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

Toward Improving Human-Robot Collaboration with Emotional Awareness

Mohammad Shayganfar · Charles Rich · Candace L. Sidner

Received: date / Accepted: date

Abstract Current computational theories 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 will be 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 [25]. There are many challenges in achieving a successful collaboration between robots and humans. To meet this challenges, it is crucial to understand what makes a collaboration not only successful, but also efficient. Existing computational

Mohammad 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 [13] [25] [41], and are derived from Bratman's BDI architecture [4]. Two theories, Joint Intentions [13] and SharedPlans [22,23,25], have been used to support teamwork and collaboration between humans and robots or virtual agents [8] [52] [74] [85]. 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 maintining 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 importantly experienced in a social context and acquire their significance in relation to this context [45]. 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 [84].

There is also a communicative aspect of emotions. For instance, emotions are often intended to convey information to others [17]. 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 [57]. The interpretation of these relational meanings are handled by the appraisal of events. Appraisal processes give us a way to view emotion as social [80]. 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 [45]. 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 iteracting 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 asscoiated with each mechanism using JESS (JAVA Expert System Shell) to be able to generate the same type of collaborative behaviors as they appear in our examples. In the future, we are going to use these rules and the processes involved in each mechanism to build 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 of their private and/or shared goals because of a blocked or a protracted subtask. 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. The following is part of an interaction between the Astronaut and the Robot during their collaboration on installing solar panels.

2.3 Agreeing on Shared Goal (Emotion-Awareness)

This hypothetical example and the emotion-ignorant version in the next section demonstrate that agreeing on a shared goal requires the Robot to be aware 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 that is used for checking the quality of the installed panel. 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 leads to acknowledgement of the Astronaut's emotion. See the absence of these utterances as the consequence of ignoring the Astronaut's emotions in the same turn in the next example (Section 2.4). 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 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 in A3 the Astronaut is no longer expressing a negative emotion (i.e., frustration), and is asking for help, the Robot can provide an alternative task as a potential solution (A4). This 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.

In the next example we show what happens in the same hypothetical situation when the Robot ignores the Astronaut's emotion and tries to save the collaboration process from failure.

2.4 Agreeing on Shared Goal (Emotion-Ignorance)

This example shows the same process of agreeing on a shared goal as in the previous section except that it diverges from reaching an agreement, despite the fact that 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 prevent it from failure without considering its collaborator's negative emotion which can be the result of the task failure. 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 procedure during collaboration.

The Robot, in this example, does not perceive the Astronaut's emotion, since it does not apply 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), the Robot's 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 disfunctional measurement tool and proposes the potential solution of replacing the broken tool and asking whether the Astronaut agrees on that (B6).

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, and therefore 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 dissociated reasoning about the problem from the Astronaut's mental state. Finally, the poor interaction of the Robot caused by its emotionally-ignorant behavior which developed from the Robot's very first response and has continued till the end, 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 until such time as they are critical.

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. <u>Don't worry.</u> 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 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 first example in Section 2.3. The underlined utterance in the Robot's turn C2, shows the Robot's awareness of the Astronaut's emotion. Also, the Robot does not ask the Astronaut (because of perceiving her worry) if it is okay to leave the current task which was helping the Astronaut to install the first panel. Because the Robot knows that 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 the actual detailed questions about the delegated task, declares the possibility of asking some follow-up questions whenever the Astronaut's answers are required to make progress in executing some primitive tasks while installing the second panel (C4). Thus, the Robot not only prevents overwhelming the Astronaut with several questions related to executing the next steps in the future, but its utterance implicitly reveals its knowledge of 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.

The next section shows the same task delegation situation with the negative impact of the Robot ignoring the Astronaut's emotion.

2.6 Task Delegation (Emotion-Ignorance)

This last example illustrates how ignoring the collaborator's emotions in a task delegation procedure 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 utterance of the Astronaut 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 worrying 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 to 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 uses the Collaboration mechanism (see Section 3.1) to obtain required information such as task dependencies, existing preconditions and required resources. Subsequently, the absence of required information and dependencies to the Astronaut's actions causes the Robot to ask a question about whether the Astronaut will be able to help in some parts of the task (D4). Although this type of interactive behavior is crucial in many collaborative contexts, here it is counter productive. Thus, the Astronaut succinctly responds to the Robot's question while she is still worried about finishing installation of the first panel (D5). The Robot asks another question about the required inputs for the task which are dependent on the Astronaut without considering her worry (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, the Robot continues to attempt to develop a full plan due to ignorance of the Astronaut's frustration. Then, the Astronaut does not recall 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).

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 based on the *Affective Motivational Collaboration Theory*. Notice, in Fig. 1, there are two mechanisms, i.e., Perception and Action, which are not part of *Affective Motivational Collaboration Theory*. They only provide required input and output to our framework which can differ based on the capabilities of any 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 [22,23,25]. *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.

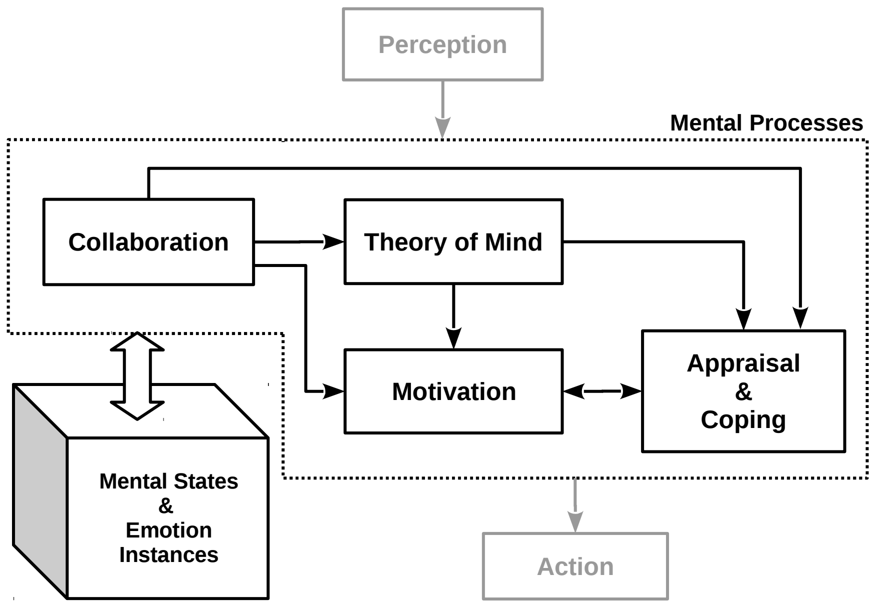

Fig. 1 Computational framework based on *Affective Motivational Collaboration Theory* (arrows indicate primary influences between mechanisms).

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 [23, 24] 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) [48] [71]. 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 [14], 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 [25] 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 are an essential category of the strategies in the Coping mechanism [48]. 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. The following four walkthrough examples are in the same order as the four examples in Section 2. The name of the mechanisms written in parantheses and bold indicates below which mechanism is involved in each step. There are also specific processes written after a colon in front of the mechanism, if appropriate.

4.1 Agreeing on Shared Goal (Emotion-Awareness)

This section provides a step-by-step walkthrough explanation of the example presented in Section 2.3. In this example, the explanation between Astronaut's utterance A1 and Robot's utterance A2 illustrates how the Robot perceives and interprets events including Astronaut's utterances and emotional expressions, and how the Robot takes an appropriate action whenever it is required.[1]

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.

[1] Since our walkthrough explanation of underlying processes is based on collaborators' utterances, we use **verbal** expression of emotions within the utterances to emphasize their existence in certain parts of the collaboration. However, although the nonverbal emotional expressions (e.g., facial expressions) can provide the same impact during collaboration, the automatic recognition of them is out of our research context.

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 both verbally and nonverbally. 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 are about:

- 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 correspondingly.

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:

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

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

The Appraisal mechanism simultaneously uses distinct processes to compute values for each individual appraisal variable. The output of these processes provides a vector of values describing the Robot's interpretation of the current event (A1). In this example, the beliefs listed above including the negative-valenced emotion provide a vector of values that would be read as frustration. The Action component has the task of expressing this emotion.

(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 is a function of a) the Astronaut's emotional state as an admissible 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 to the collaborators' mutually accepted shared goal. The Robot ultimately selects the most related motive and forms a new intention with respect to the current status of collaboration. After this whole process, the Robot uses the coping mechanism to take an action based on the available intention.

(Coping) Based on the current mental state, the Robot chooses an emotion-focused coping strategy, decides to acknowledge the Astronaut's emotion, and to provide an alternative solution. Subsequently, the Robot responds to the Astronaut with A2.

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 new utterance (A3) provides the Robot with a new question about whether the Robot can fix the measurement tool. The following paragraphs show how the robot employs the same mechanisms described in Section 3 to negotiate with the Astronaut to reach an agreement on the shared goal.

A3. Astronaut: Can you fix the measurement tool?

Using the same processes as before, 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.

(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. Robot finds the replacement of the measurement tool *relevant*, *desirable*, and *controllable*.

(Motivation: *Motive & Intention Formations*) The Robot forms new motives based on the next task according to the shared plan and outcome of the appraisal of the possibility of replacing 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 an action based on the recent intention.

(Coping) Based on the current mental state, Robot chooses to negotiate and offer an alternative action to the Astronaut (A4).

> ***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, Astronaut is content with the way Robot outlined the shared goal and responds respectively (A5). The Robot perceives and interprets the Astronaut's response as an agreement on their new shared goal as we discussed above.

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

4.2 Agreeing on Shared Goal (Emotion-Ignorance)

This walkthrough example begins with the same utterance as the previous one, and it provides the computational walkthrough of the emotional-ignorance example in Section 2.4. In emotional-ignorance examples, we assume Robot always perceives neutral emotion expressed by the Astronaut. To avoid redundant explanations, we refer to similar processing in the previous walkthrough example, when possible.

> ***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 only perceives the Astronaut's utterances (B1).

Here, in the first step, the Robot perceives the Astronaut's utterances and ignores her expressed emotion in B1, i.e., frustration. 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. These beliefs include:

- the existence of a *neutral valenced* emotion,
- which maps into a three-value vector of *pleasure*, *arousal*, and *dominance*,
- and is verbally conveyed as *neutral* emotion.

(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 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 with neutral emotion, the corresponding anticipated 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 Robot appraises the Astronaut's utterances.

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 Astronaut's emotion, it maps the appraisals to a different emotion, i.e., hope, than the one elicited in previous example. The Robot elicits 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*) As we discussed earlier, 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.

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 above three sources of motives forms a different motive because it holds a different value, which is caused by 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) in emotional-ignorance example 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 motive and forms a new intention with respect to the current status of collaboration. After this whole process, the Robot uses the coping mechanism to take an action based on the available intention.

(Coping) Based on the current mental state, 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 in Section 2.4, 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 we 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, Robot's question makes an ambiguous assumption for the Astronaut on whether the Robot can fix the disfunctional measurement tool for her. The ambiguity of Robot's question does not help Astronaut's frustration and causes her to ask a clarification question (B5).

B5. Astronaut: Can you fix my measurement tool?

Again, the same processes will be run (similar to what we had above) to form beliefs for the new utterance of the Astronaut (B5) before the procedure that follows. The new beliefs based on the Astronaut's new utterance are as follows. Notice that the Robot still believes that the Astronaut's emotion is neutral.

- the precondition associated to 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 fixing the measurement tool,
- also, the Astronaut's *emotion* is still *neutral*.
- but, her user model *has changed* to having medium-autonomy and being highly communicative.

(Collaboration) The change in the focus of attention to fixing the measurement tool causes the Robot to check the availability of a recipe to fix or replace the disfunctioning measurement tool. Similarly to the previous example, 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. The Robot finds the replacement of the measurement tool *relevant*, *desirable*, and *controllable* just as before.

(Motivation: *Motive & Intention Formations*) The Robot forms new motives based on the next task according to the shared plan and the outcome of the appraisal of the possibility of replacing the measurement tool. The Robot forms the corresponding intentions with respect to the new motives.

Once again, the Robot decides to take an action based on the recent intention.

(Coping) Based on the current mental state, first, the Robot responds to the Astronaut's question, and then, chooses 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 answer the Robot tries to negotiate (using the same procedure as we discussed above) with the Astronaut to protect the collaboration and the shared goal from failure. Therefore, the Robot asks about the possibility of pursuing another task.

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!

As shown in this example, 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 which is a crucial point in our first two examples. Although the collaboration mechanism provides the required structural details of collaboration between the Robot and the Astronaut, these structural details are not enough for saving a collaboration from a 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 among 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. The Robot also misses the right time to begin a negotiation process to save the collaboration from failure. Therefore, it causes the Astronaut to reject the Robot's proposal which again aggravates the Astronaut's negative emotion. Consequently, the same collaboration fails even though that the Robot uses the same computational mechanisms, as we showed above. In Sections 4.3 and 4.4, as another example of collaborative behavior, we are going to show the importance of emotional-awareness and its underlying computational mechanisms in a task delegation procedure.

4.3 Task Delegation (Emotion-Awareness)

This walkthrough example is focusing on the delegation of a task by the Astronaut during collaboration (see also the example in Section 2.5). In this example, the explanation between the Astronaut's utterance C1 and the Robot's utterance C2 provides the details of how different mechanisms discussed in Section 3 are involved in making the Robot show collaborative behaviors in acceptance of a new delegated task. To avoid redundant explanations, we refer to similar procedures in previous walkthrough examples.

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.

(Perception) The Robot perceives the Astronaut's utterances and emotion in C1.

The perception mechanism forms beliefs based on the Astronaut's utterances and her emotion (i.e., worry) which she has expressed nonverbally. We have shown some examples of the beliefs formed by the perception mechanism in the example in Section 4.1.

(Collaboration: *Interruption & Constraint Management*) The Robot infers the interruption and uses the constraint management process to retrieve required resources and preconditions. the Robot also checks whether there is an available associated recipe for the delegated task.

(Theory of Mind: *Reverse Appraisal & User Modeling*) The Robot uses reverse appraisal to understand the meaning of the Astronaut's worry with respect to the collaborative task status retrieved in the previous step (e.g., precondition status, postcondition status, required resources, shared goal). the Robot also updates the Astronaut's user model and forms beliefs that a) the Astronaut has *high autonomy*, and b) Astronaut is a highly communicative collaborator. We have discussed more details about reverse appraisal and user modeling processes in our example in Section 4.1.

(Appraisal) The Robot appraises the Astronaut's utterances and emotion. The Robot interprets the Astronaut's utterances and emotional state as a *relevant*, *unexpected*, *undesirable*, *urgent*, but *controllable* event.

(Motivation: *Motive & Intention Formations*) The Robot forms new motives based on the result of the same processes we discussed in Section 4.1, and compares the available motives in the same way we discussed in that section. The Robot forms new intention(s) with respect to the selected motive.

(Coping) Based on the current mental state, the Robot chooses an emotion-focused coping strategy and decides to acknowledge the Astronaut's emotion, and provide a proper response (C2) without asking questions about the delegated task.

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

The Astronaut perceives Robot's acknowledgment of her emotion as well as the Robot's positive response to the Astronaut's delegated task. The Astronaut knows that the Robot needs her to help with some of the primitive tasks in her own delegated task to the Robot. Therefore, the Astronaut, while she is still worried about time, informs the Robot that she will try to finish her current task quickly.

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

The Robot perceives Astronaut's utterance and emotional expression. The same process happens from updating beliefs to taking actions as we discussed above or in previous examples. Since the Robot believes that the Astronaut is still worried about time, it only informs the Astronaut about some potential questions in the future. The Robot knows about these questions since there is either missing information according to the partial plan, or required resources and sub-tasks that can be provided by the Astronaut. The Robot chooses a proper utterance about missing information according to the human's emotion (C4).

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

The Astronaut finds the Robot's response appropriate for the delegated task. Therefore, the Robot's proper response mitigates the Astronaut's negative emotion which was caused by the lack of time for a successive installing procedure of the first and second panels. As a result, the Astronaut properly responds to the Robot's needs (C5).

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

4.4 Task Delegation (Emotion-Ignorance)

This walkthrough example begins with the same exact utterance as the previous one in Section 4.3. This section briefly provides the corresponding details of the emotional-ignorance example in Section 2.6 which is focusing on the delegation of a task by the Astronaut during collaboration. In this example, the explanation between the Astronaut's utterance D1 and the Robot's utterance D2 provides some details to show that even though the same mechanisms (discussed in Section 3) are used to make the Robot obtain collaborative behaviors, ignoring the Astronaut's expressed emotion changes the output of different computational mechanisms (see also Section 4.2) which ultimately causes unsuccessful termination of the task delegation process. To avoid redundant explanations, we group some of the Astronaut and Robot's utterances which constitute the representation of repetitive interaction between them. We also refer to similar procedures in previous walkthrough examples.

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) The Robot only perceives the Astronaut's utterances (D1). These beliefs are about the unsatisfied postcondition of the first task, i.e., installing the first solar panel, and the Astronaut's proposal of installing the second panel. We have shown some examples of the beliefs formed by perception mechanism in Section 4.1. Also, as we have shown in Section 4.2 that the Robot does not perceive the Astronaut's emotion (i.e., worry). Therefore, the Robot misses beliefs about the Astronaut's emotion.

(Collaboration: *Interruption & Constraint Management*) The Robot infers the interruption and uses the constraint management process to retrieve required resources, and preconditions. Similarly to the previous example, the Robot also checks whether there is an available associated recipe for the delegated task. Ignoring the Astronaut's expressed emotion does not change the beliefs formed based on the output of the collaboration mechanism.

(Theory of Mind: *Reverse Appraisal & User Modeling*) Similar to the example in Section 4.2, the Robot uses reverse appraisal to understand the meaning of the Astronaut's emotion with respect to the collaborative task status. However, since the Robot ignores the Astronaut's actual emotion, the output of the reverse appraisal does not help the Robot's inference about its own collaborative behavior (see Section 4.2). Similarly, the Robot updates the Astronaut's user model based on wrong beliefs achieved by ignoring the Astronaut's expressed emotion. We have discussed more details about reverse appraisal and user modeling processes in our example in Section 4.1, and we have shown similar effects in Section 4.2.

(Appraisal) The Robot appraises the Astronaut's utterances with a wrong assumption of her expressing neutral emotion. Consequently, similar to the example in Section 4.3, the Robot interprets the Astronaut's utterances as a *relevant*, *unexpected*, *undesirable*, *urgent*, and *controllable* event.

(Motivation: *Motive & Intention Formations*) The Robot forms and compares new motives in the same way as we have discussed in previous examples. However, both processes of the motive formation and comparison are performed based on false outputs of the other processes (see also Section 4.2). Therefore, the Robot forms a new intention with respect to the selected motive which can be formed or selected for a wrong reason.

(Coping) Based on the current mental state (including a wrong intention), the Robot decides to use problem-focused coping strategy of seeking social support (seeking information) to be able to drop the intention of working on the same task as the Astronaut. Respectively, Robot asks a question (D2).

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

The rest of the Robot's utterances from D4 to D10 shows 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. On the other hand, the Astronaut who was worried about the lack of time since the beginning, tries to succinctly respond the Robot and successfully delegate the task. However, the Robot, by asking several questions instead of understanding the Astronaut's mental state and accepting the delegated task, causes the Astronaut to quickly become frustrated and express her frustration (D9).

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?

The Robot perceives the Astronaut's last utterance and once again ignores her emotion. This time the Robot's tries to provide some examples to clarify its own point which is another divergence from achieving the shared goal or planning for the new delegated task (D10). Finally, the Astronaut terminates the collaboration due to the lack of time and the Robot's failure in incorporating proper collaborative behavior (D11).

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

5 Related Work

The prominent collaboration theories are mostly based on plans and joint intentions [13, 25, 41], and they were derived from the BDI paradigm developed by Bratman [4] which is fundamentally reliant on folk psychology [62]. The two theories, Joint Intentions [13] and SharedPlans [25], have been extensively used to examine and describe teamwork and collaboration. There are many research focusing on different aspects of collaboration based on different collaboration theories, i.e., SharedPlans [22, 23, 25], Joint Intentions [13], and hybrid theories of collaboration, e.g., STEAM [77]. 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. In the rest of this section, we provide a review of several applications of different prominent collaboration theories which puts emphasis on the importance of the collaborative robots and their applications. And at the end, we also provide some related works on applications of artificial emotions and appraisal theory to express the importance of their applicability in robots and autonomous agents.

There are some works focusing on the concepts of robot assistants [12], or teamwork and its challenges in cognitive and behavioral levels [54, 68]. Some researchers have an overall look at a collaboration concept at the architectural level. In [16] authors present a collaborative architecture, COCHI, and argue the need to support emotional-awareness in the design and implementation of groupwares. In [15] authors present the integration of emotional competence into a cognitive architecture which runs on a robot, MEXI. In [75] authors discuss the challenges of integrating natural language, gesture understanding and spatial reasoning of a collaborative humanoid robot situated in space. The importance of communication during collaboration has also been considered by some researchers from human-computer interaction and human-robot collaboration [11, 50, 65] to theories describing collaborative negotiation, and discourse planning and structures [2, 24, 73]. There are other concepts such as joint actions and commitments [21], dynamics of intentions during collaboration [40], and task-based planning providing more depth in the context of collaboration [9, 63]. The concept of collaboration has also received attention in the industry and in research in robotic laboratories [20]. Some of these works emphasize the applicability of emotions in their architectures, and some others emphasize the collaborative aspect of their robots. In the following pages, we review some of the applications of each prominent collaboration theory.

Applications of SharedPlans Theory – COLLAGEN [64,65] is the first implemented system based on the SharedPlans theory. It incorporates certain algorithms for discourse generation and interpretation, and is able to maintain a segmented interaction history, which facilitates the discourse between the human user and the intelligent agent. The model includes two main parts: (1) a representation of a discourse state and (2) a discourse interpretation algorithm for the utterances of the user and agent [66]. In [26] Heeman presents a computational model of how a conversational participant collaborates in order to make a referring action successful. The model is based on the view of language as goal-directed behaviour, and in his work, he refers to SharedPlans as part of the planning and conversation literature. In [44], Lochbaum and Sidner modify and expand the SharedPlan model of collaborative behavior [25]. They present an algorithm for updating an agent's beliefs about a partial shared plan and describe an initial implementation of this algorithm in the domain of network management. Lochbaum, also in [43], provides a computational model (based on the collaborative planning framework of SharedPlans [23]) for recognizing intentional structure and utilizing it in discourse processing. In short, she presents a SharedPlans model for recognizing Discourse Segment Purposes (DSPs) [25] [73] and their interrelationships. CAST (Collaborative Agents for Simulating Teamwork) [85] [86] is a teamwork framework based on the SharedPlans theory. CAST focuses on flexibility in dynamic environments and on proactive information exchange enabled by anticipating what information team members will need. Petri Nets are used to represent both the team structure and the teamwork process, i.e., the plans to be executed. Researchers in [29] discuss developing an ontology of microsocial concepts for use in an instructional system for teaching cross-cultural communication. They believe being acquainted with one another is not a strong enough relationship from which to create a society. Hence, there is a need for commitment and shared plans (as the basis of social life) to achieve a shared goal. In this work, Grosz and Sidner's SharedPlans theory [25] is used to explain the concept of shared plans within the interpersonal relationships of societies in an industrial environment. In [31] Hunsberger and Grosz discuss the idea of whether the rational, utility-maximizing agents should determine commitment to a group activity when there is an opportunity to collaborate. They call this problem the "initial-commitment decision problem" (ICDP) and provide a mechanism that agents can use to solve the ICDP. They use the representation of action, act-types and recipes in the SharedPlans theory. In [87] an integrated agent-based model for Group Decision Support Systems is proposed and discussed. The decisional model that authors outline in this paper is based on the SharedPlans theory. Rauenbusch and Grosz in [61] formally define a search problem with search operators that correspond to the team planning decisions. They provide an algorithm for making the three types of interrelated decisions by recasting the problem as a search problem. Their model respects the constraints on mental states specified by the SharedPlans theory of collaboration.

Babaian et. al. in [3] describe Writer's Aid, a system that deploys AI planning techniques to enable it to serve as an author's collaborative assistant. While an author writes a document, Writer's Aid helps in identifying and inserting citation keys and by autonomously finding and caching potentially relevant papers and their associated bibliographic information from various on-line sources. They believe the underlying concepts of SharedPlans is relevant since in collaborative interfaces like Writer's Aid, the users establish shared goals with the system and user and the system both take initiative in satisfying them. In [52] researchers address high-level robot planning issues for an interactive cognitive robot that acts in the presence of or in collaboration with a human partner. They describe a Human Aware Task Planner (HATP) which is designed to provide socially acceptable plans to achieve collaborative tasks. They use notions of plans based on SharedPlans theory. In [74] Sidner and Dzikovska argue that robots, in order to participate in conversations with humans, need to make use of conventions of conversation and the means to be connected to their human counterparts. They provide an initial research on engagement in human-human interaction and applications to stationary robots in hosting activities. They believe hosting activities are collaborative because neither party completely determines the goals to be undertaken nor the means of reaching the goal. To build a robot host, they rely on an agent built using COLLAGEN which is implemented based on the SharedPlans theory.

Applications of Joint Intentions Theory – In [36] authors introduce a language for representing joint plans for teams of agents. They describe how agents can organize the formation of a suitably skilled team to achieve a joint goal, and they explain how such a team can execute these plans to generate complex, synchronized team activity. In this paper, authors adopt the underlying concepts of the Joint Intentions theory as the structure of their collaborative agents. Breazeal et. al. in [8] present an overview of their work towards building socially intelligent, cooperative humanoid robots, Leonardo, that can collaborate and learn in partnership with humans. They employ the Joint Intentions theory of collaboration to implement the collaborative behaviors while performing a task in collaboration with humans. In [76] the researchers' goal is to develop an architecture (based on the concepts of Joint Intentions theory) that can guide an agent during collaborative teamwork. They describe how a joint intention interpreter that is integrated with a reasoner over beliefs and communicative acts can form the core of a dialogue engine. Ultimately, the system engages in dialogue through the planning and execution of communicative acts necessary to attain the collaborative task at hand. Mutlu et. al. in [53] discuss key mechanisms for effective coordination toward informing the design of communication and coordination mechanisms for robots. They present two illustrative studies that explore how robot behavior might be designed to employ these mechanisms (particularly joint attention and action observation) to improve measure of task performance in human-robot collaboration. Their work uses Joint Intentions theory to develop shared task representations and strategies for task decomposition. The system GRATE* by Jennings [33] is based on the Joint Intention theory. GRATE* provides a rule-based modelling approach to cooperation using the notion of Joint Responsibilities, which in turn is based on Join Intentions. GRATE* is geared towards industrial settings in which both agents and the communication between them can be considered to be reliable.

Applications of Hybrid Theories – The domain independent teamwork model, STEAM, has been successfully applied to a variety of domains. From combat air missions [28] to robot soccer [38] to teams supporting human organizations [60] to rescue response [69]. In [46] authors provide their RoboCup (robotics soccer testbed) in which their focus is on teamwork and learning challenges. Their research investigation in RobotCup is based on ISI Synthetic, a team of synthetic soccer-players. They also investigate the use of STEAM as their model of teamwork. In [34] researchers propose a behavioral architecture C^2BDI that allows the enhancement of the knowledge sharing using natural language communication between team members. They define collaborative conversation protocols that provide proactive behavior to agents for the coordination between team members. Their agent architecture provides deliberative and conversational behaviors for collaboration, and it is based on both of the SharedPlans and Joint Intentions theories.

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 [83]. Some of these works are inspired by specific psychological theories (we provide several examples in this section), some are freely using the concept of emotion without using the theoretical background in social sciences, and some are using a combination of concepts from the psychological theories. For instance, in PECS [79] which is designed for modeling human behaviors, the agent's architecture is not based on a certain kind of social or psychological emotion theory. In fact, it is intentionally designed and described in a way which enables the integration of a variety of theories. The PECS' design enables an integrative modeling of physical, emotional, cognitive and social influences within a component-oriented agent architecture. Also, in [49] the computational architecture which is designed to provide information about the possible overall behavior of a work team is not based on any specific theory. Some researchers apply combinations of emotion theories in their work [37]. For instance, in [10] Cañamero shows how an agent can use emotions for activity selection while taking into account both dimensional and discrete approaches in an action selection mechanism. 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 [7] [39] [56] [72]. Robots also use emotions theories for automatic affect recognition using different modalities [27] [88]. Moreover, in some works, researchers have explored the user's affective state as a mechanism to adapt the robot's behaviors during the interaction [6] [42].

Applications of Appraisal Theory – The emphasis of models derived from appraisal theories of emotion is on making appraisal the central process. 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 [48] 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 [19] [47]. There are several examples in artificial intelligence and robotics of applying appraisal theory [1] [35] [48]. In robotics, appraisal theory has been used to establish and maintain a better interaction between a robot and a human. For instance in [35] researchers provide their computational model of emotion generation based on appraisal theory to have a positive human-robot interaction experience. In [67] authors describe a system approach to appraisal processes based on Scherer's work on appraisal and the Component Process Model [70]. They show how the temporal unfolding of emotions can be experimentally tested. They also lay out a general domain-independent computational model of appraisal and coping. In [82] researchers consider their robot's (INDIGO) emotion, speech and facial expressions as a key point to establish an effective communication between the robot and a human during their interaction. They apply concepts of appraisal theory in INDIGO's emotion modeling. MAGGIE, a sociable robot, also applies the appraisal theory of emotions to consider fear in its decision making system [18]. Velasquez developed Cathexis which is a distributed computational model for generation of emotions and their influence in the behavior of the autonomous agents [81]. The emotion model in this work is based on Roseman's work on appraisal theory. Marinier and Laird in [32] focus on the functional benefits of emotion in a cognitive system. In this work, they integrate their emotion theory (which is based on the appraisal theory) with Soar cognitive architecture, and use emotional feedback to drive reinforcement learning. In [30] Hudlicka provides a model of a generic mechanism mediating the affective influences on cognition based on cognitive appraisal. This model is implemented within a domain-independent cognitive-affective architecture (MAMAID). In the virtual agents community, empathy is a research topic that has received much attention in the last decade [5] [51] [55] [59] [78]. In [58] researchers developed an agent with capability of affective decision-making based on appraisal theory to establish an affective relationship with its users. Then, they compared the performance of their agent with a human (based on a WoZ study) in a speed-dating experiment.

6 Conclusion and Future Work

In this paper, we argued the missing part of computational collaboration theories and the need for a theory explaining the underlying processes involved in a collaboration. Then, we discussed the importance of social functions of emotions and how emotions are involved in social contexts to reveal or enrich the meaning of interactants' messages through communication. In fact, there is a correspondence between what a collaboration needs and what social functions of emotions provide within a social context. Next, in Section 2, we provided four hypothetical examples in two pairs. Each pair of examples was about a distinct collaborative behavior. The first pair was about agreeing on a shared goal between a robot and an astronaut, and the second pair was about delegation of a new task to the robot by the astronaut. Each pair of these examples contains a successful collaboration because the robot was aware of the astronaut's emotion, and a failure in collaboration as the consequence of the robot ignoring the astronaut's emotions. We provided a brief description for each example as well as the utterances of both the robot and the astronaut during their collaboration. These examples illustrated the importance of emotional-awareness to attain successful collaborative behavior. Then, in Section 3, we continued by briefly introducing the main components of *Affective Motivational Collaboration Thoery* as our computational framework which incorporates emotion-regulated mechanisms. This framework let us describe the same examples in more detail (in Section 4) by explaining certain mechanisms and their underlying processes, and how they are involved in helping the robot obtain a collaborative behavior by observing the astronaut's emotions. At last, in Section 5, we provided some related works which use computational collaboration and emotion theories.

We have implemented the rules associated with these examples using JESS (Java Expert System Shell) which is a rule engine for the Java platform. In our current implementation we have categorized the rules in different modules associated with the mechanisms and the underlying processes in *Affective Motivational Collaboration Thoery* (see Fig. 1). In our future work, we will implement algorithms for each individual mechanism to be able to automatically generate the required facts within each mechanism to fire the existing rules. Ultimately, we are going to offer a platform which operates based on the collaboration structure discussed by SharedPlans theory [24], and employs emotion-driven processes such as the appraisal process in [48] to enable a robot to obtain and demonstrate collaborative behaviors.

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. J. Andriessen, K. de Smedt, and M. Zock. Discourse planning: Empirical research and computer models. In A. Dijkstra and K. de Smedt, editors, *Computational psycholin-*

guistics: AI and connectionist models of human language processing, pages 247–278. Taylor & Francis, 1996.
3. T. Babaian, B. J. Grosz, and S. M. Shieber. A writer's collaborative assistant. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI2000)*, pages 7–14. ACM Press, 2002.
4. M. E. Bratman. *Intention, Plans, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 1987.
5. S. Brave and C. Nass. Emotion in human-computer interaction. In J. A. Jacko and A. Sears, editors, *The Human-computer Interaction Handbook*, pages 81–96. L. Erlbaum Associates Inc., 2003.
6. C. Breazeal. *Designing Sociable Robots*. MIT Press, 2002.
7. 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.
8. 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.
9. C. Burghart, R. Mikut, R. Stiefelhagen, T. Asfour, H. Holzapfel, P. Steinhaus, and R. Dillmann. A cognitive architecture for a humanoid robot: A first approach. In *5th IEEE-RAS International Conference on Humanoid Robots*, pages 357–362, 2005.
10. L. D. Canamero. Designing emotions for activity selection in autonomous agents. In R. Trappl, P. Petta, and S. Payr, editors, *Emotions in Humans and Artifacts*, pages 115–148. MIT Press, 2003.
11. A. B. S. Clair and M. J. Matarić. Modeling action and intention for the production of coordinating communication in human-robot task collaborations. In *21st IEEE International Symposium on Robot and Human Interactive Communication: Workshop on Robot Feedback in HRI*, Paris, France, 2012.
12. 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.
13. P. Cohen and H. J. Levesque. *Teamwork*. SRI International, 1991.
14. 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.
15. 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.
16. 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.
17. E. Goffman. *The Presentation of Self in Everyday Life*. Anchor, 1959.
18. 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.
19. 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.
20. S. A. Green, M. Billinghurst, X. Chen, and J. G. Chase. Human-robot collaboration: A literature review and augmented reality approach in design. *International Journal of Advanced Robotic Systems*, 5(1):1–18, 2008.
21. B. J. Grosz and L. Hunsberger. The dynamics of intention in collaborative activity. *Cognitive Systems Research*, 7(2-3):259–272, 2007.
22. B. J. Grosz, L. Hunsberger, and S. Kraus. Planning and acting together. *AI Magazine*, 20(4):23–34, 1999.
23. B. J. Grosz and S. Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.
24. B. J. Grosz and C. L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, July 1986.

25. 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.
26. P. A. Heeman. *A Computational Model of Collaboration on Referring Expressions*. PhD thesis, University of Toronto, 1991.
27. F. Hegel, T. Spexard, B. Wrede, G. Horstmann, and T. Vogt. Playing a different imitation game: Interaction with an empathic android robot. In *Proceedings of 2006 IEEE-RAS International Conference on Humanoid Robots (Humanoids06)*, 2006.
28. R. W. Hill, Jr., J. Chen, J. Gratch, P. Rosenbloom, and M. Tambe. Intelligent agents for the synthetic battlefield: A company of rotary wing aircraft. In *Innovative Applications of Artificial Inttigence (IAAI-97)*, pages 227–262, 1997.
29. J. R. Hobbs, A. Sagae, and S. Wertheim. Toward a commonsense theory of microsociology: Interpersonal relationships. In *Formal Ontology in Information Systems - Proceedings of the Seventh International Conference*, pages 249–262, 2012.
30. 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.
31. L. Hunsberger and B. J. Grosz. A combinatorial auction for collaborative planning. In *In Proceedings of ICMAS*, 2000.
32. R. P. M. III and J. E. Laird. Emotion-driven reinforcement learning. In *CogSci 2008*, 2008.
33. N. R. Jennings. Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence*, 75(2):195–240, 1995.
34. 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.
35. 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.
36. D. Kinny, M. Ljungberg, A. Rao, G. Tidhar, E. Werner, and E. Sonenberg. Planned team activity. In *Lecture notes in artificial intelligence*. Springer-Verlag, 1992.
37. 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.
38. 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.
39. I. Leite, A. Pereira, S. Mascarenhas, C. Martinho, R. Prada, and A. Paiva. The influence of empathy in human-robot relations. *International Journal of Human-Computer Studies*, 71(3):250–260, 2013.
40. 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.
41. 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.
42. 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.
43. K. E. Lochbaum. A collaborative planning model of intentional structure. *Computational Linguistics*, 24(4):525–572, 1998.
44. 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.
45. 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.
46. 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.

47. S. Marsella, J. Gratch, N. Wang, and B. Stankovic. Assessing the validity of a computational model of emotional coping. In *International Conference on Affective Computing and Intelligent Interaction*. IEEE, 2009.
48. S. C. Marsella and J. Gratch. EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1):70–90, March 2009.
49. J. Martínez-Miranda, A. Aldea, and R. Bañares-Alcántara. Simulation of work teams using a multi-agent system. In *The Second International Joint Conference on Autonomous Agents & Multiagent Systems AAMAS, July 14-18, Melbourne, Victoria, Australia*, pages 1064–1065, 2003.
50. L. Matignon, A. B. Karami, and A.-I. Mouaddib. A model for verbal and non-verbal human-robot collaboration. In *AAAI Fall Symposium Series*, pages 62–67, 2010.
51. 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.
52. 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.
53. 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.
54. 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.
55. 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.
56. 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.
57. S. Planalp. *Communicating Emotion: Social, Moral, and Cultural Processes*. Cambridge University Press, 1999.
58. 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.
59. H. Prendinger and M. Ishizuka. The empathic companion: a character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19(3-4):267–285, 2005.
60. D. V. Pynadath and M. Tambe. An automated teamwork infrastructure for heterogeneous software agents and humans. *Journal of Autonomous Agents and Multi-Agent Systems, Special Issue on Infrastructure and Requirements for Building Research Grade Multi-Agent Systems,*, 7(1-2):71–100, 2003.
61. T. W. Rauenbusch and B. J. Grosz. A decision making procedure for collaborative planning. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, 2003.
62. I. Ravenscroft. *Folk Psychology as a Theory*. Stanford Encyclopedia of Philosophy, 2004.
63. C. Rich. Building task-based user interfaces with ANSI/CEA-2018. *IEEE Computer*, 42(8):20–27, July 2009.
64. 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.
65. C. Rich, C. L. Sidner, and N. Lesh. COLLAGEN: Applying collaborative discourse theory to human-computer interaction. *AI Magazine*, 22(4):15–26, 2001.
66. 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.
67. D. Sander, D. Grandjean, and K. R. Scherer. A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18(4):317–352, 2005.

68. 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.
69. 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.
70. 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.
71. K. R. Scherer, A. Schorr, and T. Johnstone. *Appraisal Processes in Emotion: Theory, Methods, Research.* Oxford University Press, 2001.
72. M. Shayganfar, C. Rich, and C. L. Sidner. A design methodology for expressing emotion on robot faces. In *IROS*, pages 4577–4583. IEEE, 2012.
73. C. Sidner. An artificial discourse language for collaborative negotiation. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 814–819. MIT Press, 1994.
74. 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.
75. 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.
76. R. A. Subramanian, S. Kumar, and P. Cohen. Integrating joint intention theory, belief reasoning, and communicative action for generating team-oriented dialogue. In *AAAI*, pages 1501–1507. AAAI Press, 2006.
77. M. Tambe. Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:83–124, 1997.
78. M. Tambe. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327, 2005.
79. C. Urban. Pecs: A reference model for human-like agents. In *Deformable Avatars*. Netherlands: Kluwer Academic Publishers, 2001.
80. S. van Hooft. Scheler on sharing emotions. *Philosophy Today*, 38(1):18–28, 1994.
81. 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.
82. 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.
83. T. Wehrle. Motivations behind modeling emotional agents: Whose emotion does your robot have?, 1998.
84. A. K. Wisecup, D. T. Robinson, and L. Smith-Lovin. *The Sociology of Emotions*. SAGE Publications, Inc., 2007.
85. 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.
86. 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.
87. Zamfirescu and Candea. On integrating agents into gdss. In *Preprints of the 9th IFAC / IFORS / IMACS / IFIP/ Symposium on Large Scale Systems: Theory and Applications*, 2001.
88. 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.