the non-emotional content of the Astronaut's utterances (B1).

Similarly to the previous example, Perception forms beliefs about the task in the Astronaut's focus of attention. These beliefs include:

- the Astronaut's proposal of *stopping* the task,
- which is a *future* event,
- and is *caused by* the measurement tool problem.

Notice that beliefs about the Astronaut's emotion are formed differently in comparison with the previous example, and are based on the neutral emotion of the Astronaut, since the Robot ignores the Astronaut's actual emotion instance, i.e., frustration.

(Collaboration) The Robot uses the Collaboration mechanism to form new beliefs about the collaboration status based on its perception. These new beliefs are the same as those generated by the Collaboration mechanism in the previous example.

(Theory of Mind: *Reverse Appraisal & User Modeling*) The Robot uses reverse appraisal to understand the meaning of the Astronaut's neutral emotion according to the collaborative task status (e.g., precondition and shared goal status). The Robot updates the Astronaut's user model respectively.

In this example, since the Robot misses the actual expressed emotion by the Astronaut (i.e., frustration) and incorrectly perceives her as having neutral emotion, the corresponding reverse appraisal values lead to the wrong interpretation of the event. The Robot thinks the Astronaut interprets the event as *desirable*, *controllable*, *non-urgent*, and *expected* (all of which are incorrect). Furthermore, the user modeling process forms incorrect beliefs:

- Astronaut has *high autonomy*,

- Astronaut is a *moderately communicative* collaborator.

(Appraisal) The Appraisal mechanism operates similarly to what we discussed in Section 2.3. The output of these processes provides a vector of values describing the Robot's interpretation of the current event (B1). The outcome will also be mapped to a particular emotion instance, but since the Robot misses the Astronaut's emotion, it maps the appraisals to a different emotion, i.e., hope, than the one elicited in previous example. The Robot generates hope because it believes the Astronaut's emotion is neutral and the current task is blocked. Therefore, the Robot wants to come up with an alternative solution immediately.

(Motivation: *Motive & Intention Formation*) Although the process of comparing and sorting available motives here is similar to the previous example, all of the new motives are different. The reason is that each of the three sources of motives forms a different motive due to the ignorance of the Astronaut's actual emotion. For instance, the motive generated with the influence of appraisal in the emotional awareness example urges the Robot to postpone asking questions about the alternative solutions while the motive with the same cause (i.e., appraisal) here urges the Robot to immediately try to fix the problem and come up with alternative solutions by asking questions. The Robot, similarly to the previous example, selects the most related motives and forms new intentions with respect to the current status of collaboration. After this whole process, the Robot uses the Coping mechanism to take actions based on the available intentions.

(Coping) Based on the current mental state and with a similar process as in the previous example, the Robot decides to use a problem-focused coping strategy of seeking information to be able to choose between two available actions and reduce the current amount of uncertainty. Therefore, the Robot, without acknowledging the Astronaut's emotion, asks the Astronaut to choose between two alternative solutions (B2).

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

As we mentioned earlier, the Robot's response does not make any progress in the collaboration status. Hence, the Astronaut repeats herself about the task status (B3).

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

The Robot perceives the Astronaut's new utterance (B3) while, again, ignoring her frustration. The Robot goes through the same process as described above, and since the Astronaut has just repeated herself, her new utterances do not change the Robot's mental state. Having the same mental state causes

the Robot to ask a similar question (B4).

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

This time, the Robot creates ambiguity for the Astronaut because part of the Robot's utterance (i.e., "fix this problem") does not make it clear whether the Robot can fix the measurement tool, or whether Robot has another potential solution. Hence, the ambiguity of Robot's question does not help the Astronaut's frustration and causes her to ask a clarification question (B5).

> **B5. Astronaut:** Can you fix my measurement tool?

All the beliefs formed at this point are the same as those of turn A3 in Section 4.1, except the following ones:

- the Astronaut's *emotion* is still believed to be *neutral*,
- but, her user model *has changed* to having *medium-autonomy* and being *highly communicative*.

Notice that similar to the process we discussed in turn A4, the Collaboration and Appraisal mechanisms operate to form required beliefs for the Motivation mechanism.

(**Motivation:** *Motive & Intention Formations*) The Robot forms new motives similar to the turn A4 based on the same set of beliefs, except that in this example the Robot misses a belief about a positive change in the Astronaut's emotion (since the Robot always believes the Astronaut has neutral emotion).

(**Coping**) Here, the Robot does not have an intention to maintain the recent improvement in the Astronaut's emotional state. However, since the Robot does have an intention to alleviate the pressure of the shortage of time, and an intention to fetch another measurement tool, the Robot chooses a planning strategy, to immediately respond to the Astronaut's question, and try to negotiate and offer an alternative action to the Astronaut (B6).

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

The Astronaut's strong emotion, shortage of time, and the Robot's mismatching answer to the Astronaut's assumption causes the Astronaut to reject the Robot's proposal (B7).

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

After perceiving the Astronaut's response, once again, the Coping mechanism makes the Robot negotiate with the Astronaut to protect the shared

goal from failure.

> **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.

The Astronaut terminates the collaboration due to the lack of time and failure in the Robot's collaborative behavior (B9).

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

In summary, this example shows that ignoring the Astronaut's emotion impacts the Robot's perception and corresponding beliefs. The output of the Collaboration mechanism remains unchanged in comparison with the emotional awareness example. Although the Collaboration mechanism provides the required structural details of collaboration between the Robot and the Astronaut, these structural details are not enough to save a collaboration from failure. As we continue, we can see that ignoring the actual emotion of the Astronaut causes misfunctioning of the processes in the Theory of Mind mechanism, i.e., reverse appraisal and user modeling. Comparing the result of these two processes with the results in the emotional awareness example shows the importance of correctly perceiving a collaborator's emotion. This problem continues even with the Appraisal mechanism which maps the Robot's interpretation of the environment to a wrong emotion. Consequently, all sources of the Motivation mechanism provide incorrect values which drastically influence the formation of the underlying motives of the required intentions. Finally, the Coping mechanism operates based on wrong newly formed intentions which leads to a totally different behavior of the Robot in comparison with the same turn in the emotional awareness example. The divergence of the Robot's collaborative behavior from its successful path continues throughout the Robot and the Astronaut's interaction which increases the required time for achieving the shared goal, and perpetuates the negative feeling of the Astronaut. Consequently, the same collaboration fails even though that the Robot uses the same computational mechanisms, as we showed above.

4.3 Task Delegation (Emotion Awareness vs. Emotion Ignorance)

In this section, we present the last two walkthroughs which focus on task delegation with and without emotional awareness. To avoid redundant explanations, these walkthrough examples are limited to comparison of the key turns in each example emphasizing the crucial differences in the results of the underlying mechanisms.

> ***C1 & 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.

(Perception) Although the Astronaut's first utterance is the same for both examples (C1 & D1), in the second example the Perception mechanism misses the Astronaut's expressed emotion, i.e., worry. The Perception mechanism forms beliefs in both examples similar to what we discussed in Sections 4.1 and 4.2.

(Collaboration) The Collaboration mechanism forms the same beliefs (e.g., about required resources and availability of a recipe) in both examples. In other words, ignorance of the Astronaut's emotion by the Robot does not impact beliefs formed based on the output of the Collaboration mechanism.

(Theory of Mind: *Reverse Appraisal & User Modeling*) Similar to before, in the emotion awareness case, the Robot uses reverse appraisal to understand the meaning of the Astronaut's emotion (i.e., worry) with respect to the collaborative task status, and finds the anticipated appraisals as *relevant*, *undesirable*, *uncontrollable*, *urgent* and *unexpected*. In the emotion ignorance case, the Robot finds most of the anticipated appraisals the same, except that ignoring the Astronaut's negative emotion causes an interpretation of the event as *controllable* and *expected*. As a result, the Robot updates the Astronaut's user model and forms beliefs that a) the Astronaut has *low autonomy* in the emotion awareness example versus *high autonomy* in the emotion ignorance example, and b) the Astronaut is a highly communicative collaborator in both examples.

(Appraisal) The Robot's interpretation of the current event leads to the same appraisal values in both cases as *relevant*, *unexpected*, *undesirable*, *urgent*, and *controllable*.

(Motivation: *Motive & Intention Formation*) The Robot forms new motives in the same way as we discussed in Section 4.1. The motives in the emotion awareness case are formed to acknowledge the Astronaut's negative emotion, and postpone any question regarding required information (e.g., unsatisfied postcondition of the current task) because of the Astronaut's perceived negative emotion. However, in the emotion ignorance example one motive is formed based on the availability of a partial recipe for accepting the delegated task, and another motive is formed based on the unsatisfied precondition of the Robot's ongoing task. The Robot forms corresponding intentions with respect to these motives.

(Coping) In the emotion awareness case, the Robot chooses similar coping strategies as discussed in Section 4.1, to a) acknowledge the Astronaut's negative emotion, b) accept the delegated task, and c) postpone asking questions

about the delegation process. Hence, the Robot provides a proper response (C2) without asking any questions about the delegated task.

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

In contrast, in the emotion ignorance case, not only does the Robot not have an intention to acknowledge the Astronaut's emotion, it also selects an incorrect coping strategy to seek information about dropping its own current unachieved task (D2).

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

After this point, in the emotion awareness case, the Robot, by choosing the restraint coping strategy, waits till the appropriate opportunity (i.e., Astronaut's mitigated emotion) to ask required questions. Therefore, the Robot tries to mitigate the Astronaut's negative emotion which helps them to successfully achieve the goal of a task delegation. Notice that in utterance C4, since the Robot knows that the Astronaut is still worried, it only informs the Astronaut about some potential questions in the future. Consequently, the Robot's collaborative behavior also mitigates the Astronaut's negative emotion.

> **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 end, the Astronaut finds the Robot's response appropriate for the delegated task. Thus, the Robot's proper response mitigated the Astronaut's negative emotion.

In the emotion ignorance case, the rest of the Robot's utterances from D4 to D10 (see Section 2.6) show that the Robot wants to do planning in the most efficient manner possible by asking many questions about missing information according to the shared plan. However, the Robot's improper collaborative behavior, caused by choosing inappropriate coping strategies diverges the task delegation process from achieving the goal. We can see as time passes the content of the Astronaut's utterances becomes less informative for the Robot, e.g., D5 and D9. Consequently, the Robot needs more confirmations from the Astronaut, e.g., D4 and D8, to obtain required information for the execution of the delegated task. Finally, the Astronaut terminates the collaboration due to the lack of time and the Robot's failure to incorporate proper collaborative behavior (D11).

5 Related Work

The prominent collaboration theories are mostly based on plans and joint intentions [11, 22], and they were derived from the BDI paradigm developed by Bratman [7] which is fundamentally reliant on folk psychology [46]. The two theories, Joint Intentions [11] and SharedPlans [22], 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 [19, 20, 22]. 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 [60]. Some of these works present algorithms and computational models in a teamwork environment based on the underlying structure of the SharedPlans theory [33, 34, 67, 68], and Joint Intentions theory [9, 40]. The hybrid teamwork model, STEAM [60], has also been successfully applied to a variety of domains [25, 28, 36, 52]. 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 [19, 20, 22], and it focuses on the processes that generate, maintain and update this structure based on mental states. COLLAGEN [47, 48] 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 [49]. 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 [41, 53]. Some researchers have an overall look at a collaboration concept at the architectural level [13, 14, 59]. There are other concepts such as joint actions and commitments [18], dynamics of intentions during collaboration [30] 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 [65]. 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 [61], and some are using a combination of concepts from the psychological theories [27]. 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, 43, 57]. Robots also use emotions theories for some other purposes such as automatic affect recognition using different modalities [69], and behavior adaptation [32].

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 [29] and then were actively developed in the early 80s by Ellsworth and Scherer and their students [50, 51, 54–56]. 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 [37] 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 [17]. 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, 26, 37]. In robotics, appraisal theory has been used to establish and maintain a better interaction between a robot and a human [26, 51, 64]. Appraisal theory has also been used in robots' decision making [16], or in their cognitive systems [23, 24]. In the virtual agents community, empathy and affective decision-making is a research topic that has received much attention in the last decade [38, 42, 45, 63].

6 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 Shared-Plans theory [21], and employing emotion-driven processes such as appraisal [37] 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, 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.

18. B. J. Grosz and L. Hunsberger. The dynamics of intention in collaborative activity. *Cognitive Systems Research*, 7(2-3):259–272, 2007.

19. B. J. Grosz, L. Hunsberger, and S. Kraus. Planning and acting together. *AI Magazine*, 20(4):23–34, 1999.

20. B. J. Grosz and S. Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.

21. B. J. Grosz and C. L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, July 1986.

22. 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.

23. 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.

24. R. P. M. III and J. E. Laird. Emotion-driven reinforcement learning. In *CogSci 2008*, 2008.

25. 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.

26. 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.

27. 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.

28. 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.

29. R. S. Lazarus. *Emotion and Adaptation*. OXFORD University Press, 1991.

30. 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.

31. 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.

32. 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.

33. K. E. Lochbaum. A collaborative planning model of intentional structure. *Computational Linguistics*, 24(4):525–572, 1998.

34. 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.

35. 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.

36. 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.

37. S. C. Marsella and J. Gratch. EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1):70–90, March 2009.

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

44. S. Planalp. *Communicating Emotion: Social, Moral, and Cultural Processes*. Cambridge University Press, 1999.

45. 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.

46. I. Ravenscroft. *Folk Psychology as a Theory*. Stanford Encyclopedia of Philosophy, 2004.

47. 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.

48. C. Rich, C. L. Sidner, and N. Lesh. COLLAGEN: Applying collaborative discourse theory to human-computer interaction. *AI Magazine*, 22(4):15–26, 2001.

49. 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.

50. 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.

51. D. Sander, D. Grandjean, and K. R. Scherer. A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18(4):317–352, 2005.

52. 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.

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 *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '03, pages 433–440, New York, NY, USA, 2003. ACM.

54. 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.

55. 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.

56. K. R. Scherer, A. Schorr, and T. Johnstone. *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press, 2001.

57. M. Shayganfar, C. Rich, and C. L. Sidner. A design methodology for expressing emotion on robot faces. In *IROS*, pages 4577–4583. IEEE, 2012.

58. 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.

59. 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.

60. M. Tambe. Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:83–124, 1997.

61. C. Urban. Pecs: A reference model for human-like agents. In *Deformable Avatars*. Netherlands: Kluwer Academic Publishers, 2001.

62. S. van Hooft. Scheler on sharing emotions. *Philosophy Today*, 38(1):18–28, 1994.

63. 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.

64. 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.

65. T. Wehrle. Motivations behind modeling emotional agents: Whose emotion does your robot have?, 1998.

66. A. K. Wisecup, D. T. Robinson, and L. Smith-Lovin. *The Sociology of Emotions*. SAGE Publications, Inc., 2007.

67. 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.

68. 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.

69. 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.