Affective Motivational Collaboration Theory

by

Mohammad Shayganfar (mshayganfar@wpi.edu)

Dissertation Proposal

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

DOCTOR OF PHILOSOPHY

in

Computer Science

February 2015

Professor Charles Rich, Thesis Advisor

Professor Candace L. Sidner, Thesis Advisor

Professor Stacy C. Marsella, Thesis Committee Member

Professor John E. Laird, Thesis Committee Member

ABSTRACT

In this thesis, I investigate the mutual influence of affective and collaboration processes in a cognitive theory to support the interaction between humans and robots or virtual agents. I will develop new algorithms for these processes, as well as a new overall computational model for implementing collaborative robots and agents. I build primarily on the *cognitive appraisal* theory of emotions and the *SharedPlans* theory of collaboration to investigate the structure, fundamental processes and functions of emotions in a collaboration context. As part of this work, I also address a deficiency in existing cognitive models by accounting for the influence of motivation on collaborative behaviors, such as overcoming an impasse. This motivation mechanism uses the results of cognitive appraisal to dynamically form new beliefs and intentions related to the collaboration structure.

Contents

List of Figures

CHAPTER 1

INTRODUCTION

In this thesis, I investigate the mutual influence of affective and collaboration processes in a cognitive theory to support the interaction between humans and robots or virtual agents. I will develop new algorithms for these processes as well as a new overall computational model for implementing collaborative robots and agents. I build primarily on the *cognitive appraisal theory* [96, 109, 147] of emotions and the *SharedPlans theory* [74] of collaboration to investigate the structure, fundamental processes and functions of emotions in a collaboration context. My work is also inspired by other theories in psychology and artificial intelligence that will be discussed in detail later in this document.

1.1 Motivation

In *The Society of Mind*, Marvin Minsky says: "The question is not whether intelligent machines can have emotions, but whether machines can be intelligent without any emotions." [115] Ronald De Sousa in *The Rationality of Emotion* [164] makes a good case for the claim that humans are capable of rationality largely because they are creatures with emotions.

The idea of having robots or other intelligent agents living in a human environment has been a persistent dream from science fiction books to artificial intelligence and robotics laboratories. However, there are many challenges in achieving collaboration between robots and humans in the same environment. Some of these

challenges involve physical requirements, some involve cognitive requirements, and some involve social requirements. Thus far, there has been an emphasis on the design of robots to deal with the physical requirements. Many researchers are also working on the cognitive requirements, inspired by a diverse set of disciplines. As time passes, there has been an increasing recognition of the importance of the social requirements.

One aspect of the sociability of robots and agents is their ability to collaborate with humans in the same environment. Therefore, it is important to understand what makes a collaboration effective. One's cognitive processes and the ability to understand the collaborative environment impact the effectiveness of a collaboration. Examples of cognitive capabilities that support the effectiveness of collaboration include: a) perceiving one's own internal states and b) communicating them, c) coordinating personal and group behaviors, d) identifying self and mutual interests, e) recognizing the accountability of private and shared goals, f) selecting appropriate actions with respect to events, and g) engaging others in collaboration.

Collaboration is a coordinated activity in which the participants work jointly to satisfy a shared goal. There are many important unanswered questions about the involvement of an individual's cognitive abilities during collaboration. Some of these questions are related to the dynamics of collaboration, as well as the underlying mechanisms and processes. For instance, a general mechanism has yet to be developed that allows an agent to initiate proactive collaborative behaviors when it faces a blocked task. There is also lack of a general mechanism that, in the event of a task failure, allows an agent to consider the human's anticipated mental states and emotions, while managing its own internal goals as well as the collaboration's shared goal. There are also other questions about the components involved in these processes at the cognitive level, such as the processes that are involved for evaluative, regulatory or motivative purposes as internal processes of the agent. There has also not been enough attention on the processes that are involved to maintain the social aspects of a collaboration.

Emotions have a key role in influencing the cognitive processes involved in social interaction and collaboration. Psychologists claim that emotions contribute to the modulation of assessment, selection and the evaluation of options (alternative decisions). For instance, a sad mood may result from an undesired outcome, but it can also lead to increased salience of negative attributes of options in future assessments. Emotion processing and decision-making are integral aspects of daily life and maintain their prominence during social interaction and collaboration. However, researchers' understanding of the interaction between emotions and collaborative behaviors is limited. For instance, there are situations in which the best decision would not necessarily be optimal with respect to the private or the shared goal during the collaboration. It might be more favorable for the collaborators to make decisions within an approximation range of the optimal decision. Such an approach might help the collaboration to continue when there are similar utilities for different actions within the optimal range.

There are also situations during a collaboration in which conflicts occur between the collaborators' decisions towards achieving their private versus shared goals. Conflicts are an inherent part of collaboration. Frequently, proposed solutions to these conflicts are characterized as infeasible because they are believed to be in conflict with collaborators' shared interests. Furthermore, execution of an ongoing task could become blocked or protracted for various reasons, such as the failed pre- or post-conditions of a required task. A task blocking could also be caused by a lack of required information about the way a task should be done, which may have originated from an incomplete plan. It is important to investigate the underlying cognitive mechanism that can help humans to establish, maintain and conclude a successful and effective collaboration in such conditions. I am interested to determine how affective components enter into these decision making processes to minimize or avoid task blocking, unsatisfactory progress during collaboration, or termination without a satisfactory outcome.

I believe that the evaluative role of emotions as a part of cognitive processes helps

an agent to perform appropriate behaviors during a collaboration. It is important to think about the underlying cognitive processes of the collaborators in order to have a better understanding of the role of emotions. To work jointly in a coordinated activity, participants (collaborators) act based on their own understanding of the world and the anticipated mental states of the counterpart; this understanding is reflected in their collaborative behaviors. Emotions are pivotal in the collaboration context, since their regulatory and motivative roles enhance an individual's autonomy and adaptation as well as his/her coordination and communication competencies in a dynamic, uncertain and resource-limited environment. The collaborative behavior of the individuals can also be influenced by the tasks contributing towards a shared goal. Some tasks may be inherently insignificant, boring, unpleasant or arduous for a collaborator. Thus, knowing how to externally motivate other collaborator to perform such tasks becomes an essential skill for a participant in a successful collaboration. Such knowledge enables an individual to lead his collaborator to internalize the responsibility and sense of value for an externally motivated task.

1.2 Thesis Statement and Scope

In my thesis, I will develop and validate an *Affective Motivational Collaboration Theory* and an associated computational model that will enhance the performance and effectiveness of collaboration between agents and humans. This theory will explain the functions of emotions in a dyadic collaboration and show how affective mechanisms can coordinate social interactions by anticipating other's emotions, beliefs and intentions. This theory will also specify the influence of the underlying collaboration processes on appraisals. *Affective Motivational Collaboration Theory* will elucidate the role of motives as goal-driven emotion-regulated constructs with which an agent can form new beliefs and intentions to cope with internal and external events. An important contribution of this work is to elucidate how motives are involved not only in the appraisal and coping processes, but how they also serve

as a bridge between appraisal processes and the collaboration structure. Finally, I will validate my theory using the associated computational model in the context of a simulated human-robot collaboration.

I focus only on two-participant collaboration; teamwork collaboration is out of my scope. From another perspective, my theory is bounded by the kinds of internal events, e.g., generation of a new motive, and external events, e.g., human's utterances, it deals with. Finally, my theory focuses on a) the influence of emotion-regulated processes on the collaboration structure, and b) prediction of the observable behaviors of the other during a collaborative interaction.

I describe the cognitive processes involved in a collaboration in the context of a cognitive architecture. There are several well-developed cognitive architectures, e.g., Soar [92] and ACT-R [83], each with different approaches to defining the basic cognitive and perceptual operations. There have also been efforts to integrate affect into these architectures [39, 108]. In general, however, these cognitive architectures do not focus on processes to specifically produce emotion-regulated goal-driven collaborative behaviors. At the same time, existing collaboration theories, e.g., SharedPlans [74] theory, focus on describing the structure of a collaboration in terms of fundamental mental states, e.g., mutual beliefs or joint intentions. However, they do not describe the associated processes, their relationships, and influences on each other. *Affective Motivational Collaboration Theory* deals with all of the major processes, including affective and motivational processes, having an impact on the collaboration structure. This theory is informed by research in psychology and artificial intelligence which is reviewed in Chapter 2. My contribution, generally speaking, is to synthesize prior work on motivation, appraisal and collaboration, and thus to provide a new theory which describes the prominent emotion-regulated goal-driven phenomena in a dyadic collaboration.

1.3 Contributions of the Thesis

1. **Developing new computational models and algorithms for *Affective Motivational Collaboration Theory*; specifically for the**:

   (a) **appraisal of internal and external events during collaboration**

   I will create computational models and algorithms to compute the value of appraisal variables in a dyadic collaboration. Applying cognitive appraisal theory in the collaboration context is novel. Other models of the appraisal theory have not paid attention to the dynamics of the collaborators' mental states and behaviors.

   (b) **functions of emotions during collaboration**

   I will develop new algorithms for the various functions of emotions specifically in the context of collaboration. Existing models and implementations of emotions focus only on how emotions regulate and control internal processes and sometimes behaviors. In my work, I also investigate the details of how emotions result in collaborators' mental states. This in turn sheds light on the dynamics of collaborators' mental states and how appraisal of the self and the environment contributes to these functions.

   (c) **motivation-based belief and intention formation**

   I will create a motivation mechanism that employs appraisal processes to dynamically form new beliefs and intentions related to the collaboration structure. This mechanism will help the agent overcome impasses during collaboration. Existing cognitive models do not include an account of the influence of motivation on collaborative behaviors.

2. **Developing and implementing a computational model based on *Affective Motivational Collaboration Theory*:**

   My computational model will implement the key algorithms related to *Affective Motivational Collaboration Theory* as well as minimal implementation of other processes which are required for validation of the model but are not part of my thesis contributions. The emphasis of the model is on underlying cognitive processes embracing collaboration and appraisal concepts, rather than the Perception and the Action mechanisms.

3. **Validating *Affective Motivational Collaboration Theory:***

   I have identified eight key social characteristics (see Section 3.5) which occur during the course of a collaboration. I will validate how the various functions of emotions give rise to these characteristics during collaboration. Specifically, I will first incrementally validate one or more of the computational components in my model starting with appraisal. Finally, I will conduct an end-to-end system evaluation with human subjects and a simulated robot.

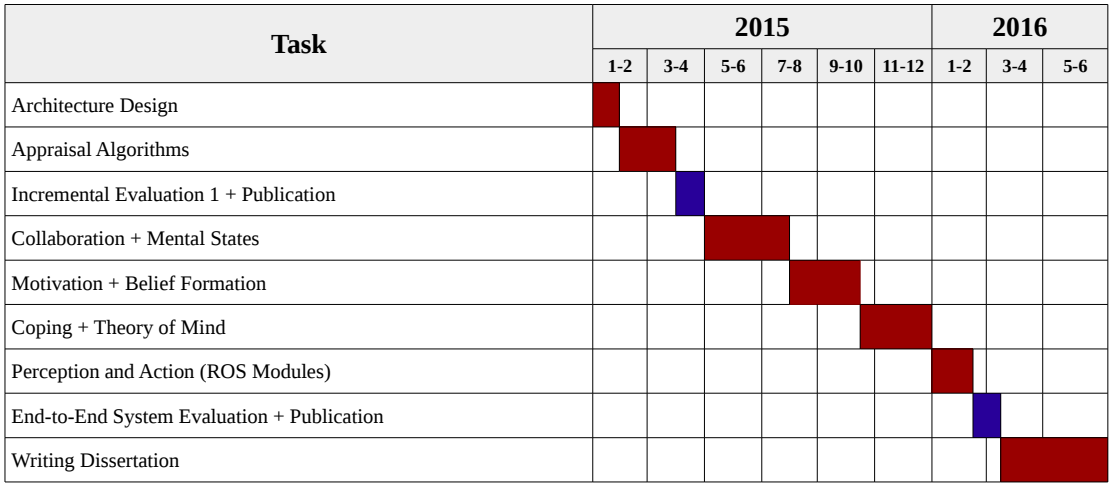

Figure 1.1: PhD thesis schedule.

1.4 Timetable

As shown in Figure 1.1, the implementation of my thesis will begin with architecture design and building models of appraisal variables (see Figure 3.1 for roadmap to the mechanisms referred to in this timeline). I will then run an online study (discussed in Section 6.1) to validate the computational model of the Appraisal mechanism. The results will be published at this point. Then I will implement the collaboration mechanism and the structure of Mental States as part of the computational model. The implementation of the computational model will continue by developing the computational model of the Motivation mechanism as well as Belief and Intention Formation mechanisms. The last part of the computational model will include the development of the Coping mechanism and the Theory of Mind. Depending on the time available, I will then run additional online studies to incrementally validate additional mechanisms such as Motivation and Collaboration (see Section 6.1). The final part of my work will be a human study based on the complete end-to-end system including simplified Perception and Action mechanisms. The results of this study also will be published. Finally, I will complete writing my dissertation.

1.5 Structure of Thesis Proposal

This thesis proposal continues by describing background and related work in Chapter 2. The sections in Chapter 2 are organized based on the role of emotions and motives within the context of collaboration. Related theoretical topics are reviewed in fields including philosophy, psychology, neuroscience, artificial intelligence and robotics. Chapter 3 describes *Affective Motivational Collaboration Theory* in detail. This chapter begins with a scenario involving collaboration between a Robot and an Astronaut which will be used throughout the thesis to provide examples clarifying various concepts. Then Chapter 3 provides a general argument about this thesis from a theoretical point of view. The remaining sub-sections discuss details of

different aspects of the theory. Chapter 4 starts with four computational examples drawn from the scenario introduced in Chapter 3. Chapter 4 primarily describes the computational model I have developed based on *Affective Motivational Collaboration Theory*. Chapter 5 briefly provides information about the computational model. Finally, Chapter 6 describes how I plan to evaluate parts of my work incrementally, as well as a final end-to-end system evaluation.

CHAPTER 2

BACKGROUND AND RELATED WORK

In this chapter, I discuss the social and communicative aspects of emotions from a psychological point of view. Understanding the social aspects of emotions is important in my work, since it is focused on collaboration which is a social phenomenon in human environments. I also present the concept of artificial emotions and provide some examples of the existing computational models of emotions. Then, I provide background about the cognitive appraisal theory of emotions as one of two underlying theories in my work as well as related concepts such as some examples of cognitive architectures and the influence of affect in decision-making procedures. This chapter continues with the description of motives and the related theories in psychology and artificial intelligence. The role of motives as goal-driven affective components is crucial in my work, since the collaboration structure is built based on the concept of a shared goal between collaborators. This chapter also contains the background about collaboration theory which is the second foundation for my work. It provides both theoretical and practical related works about the collaboration concept. Finally, a brief description and the related work in psychology and artificial intelligence is provided about the concept of theory of mind.

2.1 Affect and Emotion

Emotion affects not only what people do, but also the way they do it [38]. Aristotle in *The Nicomachean Ethics* reveals his idea about emotions. He says "Anyone can

become angry–that is easy. But to be angry with the right person, to the right degree, at the right time, for the right purpose, and in the right way–this is not easy [5]."

Intelligence is the process that humans use to explain the different degrees of adaptive success in one's behavior. It is a set of mental abilities that enables a human to comprehend, reason and adapt in the environment, and as a result, act effectively and purposefully in that environment. Emotions play a crucial role in humans' explanation of intelligent behaviors. Emotions significantly impact the procedures of action generation, execution, control, and interpretation [176] in different environments. Emotions are conceptualized as ongoing processes rooted in dynamic social contexts, which can shape both implicit and explicit emotional responses [107]. An emotion is a dynamic episode that not only makes changes in cognitive states, but also produces a sequence of response patterns on body movements, posture, voice and face [146]. Emotions typically occur in response to an event, usually a social event, real, remembered, anticipated, or imagined. They are associated with distinctive relational meanings [126]. These relations can be with the individual's past experience, the individual's surrounding objects and environment, or the other individuals with or without mutual beliefs in a dyadic or a group setting. Emotions are evaluative and responsive patterns that serve the function of providing appraisal about whether the ongoing event is harmful, threatening or beneficial for the well-being of an individual [176]. Consequently, reasoning and emotional processes have an integral and a supportive relationship, rather than an antagonistic and a conflicting one.

A better question than what emotions are, is the question of what they can do, and how they impact humans' life. Studies show that the decision making of humans is not always logical [69], and in fact, not only is pure logic not enough to model human intelligence, but it also shows failures when applied in artificial intelligence systems [48]. Emotions impact fundamental parts of cognition including perception, memory, attention and reasoning [32]. This impact is caused by the

information emotions carry about the environment and event values. The influence of emotions depends on an individual's focus of attention. For instance, a positive affect can cause a positive attitude towards an object if the individual's focus is on the object, whereas the same positive affect can be interpreted as a positive feedback towards one's partner during the course of a collaboration. As another example, a positive feedback can promote certain cognitive processes, or it can inhibit other cognitive processes according to the conditions in the environment [33]. In both cases, emotions play a regulatory role for cognitive processes [68]. Some of the effects flow from underlying shifts in the way people perceive and think under the influence of emotion.

2.1.1 Emotion in Social Context

In this section, I discuss the importance of studying emotions within a social context. This perspective is important in my research because my work is focused on collaboration as a particular social setting between individuals. Understanding the dynamics of collaboration requires one to understand influential underlying components. I have chosen to study emotion as a crucial underlying component in humans' social life which will be discussed in detail throughout this document.

Emotions are involved in developing social context. Humans are social and most of the causations and constitutions of their emotions are social. Brian Parkinson in [124] argues that many of the causes of emotions are interpersonal and communicative rather than internal and reactive phenomena. There are different social aspects of emotions influenced by various factors such as social context and social relationship type. For instance, a dominant-submissive social relationship can cause and contain different emotions with different intensities compared to a reciprocal or a friendship social relationship type. As another example, an emotion can be interpreted in a certain way when an individual is situated in an environment with other people who are expressing a particular emotion.

As mentioned earlier, the social context is an important factor influencing one's emotions. A dyadic interaction is one type of a setting in a social context. Dyadic interaction tasks allow us to study emotion in a social setting [34]. Dyadic interaction tasks make it possible to examine how individuals experience and express emotions during social interactions and how emotions shape and are shaped by the reciprocal interactions between individuals. In addition, eliciting and monitoring emotional processes yields useful information about the role emotion plays in interpersonal relationships. Compared with other emotion-eliciting events, events in a dyadic interaction can better help us study an ongoing emotional relationship between two individuals in addition to their internal emotional and cognitive processes. Dyadic interaction tasks are ideal for studying a range of emotional responses because of the fairly unstructured conversations between the individuals. Thus, dyadic interaction tasks will generate a wide range of emotions in comparison with the controlled emotion-eliciting events.

Now that we know the scope of the social setting in this research, I want to explain how emotions can be social. There are numerous ways that emotions can be social [168]. There is a consensus on the fact that social events and entities surrounding the individual play an essential role in the generation of emotion. There are several ways in which other people elicit emotional responses in us. One is that we feel the emotions of those around us. Also, we have emotions about actions of those people around us. Another is we have emotions about the things that happen to other people. Yet another is our concern about our relationship with others that elicits emotion in us. The groups to which we belong can also elicit our emotions. Moreover, we can feel emotion about the success and failure of our own group or of other groups. In addition, groups or individuals may make salient cultural concerns or societal expectations that can elicit our emotions.

Beside the fact that social context can cause eliciting emotions in individuals, social context provides information about what emotion should be expressed, by whom, and in what situations. For instance, people are well aware of the inap-

propriateness of expressing too much emotion to acquaintances [168]. However, the social knowledge of emotion expression is only partially delivered in an explicit fashion. There are studies on the regulatory role of society and social relationships on emotions, showing that people's emotions become socialized in implicit and unconscious ways. From this perspective, social context can control and direct our attention toward certain types of events and away from others.

Humans are emotional and social beings. Their emotions and the social context in which they are involved have mutual impacts on each other. But, what if humans can share their emotions with others just as they share their thoughts, resources and their environment. Sharing an emotion with others may alter the experience of an event. For instance, according to the nature of the relationship between the individuals, the expression of emotions can either restrain them from further interactions or improve their relationship. Furthermore, individuals sharing emotions might possess a shared understanding of their environment. Socially shared and regulated emotions also provide social meanings to the events happening in the environment [172]. For instance, people are likely to make social inferences based on the presence or absence of particular emotions in their social environment. Moreover, emotions can provide a basis for judgment depending on the individual's relationships with others. In other words, emotions can associate or disassociate an individual, therefore, they can change or maintain the individual's social relationships [168].

Emotions can also play the role of a motivator in a social context. There is a subset of social emotions delineated as role-taking emotions in [155]. Shott provides two categories of *reflexive* (e.g. shame or pride) and *empathic* (e.g., empathy or pity) role-taking emotions. The reflexive emotions can motivate the individual's self-control which depends on the anticipated reactions of others to the individual's behaviors. For instance, guilt might lead the individual to behave altruistically to restore a positive social stance for that individual. Empathic or vicarious emotions are based on an individual mentally placing himself in other's situation to understand how the other feels in that situation. These emotions motivate prosocial behaviors

to maintain an individual's internal well-being [166].

2.1.2 Communicating Emotions

I have discussed the importance of social context and its relation to human emotions. Humans need to communicate their emotions within the social context for different reasons. In [62] Goffman argues that human behaviors around others are performative which is often intended to convey information to others. When human's actions are visible in the social context, they behave differently in the presence of the others [175]. The social life of an individual is comprised of the individual's internal cognitive competencies and his interactions in the society. Lazarus says, if society is a fabric, then emotion is its color [95]. Although emotions undeniably have personal aspects, they are usually experienced in a social context and acquire their significance in relation to this context [107].

There are several events that can elicit emotions in social contexts. For instance, during the interaction the cause of an emotion can be verbal (an utterance during conversation), nonverbal (someone's gesture), personal thoughts (interpretation of an event), or even emotions themselves (e.g., happiness for a partner's sense of pride). An utterance can include content and relational meaning. The content carries the information about the topic or the subject of the interaction, and the relational meaning reveals the meaning between the speaker and the hearer. An emotion might seem to be elicited by the content of the utterance, but in fact it is an individual's response to the relational meaning [130].

The interpretation of these relational meanings are handled by the appraisal of the events. Appraisal processes (see Section 2.2.2) also give us a way of viewing emotion as social [169]. Meaning is created by an individual's social relationships and experiences in the social world. Individuals communicate these meanings through utterances. Utterances in emotionally charged conversations, by their very nature, are supposed to inform the others about something novel. Novelty is an essential component of an event for appraisal. Conversations also possess the concept of con-

sistency because the utterances with consistent meaning constitute the individual's underlying beliefs. Relevancy is another component of an event that can be assessed by appraisal. The degree to which the individual's personal and mutual beliefs are strong and related controls emotionally rich social contexts. In other words, the more divergent the individual's beliefs, the more effort is required to converge (to be understood) which leads to more emotional responses in individuals. From another point of view, human speech carries emotional information in the semantics and in the speech prosody. The semantics or the content of what an individual says includes obvious expression of emotion. However, the prosody holds more detailed emotional information by combining non-semantic cues in spoken language (e.g., rhythm and intonation) [104].

Interpretation of the events in the social context requires a baseline for the individual's assessment process. Goals as the pillar of collaborative interactions can provide this baseline for an individual. Goals are crucial in relational meanings of the events in a social context. The facilitation, interference and inhibition of goals are each correlated with certain type of emotions. In most conversations during collaboration goals can be categorized into three different groups: goals related to accomplishing a task, goals to reveal one's personal beliefs, and goals to regulate one's social relationships [130]. For instance, for task-related goals, utterances related to accomplished tasks reveal joyful relational meaning; utterances related to impeded tasks reveal disappointing relational meanings which can lead to anger, and utterances related to tasks with no or little progress reveal the frustration of the individuals. Lastly, all these emotional responses in a social context will not only regulate or maintain individual's actions to reveal or hinder an intention, but also can control the way that action should be taken.

A successful and effective emotional communication necessitates ongoing reciprocal adjustments between interactants that can happen by interpreting each other's behaviors [107]. It not only requires proper interpretation of the other's expressions, but also correct assessment of the extent to which others can read an individual's

expressions. In emotional communication, individuals are constantly exchanging messages about their mental states, and modifying each other's emotional responses as they occur. Individuals perceive other's emotional states through verbal and nonverbal responses during the interaction by processing relevant messages. Communication dynamics represent the temporal relationship between these communicative messages. The verbal and nonverbal messages from one participant are better interpreted inside the correct context including the history and the ongoing messages from the other individuals. Interpersonal dynamics (also known as micro-dynamics in sociology) represent this influence of relationships between individuals [117].

2.1.3 Social Functions of Emotions

Humans are able to communnicate their emotions in a social context. The social functions of emotions are the reason behind why humans try to communicate their emotions. In this section, I briefly discuss these social functions of emotions since they are directly related to my work. Ekman in [50] asserts that the primary function of emotions is to mobilize the organism to deal with important interpersonal encounters. Darwin in [40] argues the significance of social communicative functions of emotions. Emotions describe interpersonal dynamics in a way that they can constitute individuals' relationships [124, 168]. One aspect of expressing and communicating emotion in a social context is to express one's social motives and intentions [79]. Another aspect of communicating emotions is to reveal the underlying mental states of an individual [125]. In other words, emotions constitute two different functionalities of expressing communicative signals associated with one's social motives and intentions as well as expressing one's internal states and how one feels about something. In [90] Van Kleef has discussed the idea of inferential processes with which individuals can infer information about others' feelings, relational orientations and behavioral intentions based on their emotional expressions. He also argues that emotional expressions can impact social interactions by eliciting others' affective responses.

Functional accounts vary according to the kind of system being analyzed. Therefore, functional approaches to the emotions should vary by level of analysis. Social functions of emotions can be analyzed in *individual, dyadic, group* and *cultural* levels. My focus in this research is on social functions in dyadic interaction (more specifically collaboration); I also consider these functions at the individual's level especially when interpreting the other collaborator's behaviors. Studies in all these levels share a few assumptions about social accounts of emotions. They assume a) individuals are social by nature and pursue solutions to survival problems in social relationships, b) individuals apply their emotions to coordinate their social interactions and relationships to address these survival problems, c) emotions are processes mediating the individuals' relations to their dynamic environment [87]. In dyadic interactions, studies focus on how emotions impact the interactions of individuals in meaningful relationships. In [87] Keltner and Haidt discuss that in a dyadic setting, researchers mostly focus on communication of emotion (e.g. Scherer [144], DePaulo [45]), properties (e.g. emotion contingency, emotion synchrony) of dyadic emotions (e.g. Levenson & Gottman [98]), discourse (e.g. Bretherton [21]), and attachments (e.g. Hazan & Shaver [78]).

Examples of Social Emotions:

There are many different types of emotions, some of which are considered social, since they appear and provide meaning in social context. Here, I provide four examples of these emotions as well as their social functions to show how social functions of emotions impact individuals and the groups they belong to. And, what causes them to be expressed by an individual.

Guilt – The function of guilt is to positively direct our behavior toward our group. We feel guilt when we hurt someone in our group, or when we fail to reciprocate care or kindness. Guilt motivates us to not hurt people in our group and to give back to others who have given to us, and in this way we strengthen the survival

prospects of both the group and ourselves.

Shame – The function of shame is twofold. On the one hand, it keeps us within the rules and norms of society by informing us when we have done something dishonorable, disgraceful, or in some way condemned by our group. On the other hand, it informs the other members of our group that we know that we have dishonored ourselves. The main difference between guilt and shame is that guilt is focused on a behavior, whereas shame is focused on ourselves.

Embarrassment – Embarrassment is related to shame, but includes some important differences. Embarrassment can only happen in public, whereas shame can happen when we are alone. We can feel embarrassment about very minor issues that have no moral implications, such as body odor, whereas shame typically concerns more grave issues with moral implications.

Pride – The function of pride is to reinforce when we or another person has done or represented something the group finds excellent. In this way, group values are reinforced and incentivized, which again helps the group to function better and motivates us to do things the group values. There is a negative form of pride in which our internal appraisal of our worth is inflated compared to the opinions of others, which is more correctly called hubris.

2.2 Artificial Emotions

In this section, I discuss the concept of artificial emotions. I briefly review the influential cognitive theories which describe emotions. These theories provide cognitive structure of emotions and some of them describe the underlying evaluative processes of emotion eliciting mechanisms. This is important in my work because I am interested to investigate how emotions are involved in collaboration and how the dynamics of a collaboration structure impact the underlying processes of emotion.

Ortony, Clore and Collins in [123] view emotions as valenced reactions to events, agents, or objects, with their particular nature being determined by the way in

which the situation eliciting the emotions is constructed. Their goal in [123] is to construct a cognitive theory concerning the origins of emotions. In general, the analysis of the cognitive structure of emotions provides eliciting conditions for each emotion as well as information that we need to design artificial emotions. Affect serves as information about the value of the input to our cognitive processes. This information can impact the outcome of human decision-making procedures. In [32] Clore and Huntsinger claim that emotions are pervasive influences on human judgment and thought. For instance, positive affect tends to promote cognitive, relational processes, whereas negative affect tends to inhibit relational processing, resulting in more perceptual, stimulus-specific processing [33]. Understanding more details about emotions and how they impact cognitive processes has led to the unification of their computational models in research related to cognitive systems. For example, in [108] authors argue a functional symbiosis between the cognitive theory of behavior control and the theory of emotion (an appraisal theory).

Emotions, as an integral part of rational behavior, provide adaptive values for an artificial creature. They can control an agent's *attention* to focus on the most salient and relevant stimulus to solve the immediate problem. They can also help an agent to *monitor its own performance* so that the agent can make alterations on goals and plans. Emotions can act as a *memory filter* allowing a better recall of the events that are congruent with current cognitive and emotional states [19]. *Assisting the reasoning process* is another role of emotions; they assist the reasoning process by directing the cognitive information processes to the perceptual cues. Emotions impact the transformation of the agent's *decision-making behavior* [61] leading to particular type of actions in a certain type of environment [176]. Emotions can *govern behavior tendencies* by providing immediate emotional responses, e.g., avoidance of elaborate reasoning because of an unavailable time or an unconcerned situation. Furthermore, emotions *provide support for social interactions* by helping the agent to understand others' behaviors as well as making expressions of the agent's internal states more perceivable during the interaction [59].

The importance of these values of emotions for designing social agents having artificial emotions is prominent. However, the question is what problems are we facing in designing an effective social agent? In [42] authors discuss some of these problems and provide references speculating on the nature, function and mechanisms of emotions. Also, the importance of emotions and the incorporation of emotions in intelligent systems as well as implementation of emotions in several multi-agent systems are presented in [111]. Scheutz discusses the role of emotions in artificial intelligence and how we can determine the utility of emotions for the design of an artificial agent [149]. In [17] authors present a definition and theory of artificial emotions viewed as a sequential process comprising the appraisal of the agent global state; they also show how emotions are generated, represented and used in the Salt and Pepper architecture for autonomous agents. From the behavior perspective, appropriately timed and clearly expressed emotions are a central requirement for believable agents [12].

There are several architectures modelling emotions for the purpose of enhancing the believability and effectiveness of the agents and robots. But the question is how do we model emotions? Hudlika in [82] deconstructs the concept of emotion modelling into: (a) fundamental categories of processes for emotion generation and emotion effects, and (b) identification of some of the fundamental computational tasks required to implement these processes. These building blocks can be helpful as a guideline for the systematic development of new computational models, or for the assessment of existing computational models of emotions as discussed in [100] and [110]. There are also logical formalizations of emotions and emotional attitudes (including speech acts) and corresponding mental states to provide a systematic analysis of computational models of emotions [1, 64, 75].

From another perspective, the necessity of employing emotions in robotics and more specifically social robots has been argued in [129] and [162]. Social robotics and cognitive robotics have many overlapping concepts, especially when they focus on interaction between a robot and a human. The relationship between cognition

and emotion receives more attention due to the mutual influences they have on each other [118, 150] resulting in many studies. For instance, in [59] authors employ emotions in the learning procedure of a robot, and in [23] and [148] authors discuss the importance of emotions in the action selection procedure of an agent or a robot, impacting on the behavior arbitration and self-adaptation mechanisms. Ultimately, employing artificial emotions will impact the context of human-robot/computer interaction [81] and how humans and robots understand each other's emotions in a social environment [89, 120]. In [137] authors selected twelve autonomous agents that incorporate an emotion mechanism into the action selection procedure to compare. They introduced a framework based on correlations between emotion roles performed and aspects of emotion mechanisms used to perform those roles. Gratch and Marsella also present one method to evaluate a computational model of emotion in [66] which compares behavior of the model against human behavior.

2.2.1 Cognitive Architectures

There are several integrated cognitive architectures trying to produce all aspects of behavior as a single system while remaining constant across different domains [83, 92]. The comparison of underlying philosophy and functional description of the most prominent cognitive architectures have been surveyed; several criteria are provided to evaluate such architectures [29, 94, 167]. The necessity of integrating these cognitive architectures with robots has been discussed from the perspective of developmental psychology [6, 46, 86]. There are also many examples emphasizing the importance of cognitive robotics from this perspective. Some of these cognitive architectures are biologically inspired, e.g., *eBICA* [139], or [13] and [134], while some others are inspired by psychological theories, e.g., $ACT-R\Phi$ [39], or [116] and [47], in which some of them also incorporate the concept of affect in their design [25].

2.2.2 Appraisal Theory of Emotions

In this section, I describe the appraisal theory of emotions and provide some examples of computational models using this theory. The appraisal theory is important because it is one of two underlying theories in my work. This thesis borrows an approach about appraisal theory from another work, which is also introduced in this section.

The emotional experience is the experience of a particular situation [57]. Appraisal theory describes the cognitive process by which an individual evaluates the situation in the environment with respect to the individual's well-being and triggers emotions to control internal changes and external actions. According to this theory, appraisals are separable antecedents of emotion, that is, the individual first evaluates the environment and then feels an appropriate emotion [147]. The appraisal procedure begins with the evaluation process of the environment according to the internalized goals and is based on systematic assessment of several elements [145]. The outcome of this process triggers the appropriate emotions. At the end, the Coping mechanism manages the individual's action with respect to the individual's emotional state and the existing internal and/or external demands [55]. An individual can also use knowledge about the emotional reactions of others to make inferences about them. According to the appraisal patterns, different emotions can be experienced and expressed. Since expression of emotions reflects one's intentions through the appraisal process, the *reverse appraisal* mechanism helps one to infer other's mental states based on their expressions. [43, 76].

There are several examples in artificial intelligence and robotics of applying the appraisal theory as the basis of a computational model for emotions [2, 88, 109]. In [140] authors describe a system approach to appraisal processes based on Scherer's work on appraisal and the Component Process Model (CPM) [143]. They show how the temporal unfolding of emotions can be experimentally tested. In this thesis, I use the cognitive appraisal theory of emotion provided by Gratch and Marsella in

[65]. They lay out a general domain-independent computational model of appraisal and coping. I use this appraisal approach, in general, as an evaluation mechanism for the internal and external events to assist the cognition and collaboration processes in my theory.

2.2.3 Social Appraisals

Social skills have been mostly neglected in artificial intelligence and robotics. However, there is a broad discussion in natural and social sciences, e.g. psychology and primatology [15, 18, 85, 171, 173], about the role of social factors in the development of intelligence [41] (see Sections 2.1.1 to 2.1.3). Robots in the real world, e.g. domestic robots or collaborative robots, require extensive understanding of aspects of humans' behaviors within their environment as well as the ability to communicate and collaborate with them. Emotions, as coordinated responses to detected or inferred relational meanings of the environment (based on appraisal theory), can provide understanding of the social environment, and the capability of communicating internal mental states and maintaining collaborations with human partners. In fact, the emotion processes momentarily respond to the unfolding affordances and constraints offered by the dynamic context of a social interaction [126]. Appraisal can provide the assessment of goal relevance and goal congruence with focus on self or other, the event, or the object in a social context [127]. In short, the agent will be capable of appraising the social environment in order to maintain effective social interaction.

2.2.4 Affect and Decision-Making

Decision-Making is an important and complicated process for any robot or virtual agent. This process becomes more complicated when the agent needs to make a decision considering its own private goal, the collaboration's shared goal and the human collaborator's interests. I will provide more details about the following concepts in

Chapter 3.

There are examples of rational and social agents designed based on the decision theory and emotional states [61]. Agents must take a form of action after making a decision. Zhu and Thagard argue how emotions significantly affect the action generation procedure as well as action execution and control [176]. The decision-making procedure, as the basis of an agent's behaviors and actions, is a crucial process for an agent in a social environment. Decision-making is a process that unfolds over time and should be explored in more detail. According to [128], the temporal structure of the decision-making process contains three component processes:

1. Choosing among options initially involves the process of *assessing* the available options. One's affective state and appraisal evaluation of one's internal state as well as the surrounding environment helps in the assessment of all available options. For example, based on the scenario in section 3.1.2, Robot's emotion instance is fear because of an existing block in the plan and its evaluation of Astronaut's emotion as anger (for the same reason). The assessment of available options will be based on minimizing the distance to the shared goal and Astronaut's satisfaction. For instance, if Robot faces a non-critical task, it will give higher value to Astronaut's demanding task which will cause the postponement of its own.

2. This process is followed by the *selection* of an option based on the value that has been assigned to the option. This process is also augmented by affective evaluation of the world, including self, other(s) and the environment. For instance, in our scenario (see section 3.1.2), following the assessment of available options, Robot will focus on Astronaut's preferred task. Also, Robot creates and annotates meta information of the current state of the collaboration with affective evaluations.

3. Finally, the outcome associated with the selected action is *evaluated* and *incorporated* into existing knowledge for subsequent decisions which implicitly

and explicitly help the belief and appraisal emotion systems to operate coherently over time. For instance, if something goes wrong the Robot gives a negative affective attribution to the outcome of the selected action or even a certain path to that action to be used in future assessments.

People's experience of events leak into their beliefs and ultimately decisions. One aspect of these type of experiences is conscious or unconscious annotations by different emotions. For instance, one will never forget working with a friend due to the pleasant feeling of experiencing the outcome. On the other hand, a person will always remember a particular experience in life because of an utterly negative emotion that was felt at the time [128].

Emotions appear to influence the value and weight computation of available alternatives, and these computations are dynamically adjusted based on the environment and the individual's internal states [28]. This way agents can operate and take actions based on preferences. In other words, emotional states of individuals are linked to their decision-making processes, assuming that emotions affect the way gains or losses are transformed to weights and values of the alternative beliefs, actions, tasks, and, in general, plans [128]. The outcome of an action is also profoundly bound to the decision making process as a final and an important stage. The experience of an outcome and in particular, the differences between the expected and observed outcome provides an opportunity to improve one's beliefs about consequences (value) of the available alternatives and adopt a better decision policy in the future [28, 128].

2.3 Affect and Motives

Motives are essential mental components in decision-making procedures and applying them in an affect-driven collaborative agent is part of this thesis' contribution. In this section, I provide related works on computational models of motivation and discuss the nature of motives. I also explain three of the important social motives

which will be used in my work. Finally, I discuss that humans' beliefs, emotions and motives are related and influence each other.

Motives' principles and mechanisms, as the reasons behind one's intentions and actions, and the influences of motives on cognition have been discussed in philosophy, neuroscience, psychology and artificial intelligence [10, 16, 22, 159, 160]. There are several examples in AI providing computational models for different psychological theories of motivation. Bach's MicroPsi agent architecture describes the interaction of emotion, motivation and cognition of agents based on Dietrich Dörner's Psi theory [8, 9, 10, 11]. Merrick and Shafi provide a computational model for motivation based on Henry Murray's theory [119] describing the three important social motivations of *achievement, affiliation* and *power*. They focus on the role of motivation in a goal-selection mechanism [114]. There are other examples focusing on the impact of motives on different cognitive processes in robots and artificial agents [20, 27, 44, 153, 170, 174]. The motivation mechanism in my work is inspired by Murray's theory and Bach's approach on Dörner's theory. It is focused on the role of motives in cognitive processes, e.g., intention formation in coping, during collaboration, which will be discussed in chapters 3 and 4.

2.3.1 Motives

A motive consists of an urge (that is, a need indicating a demand) and a goal that is related to this urge [9]. Motives shape cognition and behavior [152]. To be motivated means to be moved to do something [138]. Motives direct behaviors towards particular goals, which makes the agent more persistent in actions it takes. They also affect cognitive processes by increasing level of attention. Motive, as the outcome of the motivation process, initiates, directs and maintains goal-oriented behaviors.

Motives are goal-driven and they move the agent towards the attainment of corresponding sets of intentions. In other words, motives as an essential part of affect can lead the agent to empower an intention. They are essentially mechanisms that in

light of beliefs tend to produce, modify or select between actions and their reciprocal intentions. Some of the motives are transient, like helping the Astronaut to hold the panel, while some are long term, like reaching to the shared goal during collaboration which in our example is installing solar panels and satisfying the Astronaut's needs in the field (see Section 3.1.2).

2.3.2 Motivation Theory

There are several motivation theories in psychology [14, 63, 93], some of which have received little attention as the basis for computational models. In [119], Murray described and studied 20 different human motives, of which three have received attention in psychology and artificial intelligence as social motives [114, 177]. The following is a brief description of these three social motives, *achievement, affiliation* and *power* [7, 177] which will be used in my work:

- **Achievement motivation:** Achievement motivation drives humans to strive for excellence by improving on personal and societal standards of performance. It involves a concern for excellence, for doing one's best. In artificial agents, achievement motivation has potential roles in focusing agent behavior and driving the acquisition of competence.

- **Affiliation motivation:** Affiliation refers to a class of social interactions that seek contact with formerly unknown or little known individuals and maintain contact with those individuals in a manner that both parties experience as satisfying, stimulating, and enriching. It involves a concern with developing friendly connections with others through the two contrasting emotional components of hope of affiliation and fear of rejection. These two components become more crucial in the collaboration domain due to the importance of social emotions and their impact on beliefs and intentions.

- **Power motivation:** Power can be described as a domain specific relationship between two individuals, characterized by the asymmetric distribution of

social competence, access to resources, or social status. It involves concern with having an impact on other people or on the world at large. There are different aspects of fear or avoidance of power which channel and moderate the expression of power into socially acceptable behavior, working as inhibitions to unseemly tendencies. Power motivation can be considered with respect to the probability of success which makes it relevant to the cognitive appraisal of emotions during collaboration.

In [177] it is shown that success of a power goal is associated with anger, confusion and disgust; success at an affiliation goal is associated with interest, happiness and feeling loved; and success at an achievement goal is associated with interest, surprise, happiness, excitement and a sense of focus. In other words, succeeding at a particular motive is associated with experiencing particular emotions.

2.3.3 Motivation Regulation of Emotions

Cognitive processes help humans to reach multiple conclusions before taking actions; motives play a role in determining which of these conclusions will be used on a given occasion. They affect reasoning through reliance on a biased set of cognitive processes including strategies for accessing, constructing and evaluating beliefs. Motivated reasoning is a form of implicit emotion regulation in which the brain converges on decisions that minimize negative and maximize positive affective states associated with threat to or attainment of motives.

Affect-driven belief systems profoundly impact the transformation of actions into choices as well as the modification of future expectations. Most theorists agree that affect is closely linked to motives and motivated behavior. Many emotion theorists believe that nearly all emotions are relevant to each motive [113]. Affective states essentially are feedback signals indicating the progress of motivated, goal-directed behaviors. Moreover, affective states are a powerful source of motivated cognition and behavior. They influence how an agent can perceive, interpret, respond to,

and communicate in social settings. However, motives and affective states can only become effective by recruiting cognitive processes to their service. Zurbriggen and Sturman studied the mutual link between motives and emotions hypothesized by McClelland [177]. Wright and Sloman discuss an implementation of an autonomous agent, MINDER1, as a partial realization of an architecture for motive processing [174]. Sellers proposes an emotion model investigating the role of emotions in cognition [153]. Cañamero and Velde have addressed the problem of how emotions ground perception and social interaction in [120]. In [84] researchers argue that the five primary emotions are involved in all goal-directed behavior in the following way:

- **Happiness:** is associated with the perception of improving progress towards a goal.
- **Anger:** is associated with a blocked plan.
- **Fear:** is associated with the occurrence of a goal conflict or a threat to self-preservation.
- **Disgust:** is associated with a perception of something to reject.
- **Desire:** is associated with a perception of something to approach.

I use this idea of regulating goal-directed behaviors by emotions in my work. For instance, what is meant by "Astronaut is angry with Robot" (see Section 3.1.2)? This statement implies that Astronaut believes that there is something Robot did or failed to do and as a result at least one of Astronaut's motives has been violated. Astronaut's belief does not suffice for anger, since Astronaut might merely regret what happened or be disappointed in Robot, without being angry. Anger also requires a new motive in Astronaut. This new retributive motive is not necessarily selected for action however intense it may be. For instance, fear of consequences may keep it inoperative [160].

2.4 Collaboration Theory

Collaboration is a special type of coordinated activity in which the participants work jointly, together performing a task or carrying out the activities needed to satisfy a shared goal [72]. Collaboration is one of the fundamental pillars of my studies and it defines the context of my research. The discourse theory, proposed by Grosz and Sidner also plays a crucial role in my research since the agent/robot needs to communicate with a human through utterances during the collaboration [73]. This theory provides the discourse structure which will be used in my model.

There are prominent collaboration theories that are mostly based on plans and often analysis of the discourse between collaborators revolving around these plans [74, 101]. In these theories the discourse analysis is based on search over these tree plans [136]. In [74], Grosz and Sidner argue that the components of the discourse structure are a trichotomy of linguistic structure, intentions structure and the attention state. In their work, the linguistic structure of a discourse is a sequence of utterances aggregating into discourse segments just as the words in a single sentence form constituent phrases. They also discuss the idea of the discourse purpose as the intention that underlies engagement in the particular discourse. They believe this intention is the reason behind performing a discourse rather than some other actions, and also the reason behind conveying a particular content of the discourse rather than some other contents. They describe mechanisms for plan analysis looking at Discourse Segment Purposes (DSPs). In fact, the DSPs specify how the discourse segments contribute to achieving the overall discourse purpose. Finally, the third component in their theory, the attentional state, provides an abstraction of the agent's focus of attention as the discourse unfolds. The focusing structure contains DSPs and the stacking of focus spaces reflects the relative salience of the entities in each space during the discourse. In short, the focusing structure is the central repository for the contextual content required for processing utterances during the discourse [74].

Shared plan is another essential concept in the collaboration context. The definition of the shared plan is derived from the definition of plans Pollack introduced in [131, 132] since it rests on a detailed treatment of the relations among actions and it distinguishes the intentions and beliefs of an agent about those actions. However, since Pollack's plan model is just a simple plan of a single agent, Grosz and Sidner extended that to plans of two or more collaborative agents. The concept of the shared plan provides a framework in which to further evaluate and explore the roles that particular beliefs and intentions play in collaborative activity [103]. However, this formulation of shared plans (a) could only deal with activities that directly decomposed into single-agent actions, (b) did not address the requirement for the commitment of the agents to their joint activities, and (c) did not adequately deal with agents having partial recipes [72]. Grosz and Kraus in [72], reformulate Pollack's definition of the individual plans [132], and also revise and expand the shared plan to address these shortcomings.

There are also some other theories with similarities and contrasts conveying collaboration concepts including Cohen and Levesque's work describing the concept of *joint intentions* in [35, 99], and Tambe's work on *STEAM teamwork model* [165]. The concept of collaboration has also received attention in the industry and in research in robotic laboratories [67]. There are many research focusing on different aspects of collaboration each of which are different than my own work. In my thesis, I focus on emotion functions and how they impact collaboration's structure and processes, and how the dynamics of the collaboration structure influences emotion-regulated processes. Some of the other works focus on the concepts of robot assistants [31], or teamwork and its challenges in cognitive and behavioral levels [35, 121, 142, 165]. Some researchers have an overall look at a collaboration concept at the architectural level. In [60] authors present a collaborative architecture, COCHI, to support the concept of emotional awareness. In [53] authors present the integration of emotional competence into a cognitive architecture which runs on a robot, MEXI. In [163] authors discuss the challenges of integrating natural language,

gesture understanding and spatial reasoning of a collaborative humanoid robot situated in the space. The importance of communication during collaboration has been considered by some researchers from human-computer interaction and human-robot collaboration [30, 112, 136] to theories describing collaborative negotiation, and discourse planning and structures [3, 73, 158]. There are other concepts such as joint actions and commitments [71], dynamics of intentions during collaboration [99], and task-based planning providing more depth in the context of collaboration [24, 135].

2.5 Theory of Mind

Theory of mind, as a crucial component in human's social interaction, plays an important role in my computational model. It discusses one's beliefs about others as intentional agents. Beside the immediate effect, an individual's action also depends on the beliefs about other's perception of that action as well as the reaction they take. In my work, we use this concept whenever the agent needs to anticipate the human's mental states. I will also use the concept of *user model* as a standard collection of properties to describe others.

The concept of theory of mind has received attention in social psychology and artificial intelligence. Eligio et al. explore what collaborators understand about each other's emotions and conclude being aware of each other's emotions helps collaborators to improve their performance [51]. Fussell and Kraus discuss the importance of perspective taking in a successful communication in a social setting [58]. Scassellati discusses the importance of attribution of beliefs, goals and desires to others. He presents two psychological theories on the development of theory of mind in humans and their potential application in building robots with similar capabilities [141]. Hiatt and Trafton present a cognitive model which borrows mechanisms from three different postulates of theory of mind and show that their model produces behaviors in accordance with various theories of experiences [80]. Si, Marsella and Pynadath discuss PsychSim, an implemented multi-agent-based simulation tool for

modeling social interaction, which has its own beliefs about its environment and a recursive model of other agents [133]. They also investigate the computational modeling of appraisal in a multi-agent decision-theoretic framework using POMDP based agents [156, 157]. Since applying the concept of theory of mind is crucial in social interaction and collaboration, I will employ a simplified mechanism inspired by the existing works for our agent.

CHAPTER 3
AFFECTIVE MOTIVATIONAL COLLABORATION THEORY

Affective Motivational Collaboration Theory is about the interpretation and prediction of the observable behaviors in a dyadic collaborative interaction. The 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 self's relationship to the collaborative environment. These reactive and deliberative processes are triggered by two types of events: *external* events, such as the other's *utterances* and *primitive actions*, and *internal* events, comprising changes in the self's mental states, such as belief formation and emotional changes (see Section 3.3). *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 self's mental states including the anticipated mental states of the other, generated according to the self's model of the other (see Section 2.5). These mental states include beliefs, intentions, goals, motives and emotion instances. Each of these mental states possesses multiple attributes impacting the relation between cognition and behavior or perception. The nature of these attributes will be discussed in Section 3.6.

There are several theories discussed in Section 2.4, which describe the underly-

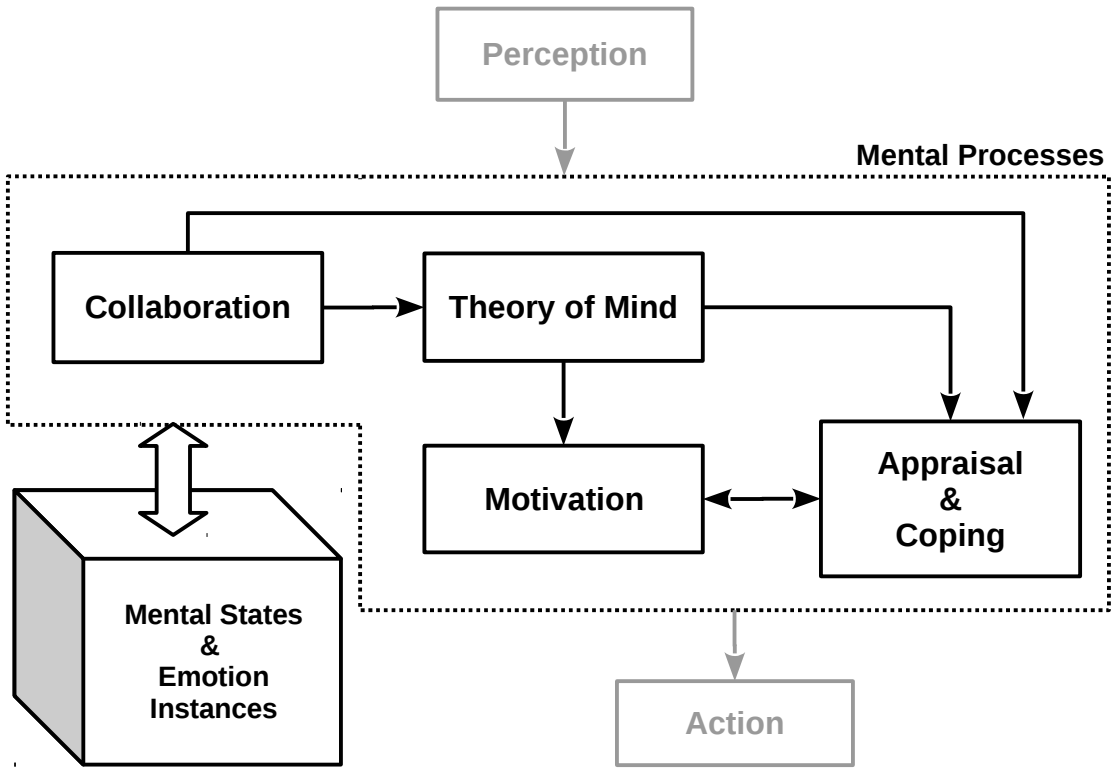

Figure 3.1: Roadmap of *Affective Motivational Collaboration Theory* showing primary influences between processes.

ing structure of a collaboration based on mental states of the collaborators. The collaboration structure of *Affective Motivational Collaboration Theory* is based on the SharedPlans theory [70]. *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 agents/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 *Appraisal & Coping* mechanism (see Figure 3.1) consists of the two processes of Appraisal and Coping. The Appraisal mechanism is responsible for evaluating changes in the self's Mental States, the anticipated Mental States of the other, and the state of the collaboration environment. Consequently, the Appraisal mechanism (see Figure 3.1) is connected to a) the Theory of Mind mechanism, to serve as an

evaluator whenever the self applies the Appraisal mechanism in reverse appraisal (see Section 2.2.2), b) the Collaboration mechanism, to interpret the progress and changes in the collaboration plan and associated Mental States, c) the Motivation mechanism, to generate and assess the self's new goal-driven motives whenever a new motive or intention is required, e.g., following the failure of a task, and d) the Perception mechanism, to interpret the external events from the collaboration environment. The Coping mechanism provides the self with different coping strategies associated with changes in the self's mental states with respect to the state of the collaboration. In other words, the Coping mechanism produces cognitive responses based on the appraisal patterns.

The *Motivation* mechanism (Figure 3.1) operates whenever the self a) requires a new motive to overcome an internal impasse in an ongoing task, or b) wants to provide an external motive to the other when the other faces a problem in a task. In both cases, the Motivation mechanism uses the Appraisal mechanism to compute attributes (see Section 3.6) of the competing motives. Also, the Motivation mechanism can serve the Theory of Mind mechanism by helping the self to infer the motive behind the other's current action. Moreover, if there is an impasse in accomplishing a collaborative task, the self requires a new intention to take a new action no matter whether the self or the other is responsible for the task. In this case, the Motivation mechanism applies the beliefs associated with the blocked task as well as the Appraisal mechanism to generate and compare a new set of motives related to the status of the collaboration. Only one of these competing motives is most likely to become a new intention. The Motivation mechanism forms a new belief and ultimately a new intention based on the winning motive. As a result, the self can take an action based on the new intention to sustain the collaboration progress.

The *Theory of Mind* mechanism (Figure 3.1) is the mechanism of inferring a model of the other's anticipated Mental States. The self will progressively update this model during the collaboration. The refinement of this model helps the self to

anticipate the other's mental state more accurately, which ultimately impacts the quality of the collaboration and the achievement of the shared goal. Furthermore, the self can make inferences about the motive (or intention) behind the other's actions using the Motivation mechanism. This inference helps the self to update its own beliefs about the other's mental state. In the reverse appraisal process (see Section 2.2.2), the self also applies the Appraisal mechanism together with updated beliefs about the other's Mental States to make inferences about the other's current mental state based on the other'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 self's model of the other will update the self's mental state.

The Collaboration mechanism maintains constraints on actions. These constraints include constraints on task states and on the ordering of tasks. The Collaboration mechanism also provides processes to update and monitor the shared plan. These processes depend on the Appraisal mechanism to evaluate the current Mental States with respect to the current status of the collaboration. The self also shifts its focus of attention according to the outcome of the Appraisal mechanism. Moreover, the Collaboration mechanism can help the self to identify the failure of a task. The Appraisal and Motivation mechanisms provide interpretation of task failure and the formation of new Mental States (e.g. intentions) respectively. Ultimately, the Coping mechanism allows the self to perform behavior appropriate to the current state of the collaboration.

3.1 Scenario

I now provide the following scenario in a simple robotic domain, which I will later use as a source of a concrete example to help explain my theory and computational model. In the scenario, there is a mission specialist astronaut, who has had a

high success rate in accomplishing space missions. He is capable of operating the manipulator system and supporting equipment. He works as a commander in the field during the operation. He is focused, organized, mentally stable, curious and patient. He is trained to collaborate with special-purpose field operation robots.

There is also a robot which is assigned to the mission to provide services to the astronaut. It has been tested in extreme environmental conditions and has a very low failure rate. It is capable of wireless communication with the astronaut and a control base and is able to recognize and understand the astronaut's voice and nonverbal behavior. It has the ability to identify and assess its own emotions and those of the astronaut.

3.1.1 The Backstory

The scenario transpires in a NASA's research center. Light, temperature and other environmental factors are simulated based on conditions on the surface of the moon. The mission is to finish installing the required solar panels to provide energy for the operation of NASA's science lab on the moon. Ninety percent of these panels have already been installed. However, the operation is now faced with low batteries which forces everyone to be cautious about consuming energy. The astronaut is inspecting the working conditions in the field and planning the installation of the remaining panels in collaboration with the robot. He determines that the sun will cast shadows over the installation structure, leading to potential difficulties. The astronaut asks control base to go through the final checks of the robot and prepare it for the operation.

3.1.2 Example of a Collaborative 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 occurring because of a blocked or a protracted sub-task. The robot and the astronaut will go through a series of assessment processes to figure out a) how did the current blocking 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 its communication skills to overcome these problems and to motivate the astronaut to propose alternative tasks. The following is part of an interaction between the astronaut and the robot during their collaboration on installing solar panels.

1. ***Astronaut***: Please hold the panel on this structure.

   [*Robot holds the panel and Astronaut begins to work on the panel.*]

   [*Both the Robot and the Astronaut continue their collaboration to achieve their shared goal.*]

2. ***Astronaut***: At this point you should be careful how you hold the panel. Turn the right side 45 degrees towards me.

3. ***Robot***: Is this what you want?

4. ***Astronaut***: Yes, please do not move it.

   [*Astronaut finishes fixing the panel onto the structure and checks the connectors to make sure there are working.*]

5. ***Astronaut***: The connectors on this panel have problems and we might not be able to finish this task.

6. ***Robot***: Don't worry! I can replace the connectors in 4 minutes. We definitely can finish this task after that.

7. ***Astronaut***: Okay, go ahead and fix the connectors.

   [*Robot fixes the issue with the connectors and passes them to the Astronaut. Astronaut connects the wires to the connectors.*]

8. ***Astronaut***: I need you to begin welding this panel and also prepare the measurement tool for me.

9. ***Robot***: Do you want me to prepare the measurement tool first? Then, I can begin welding afterwards.

10. ***Astronaut***: Yes, that's fine!

    [*Astronaut waits for the Robot to weld the panel and prepare the measurement tool for him. Robot finishes the welding task after a long time, then prepares and passes the measurement tool to the Astronaut. But, the measurement tool has an accuracy problem.*]

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

12. ***Robot***: I see. But, I can help you with the measurement and we can finish the task as originally planned.

13. ***Astronaut***: That would be great!

    [*Robot helps the Astronaut to finish the measurement task.*]

    [*Then, the Robot goes back to its own internal goal, which is to fetch the second panel to finish the overall task.*]

3.2 General Argument

Affective Motivational Collaboration Theory focuses on emotion-regulated processes involved in collaboration and builds on two well-established theories in this context. The first is Grosz and Sidner's SharedPlans collaboration theory, which is based on the concepts of mutual belief and shared plans [70, 74]. Secondly, I build on the computational model of the appraisal theory of emotions by Marsella and Gratch [65, 66, 109, 110] which explains how emotions arise from an individual's interpretation

of its relationship with the environment, and specifies the dimensions of appraisal and the appraisal patterns characteristic of different emotions [147].

Existing collaboration theories (including SharedPlans) consider the nature of a collaboration to be more than a set of individual acts. These theories argue for an essential distinction between a collaboration and a simple interaction or even a coordination in terms of commitments [70, 102]. I believe there is also a need for a computational theory to specify and characterize the underlying cognitive processes of collaborative activities. The study of these cognitive processes helps explain why and how humans collaborate with each other. For instance, SharedPlans theory can describe our scenario in Section 3.1.2 in terms of fundamental Mental States, such as mutual beliefs, intentions, and shared plans. However, it does not describe the underlying processes leading to these Mental States. *Affective Motivational Collaboration Theory* extends the SharedPlans theory by describing these processes. Furthermore, emotions, due to their evaluative and regulatory nature, provide fundamental functions (see Section 3.4) each of which plays an essential role in producing social characteristics as conceptual constructs of a collaboration (see Section 3.5). In other words, these functions explain the dynamics of a collaboration structure.

Affective Motivational Collaboration Theory specifies the processes involved in the progress of a collaboration and how they impact the collaboration's underlying structure. For instance, it is important to understand what are the social characteristics involved in a collaboration? The utterances by the Astronaut and the Robot in the example scenario in Section 3.1.2 reveal manifestation of social characteristics in their collaborative behaviors. For example in the exchange below, the Robot in order to respond appropriately to the Astronaut's requests, needs to maintain its relationship with the new situation in the environment. Different emotion functions such as adaptation, goal management and attentional focus provide self-synchronization (see Section 3.5) characteristic of the Robot's collaborative behavior:

8. ***Astronaut***: I need you to begin welding this panel and also prepare the measurement tool for me.

9. ***Robot***: Do you want me to prepare the measurement tool first? Then, I can begin welding afterwards.

What is the nature of the processes involved in a collaboration? For example in the exchange below, the Robot changes its focus of attention to something important to the Astronaut because of its perception of the Astronaut's negative emotion:

5. ***Astronaut***: The connectors on this panel have problems and we might not be able to finish this task.

6. ***Robot***: Don't worry! I can replace the connectors in 4 minutes. We definitely can finish this task after that.

And, how do these processes impact the social characteristics of a collaboration? For instance in the exchange below, emotions and the Appraisal mechanism can influence self or social awareness during collaboration:

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

12. ***Robot***: I see. But, I can help you with the measurement and we can finish the task as originally planned.

Finally, *Affective Motivational Collaboration Theory* incorporates motivation as an emotion-regulated and goal-driven mechanism, by which the self can infer a new intention based on its beliefs about it and the other, as well as the result of an Appraisal mechanism. In general, a new motive can become a new intention and the self can take a new action based on the new intention. Hence, the Motivation mechanism can help the self to reduce the probability of an impasse with respect to

the shared plan during the collaboration. The Motivation mechanism also connects the outcome of the Appraisal mechanism and the Collaboration mechanism by applying the self's belief structure and appraisal patterns. The result of this process generates a set of competing motives, each of which has a chance to become the self's intention. The self also stores beliefs about the other's motives along with their corresponding attributes. Beliefs about motives can be about self or the other, and they can impact the Appraisal mechanism. For instance, when a problem occurs during a collaboration, the self can form a new intention and provide a solution. In the following example extracted from the scenario, the Astronaut informs the Robot of a new problem, and the Robot forms a new intention to solve the problem:

5. ***Astronaut***: The connectors on this panel have problems and we might not be able to finish this task.

6. ***Robot***: Don't worry! I can replace the connectors in 4 minutes. We definitely can finish this task after that.

As another example, the Astronaut below faces a problem in his own task and informs the Robot of his decision. The Robot forms a new intention to help the Astronaut to overcome his problem and ultimately, make progress in their collaboration:

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

12. ***Robot***: I see. But, I can help you with the measurement and we can finish the task as originally planned.

3.3 External and Internal Events

The *external events* occurring in this collaborative environment include a) *utterances* spoken by the collaborators, b) *primitive actions* executed, deferred, or aborted, and c) observable *emotion instances*. The *internal events* constitute changes in the self's mental states including a) changes in the *focus of attention*, b) *belief formation* and *updates*, c) *intention formation*, d) *goal alteration*, e) *emotional changes*, and f) *motive generation*. The above events are the events that my affective collaborative agent perceives. I will discuss below the operation of individual processes in my theory based on these events. Each of the following six sub-sections describe how an individual mechanisms in Figure 3.1 handle external and the internal events.

3.3.1 Collaboration Mechanism and Events

The Collaboration mechanism is responsible for maintaining the internal structure of a collaboration, including the focus of attention, constraints on actions, updating the shared plan and, in general, monitoring of the collaboration. All of these structures require updating each time the self receives an external event. For instance, an utterance by the other can impact the self's focus of attention during the collaboration, or the effect of a primitive action can influence the self's view of an impasse on a task. As another example, the perception of the other's emotion instance can cause significant changes in the self's collaboration monitoring. The Collaboration mechanism also responds to internal events. These changes can occur under the influence of several other processes which are able to alter the self's mental states. For instance, the Motivation mechanism can generate a new motive which causes the self to add a new intention to its own Mental States.

3.3.2 Appraisal of External Events

The other's *utterances*, the effect(s) of the collaborators' *primitive actions*, and the other's *emotion instances* (expressed nonverbally) are the three types of external

events perceived by the self during collaboration. The Appraisal mechanism receives the output of the Perception and Collaboration mechanisms as well as the requisite Mental States related to current event. It appraises the external event, in terms of appraisal variables using, a) the collaboration structure, b) the social characteristics of collaboration, and c) the history of the self's related Mental States. The collaboration structure contains information about the collaboration's shared plan and the collaborators' shared goal, the temporal and the hierarchical constraints of the tasks, and the current focus of attention.

The social characteristics of collaboration (see Section 3.5) are conceptual properties associated with collaboration. These concepts provide deeper understanding of the collaborative interactions in terms of the underlying processes from both the personal and the social perspectives. Furthermore, these concepts determine when each process is used during a collaboration and why. Lastly, the self will progressively generate and update various types of Mental States (discussed in Section 3.6 and also Chapter 4) during collaboration. The occurrence of a new external event causes a change in the self's Mental States. The construct of the new mental state, e.g., beliefs, are semantically connected to the older ones. The Appraisal mechanism uses the history of the Mental States to consistently evaluate a new external event.

3.3.3 Appraisal of Internal Events

The formation of a *new belief* (one that has not already been evaluated by the Appraisal mechanism, such as a belief about a new motive), and the formation of a *new motive*, are the two types of internal events that will be evaluated by the Appraisal mechanism. Thus, the Appraisal mechanism is involved in the formation of a new belief to ultimately create balanced behaviors during the course of a collaboration. The Appraisal mechanism receives the output of the Theory of Mind and Motivation mechanisms as well as the available Mental States (e.g., beliefs about the current emotion instances of the self and the other) and evaluates the environment in terms of the appraisal variables whenever a new belief is generated. As mentioned

in Section 3.3.2, a) the collaboration structure, b) the social characteristic of the collaboration, and c) the history of the self's Mental States, are also crucial in the evaluation of internal events. The outcome of this appraisal process can result in a new belief about emotion instances, e.g. believing to have hope, or it can generate beliefs about different attributes of the Mental States of the self or the other.

3.3.4 Coping with Events

External events do not directly cause the self's Coping mechanism to operate. Instead, it is the formation or update of Mental States that cause the Coping mechanism to choose an appropriate cognitive response to these events. The cognitive responses (also known as "coping strategies") are considered to act upon the self's relationship to the world and its own Mental States. External events also trigger other processes, which impact the self's Mental States. The changes in Mental States cause the Coping mechanism to provide a consistent and appropriate cognitive responses to the world. For instance, suppose the self perceives an utterance and evaluates it in terms of the appraisal variables. The values of these variables and the corresponding emotion instances will cause new beliefs to be formed, which then cause the Coping mechanism to appropriately choose the self's behavior.

3.3.5 Motivation and Events

The Motivation mechanism acts to regulate the self's Mental States and goal-directed behaviors for internal and social purposes. The Appraisal mechanism evaluates the state of self, the environment, or the anticipated mental state of the other. In each of these cases, the outcome of the Appraisal mechanism might indicate the need for internal or behavioral regulation. In such cases, the Motivation mechanism uses the Mental States associated with the state of self, the environment or the other's anticipated Mental States as well as the pattern provided by the Appraisal process to generate motives aligned with private or shared goals. Thus, via the Ap-

praisal mechanism the Motivation mechanism implicitly responds to both external and internal events. The attributes of the generated motives (see Section 3.6.4) will be updated every time a new event occurs. For instance, the Appraisal mechanism may evaluate the outcome of the current task as *unexpected, undesirable, uncontrollable* and *urgent* which is indicative of the failure of a task. Then, the *alarm* function of the Appraisal mechanism (see Section 3.4.5) will activate the Motivation mechanism to provide goal-directed motives, each of which has the potential to become a new intention.

3.3.6 Theory of Mind and Events

Theory of Mind operates in two different situations: first, when an external event occurs and the self wants to infer and interpret the other's mental state; and second, when an internal event occurs. Thus, Theory of Mind helps the self to choose the behavior best matched to the other's anticipated Mental States. The Theory of Mind mechanism infers the mental state of the other, which helps the self to update the user model of the other (see Section 2.5). The Motivation and the Appraisal mechanisms are also involved in this procedure. For instance, suppose the self infers the other's mental state through a reverse appraisal procedure. The Motivation mechanism includes another inverse procedure to infer the other's active motives, which can lead to inferring the other's goal, beliefs, motives and intentions.

3.4 Functions of Emotions

I have talked about the crucial role of emotions in communicating Mental States, motivating actions, and evaluating and interpreting internal states and the environment. Emotions generally speaking provide a set of intra- and interpersonal functions which regulate internal processes and the self's relationship to the other during the collaboration. An emotion instance, e.g., *anger*, occurring during collaboration, can lead to these emotion functions, e.g., *Alarm Mechanism* and *Action*

Selection, via interpreting the social meaning of that emotion (see Figure 3.2). Emotions have meanings in a social context (see Section 2.1.1) which can be interpreted by an observer. The self uses these emotion meanings to trigger appropriate emotion functions with respect to the current social context. Ultimately, the elicited emotion functions impact the self's Mental States and consequently behaviors. The changes in the self's behaviors manifest as what I call the social characteristics of collaboration which will be discussed in Section 3.5. In other words, emotion functions are important because they provide social characteristics that the self needs to manifest in its collaborative behavior.

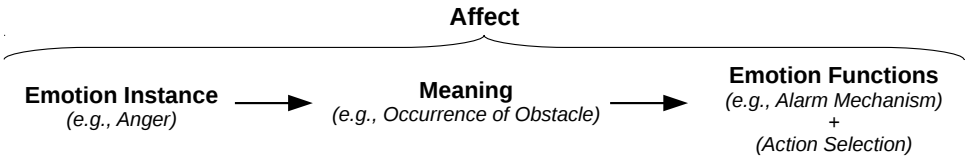

Figure 3.2: Interpretation of an emotion instance to emotion functions.

In the rest of this section, I briefly describe how ten different emotion functions are related to the collaboration context. There are other emotion functions, such as learning and memory control, which are outside of the scope of this thesis. In my implementation, I will initially focus on three of the most important of these emotion functions, namely *Adaptation, Social Regulation* and *Motivation*.

3.4.1 Action Selection

Action selection is the function in which emotion instances influence choosing the most appropriate action out of a repertoire of possible actions at a point in time. This function influences the Coping mechanism and results in the consistency of the self's actions based on anticipated emotional responses of the other and the satisfaction of the shared goal.

3.4.2 Adaptation

Adaptation is the raison d'être of emotions. It helps the self to properly respond to changing challenges in a dyadic social context by accommodating its behavior. Adaptation is a specialized problem-solving technique implicating the necessity of the self's emotional states for short and long term behavior changes during collaboration.

3.4.3 Social Regulation

Social regulation by emotions is the process which enables the self to communicate internal Mental States through the expression of emotions in a social context. It can assist the self to regulate various social interactions required in the course of a collaboration, such as conflict resolution and negotiation. Emotional expressions influence the other's behavior by triggering the other's inferential processes and emotional reactions [90].

3.4.4 Sensory Integration

Sensory integration can guide the self through the course of a collaboration by sustaining rich-sensory tasks to demonstrate more effective collaborative behaviors. It benefits the self by anticipating a certain type of inferential process to the other's mental and emotional states. For instance, perceiving fear in the other can lead to an increased focus of attention on the ongoing task, or discerning anger can raise the probability of avoiding current events (generated by the self) by the other.

3.4.5 Alarm

The alarm mechanism is a purely reactive and pattern-driven process [161]. It accounts for persuading the self that an undesired or unsatisfactory condition happened in the past, and since then, has persisted in the self's mental states. The alarm mechanism also provides the self a rapid reaction to the external or the in-

ternal events. The self will be able to interrupt the deliberative processes and show quick behavioral reactions. For instance, the self can consider corrective actions when a high probability of anticipated failure occurs during the collaboration.

3.4.6 Motivation

Motivation is a goal-driven emotion function associated with the self's behaviors. There is a motive behind every intentional action created by the Motivation mechanism. This motive is computed based on underlying beliefs relying on the evaluative role of emotions. Therefore, the motive behind any behavior carries an anticipated value of the future consequence for that behavior. It also reveals the belief foundation of a behavior. Consequently, the self can apply this function of emotions to a) cope with certain types of problems, and b) infer the other's mental state based on each action.

3.4.7 Goal Management

The goal management function identifies the existence or the need for a high priority goal for the self. These goals include both private goals and shared goals. Emotions provide an evaluation mechanism for the self to choose or reprioritize goals at each point in time. This function of emotions can impact the self's behavior with respect to the dynamics of interaction during the course of a collaboration.

3.4.8 Attentional Focus

Emotion instances and the patterns generated by the Appraisal mechanism are directly linked to the attention mechanism of the self. Both positive and negative results of a cognitive evaluation of events can change, maintain, or intensify the self's focus of attention. For instance, negative emotions, e.g., fear or anger, can influence the self's focus of attention by orienting the self towards the events [54]. Positive emotions, e.g., happiness, can broaden or expand the self's focus of attention from

details of the events to their general features [56].

3.4.9 Strategic Processing

The occurrence of new events can lead the self to rapid and/or strategic responses. The Coping mechanism contains various strategies associated with different components of the Mental States, e.g., belief or intention-related strategies. The content of the self's Mental States changes as time passes, which creates internal events causing the Coping mechanism to choose an appropriate action. The Appraisal mechanism allows the self to demonstrate a rapid response or strategically prioritize the current internal events generated based on the changes in the Mental States. For instance, is a mild, reactive facial expression an adequate response to the other's current utterance or does the self need to show a stronger behavior? Is it the new belief about the current state of the collaboration, or is it the new intention pursuing the self's private goal that the self needs to cope with? Thus, appraisal patterns and emotion instances impact the self's strategic processing.

3.4.10 Self Model

Emotions can be a representation of how the self interprets the collaboration environment. The self can generate or update beliefs about its self-model when faced with unambiguous events and apply the same self model when confronted with events possessing more ambiguity and uncertainty. Creating a self model can also help the self to demonstrate more consistent and coherent behaviors when similar situations occur during the collaboration. This reliability in the self's behavior can help the other to predict the self's responses during collaboration.

3.5 Social Characteristics of Collaboration

Collaboration is a coordinated social activity in which participants work together to perform tasks achieving a shared goal. It enables individuals to work together in

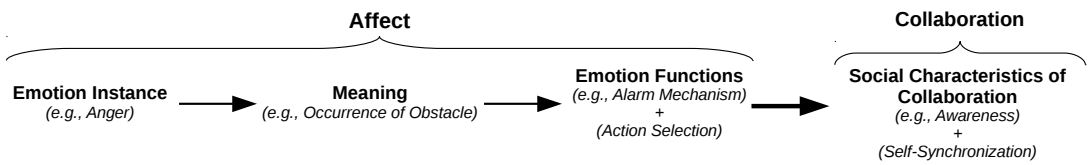

Figure 3.3: Relation between affect and social characteristics of collaboration.

a shared workspace. Conceptually, collaboration involves a number of social characteristics, each of which represents a "social skill." Each of these social skills is associated with the role emotions play in one's social life. These social characteristics manifest the self's collaborative behaviors generated based on emotion functions. In other words, the interpretation of an elicited emotion to particular emotion functions (see Section 3.4), results in the manifestation of social characteristics. More specifically, emotion functions are the cause and social characteristics are the effect of their influence on the self's behavior during collaboration. For instance in Figure 3.3, *Alarm Mechanism* and *Action Selection* are two emotion functions that result in *Awareness* and *Self-Synchronization* which are perceivable social characteristics of collaboration. This figure shows that the elicitation of anger is interpreted as the occurrence of an obstacle (see Section 2.1.1). The meaning of the anger activates different emotion functions depending on the situation, which can be interpreted as awareness and self-synchronization by the human collaborator. The same mechanism can be applied by the self based on the anticipated beliefs of the other (theory of mind) to update the other's model to cope with the situation.

In the rest of this section, I describe eight different social characteristics of collaboration and how they are related to emotion functions. In my thesis, I will focus on two or three social characteristics of collaboration, e.g., motivation and self-synchronization, that can be achieved by the implementation of emotion functions mentioned in Section 3.4.

3.5.1 Awareness

Awareness is the self's ability to manage and understand its own emotions and those of the other, and the ability to express emotions accordingly. Managing emotions refers to the self's capacity to regulate its internal emotions and to direct them towards constructive activities. Emotions, as a crucial evaluative substance of the cognition, can increase the accessibility and significance of discrepancies between Mental States, leading to higher awareness of one's sense of self and one's collaborator. In my theory, beliefs about the outcome of the appraisal and reverse appraisal [43] provide an internal and social understanding of the world. Ultimately, the self and social awareness helps our agent to become part of a collaboration process, while adopting and maintaining a shared goal.

3.5.2 Motivation

Motivation as a social characteristics of collaboration is goal-driven and concerned with self's actions and what determines those actions. There are several motivation theories describing various functional aspects of motivations. These theories can be categorized into two different groups, the *self-regulatory* and the *purposive* motivations [63]. As a regulatory process, motivation restores internal cognitive equilibrium of the self as well as self's social stance during collaboration. From a purposive perspective, the motivation process emphasizes the goal-directed nature of behaviors by anticipating the utility of self's individual actions or series of actions reaching to the private or the shared goal. In both self-regulatory and purposive perspectives, emotions serve the Motivation mechanism, offering their evaluative nature to assess the internal and external events. This process drives our agent to gain consensus in problem solving or development during collaboration.

3.5.3 Self-synchronization

Self-synchronization represents the temporal and spatial relationship of the self with the collaboration environment [106]. Adaptation as a function of emotions (see Section 3.4) play a crucial role in short and long-term behavior changes in our agent enabling self-synchronization. Furthermore, the sensory integration function of emotions can impact the synchronization procedure in the sensory level by filtering or blocking of the input data. For instance, the self can perceive the other's anger and consequently infers that the other will not totally pay attention to what the self says. The prominent role of emotions in the synchronization process during collaboration extends to their other functions in different levels of cognition including action selection, not only based on a plan but according to the self's emotional states; goal management, whenever a new goal or a sub-goal is required to be created or reprioritized; alarm mechanisms for informing the self by interrupting other cognitive processes; attentional focus, by changing the self's focus of attention between different existing saliencies in the collaboration environment; and strategic processing, by influencing the actual decision making processes of the self. As a result, our agent decides when things need to happen during the course of a collaboration.

3.5.4 Participation

It is suggested that the benefits of the collaboration arise from active participation in interaction and verbal communication with a partner who has a different perspective, either due to more knowledge, or a different viewpoint [91]. Active participation in doing tasks and communicating during the collaboration is directly related to the evaluative factors of one's cognitive processes. These evaluative factors are intertwined with different aspects of performance assessment hinged on the core concept of the collaboration, the shared goal . In order to achieve the shared goal, the self employs the evaluative nature of the appraisal process to be able to assess and participate in the ongoing collaboration. Emotion instances provide appropri-

ate, accountable and communicative signals (verbal and non-verbal behaviors) both for the self, through the manipulation of the Appraisal and Belief Formation mechanisms, and for the other, through the multi-modal emotional expressions which make the other socially aware of the self's internal states. Consequently, our agent participates in collaboration while expecting the other to participate to achieve the shared goal together.

3.5.5 Mediation

Collaborations can include disputes needed to be intervened and solved by the collaborators. Negotiation is the bargaining process between parties seeking to discover a common ground and reach an agreement to settle a matter of mutual concern or resolve a conflict. Collaborative negotiation is an interest-based, constructive negotiation (as oppose to competitive negotiation) in which the self and the other are seeking a fair and equitable agreement without having an always-conceding approach. During the collaborative negotiation the collaboration parties openly discuss their needs and try to create as much mutual value as they can. Therefore, the collaborators try to use and understand the feelings, deeper interests and motives of their collaboration partner as well as their own. Emotions, once again, become important for the self in assessing the other's offers and counter offers. Additionally, the self can communicate the result of this internal assessment through the expression of emotion instances. In the inverse of the same procedure, the reverse appraisal assists the self to perceive and interpret the other's emotional expressions and observable behaviors. Hence, our agent would be able to assess the negotiation events and to communicate its internal states during negotiation and therefore, find a middle ground according to the mutual interests and agreements.

3.5.6 Reciprocity

Reciprocity is an adaptation process by which individuals monitor their contributions in light of their partner's contributions and make adjustments accordingly [36]. Reciprocity is fundamental for creating mutually beneficial and non-zero-sum outcomes through social exchange of resources and coordination of joint activities. In fact, the adaptations associated with coordination using mutual beliefs and the shared goal concepts create a different, yet another important aspect of the reciprocity in collaboration. It leads to punishment or avoidance of collaborators with detrimental actions and to reward in contrasting situations.

There are several underlying mechanisms helping the self to perform reciprocal behaviors [36]. First, a tacit *resource monitoring mechanism* is required for the self to monitor its own as well as the other's resources contributions over the course of the collaboration. This monitoring process essentially requires the existence of an assessment process on collaborators' behaviors. The Appraisal mechanism provides this required evaluative functionality. This notion of evaluation process helps the self to a) perceive other's affective responses, and b) communicate its own. Second, a *reciprocal behavior failure detection mechanism* is needed to detect the other's failure in vital reciprocal situations (which can be met with the guidance of the Appraisal mechanism). Third, a *resource adjustment mechanism* is crucial for a collaborative agent to adjust its own social exchange behavior and also to influence the other's reciprocal behavior. The self can apply the emotion instances, e.g., guilt and gratitude, as inducements to get the other to return favors. The same mechanism, e.g., anger and sadness, can be applied by the self to punish or avoid the other based on his behavior. As a result, our agent will share resources and thoughts, and will expect sharing in return through reciprocity.

3.5.7 Reflection

Reflection is an attempt to make the implicit, explicit in order to learn from experience. There are two different types of reflection [151]. The first, *reflecting on action*, helps the self to think back on behaviors of the other occurring because of the self's last action. This process uses the Appraisal mechanism to evaluate the other's response, and serves the self by reshaping what the self does, by modification of the underlying beliefs and intentions. Additionally, reflection helps the self to learn the consequences of each action, using the content of its own Mental States, including beliefs about the other's Mental States. Consequently, the outcome of this evaluation enables the self to update its user model of the other. These processes represent the deliberative side of the emotion functions. The second type of reflection, *reflecting in action*, helps the self when a familiar routine produces an unexpected result; an error stubbornly resists correction; or, the meaning of an event has changed because of changes in the self's evaluative and/or interpretive processes. All these changes produce pleasant or unpleasant results for the self leading to different emotion instances, e.g., surprise, during collaboration. The self can intentionally ignore the events signaling these emotions and not attempt to change its own focus of attention, or in contrast, can show a quick emotional reaction, for instance because of the occurrence of an unexpected event, in which the latter represents the reflexive function of emotions. Subsequently our agent can process the collaboration status and consider alternative actions at each point in time, when required.

3.5.8 Engagement

There are three different forms of engagement: *behavioral, emotional* and *cognitive* engagements [4]. The *behavioral engagement* is associated with self's sustained behavioral involvement by taking actions during collaboration. The self's behaviors that can be indicative of behavioral engagement include persistence in taking ac-

tions, maintaining the focus of attention during collaboration, asking questions from the other whenever it is required in the current state of the world, and contributing to achieving the shared goal. The *emotional engagement* is associated with self's emotional reactions to its collaborator. These reactions are aimed at showing empathy to the other in different cases, for instance, the occurrence of the failure or disruption in one of the collaborators' tasks, e.g., anger or frustration. In general, the self can empathize with the other in the event of the appearance of any anticipated negative emotion. The emotional engagement can also help the self to evaluate the level of satisfaction based on the performance of the individual tasks from its own and the other's point of view. This evaluative process will help the Motivation mechanism to form the most appropriate intention to maximize the performance of the collaboration according to the shared plan. The *cognitive engagement* is associated with whether the self is willing to get involved in solving problems occurred during the collaboration. The Appraisal and the Motivation mechanisms serve the self to evaluate the status of the collaboration and form new intentions for actions. The Intention Formation process helps the self to find the proper solution for an ongoing issue even if the action is a part of the other's task. Consequently, our agent will proactively engage in cases when an unexpected problem occurs during collaboration, rather than just waiting to see what the other's demand is.

3.6 Attributes of Mental States

Mental states are conscious states of the mind providing the content for cognitive processes. *Affective Motivational Collaboration Theory* operates with the following Mental States: beliefs, intentions, motives, goals and emotion instances. These Mental States possess attributes, each of which provides a discriminating and unique interpretation of the related cognitive entities. The self uses Mental States' attributes whenever there is an arbitration in the internal cognitive processes. The Appraisal mechanism and the Motivation mechanism play an essential role in com-

puting the value of these attributes. I provide more details about these attributes in this section.

3.6.1 Attributes of Beliefs

The attributes of a belief are involved in arbitration procedures within different processes in the *Affective Motivational Collaboration Theory*. They impact a range of these processes from, the formation of new beliefs, the evaluation of an external event by the Appraisal mechanism, generation of new motives, updates on collaboration plan, to the activation of coping strategies and ultimately the self's behavior. The following six attributes of beliefs are most related to *Affective Motivational Collaboration Theory*.

- ***Strength:*** Belief strength is about how strong the self holds salient beliefs about an object, an entity, or an anticipated behavior. It can be measured through scales, for instance, how probable or likely that belief is, or just whether it is true or false. The strength of a belief can impact the self's intention attributes such as the *certainty* or *ambivalence*. A belief can be strong, but not accurate.

- ***Accuracy:*** Accuracy of a belief is the relation between that belief and the truth which that belief is about. The accuracy of a belief can be measured by looking at how closely that belief can relate to the truth. The accuracy of a belief as a gradational property can be used in evaluative processes of the self, i.e., Appraisal. It can also impact the self's other goal-driven processes, e.g., Motivation, by updating the utility function(s) with respect to the estimated belief accuracy.

- ***Frequency:*** The frequency of a belief is related to how regularly it appears as the result of an internal or an external event. The frequency of beliefs can impact attributes of the self's other Mental States. For instance, beliefs

forming or maintaining intentions with direct experiences (see Section 3.6.2) are more likely to occur frequently.

- ***Recency:*** The recency of a belief refers to how temporally close a particular belief is to the current state of collaboration. The recency attribute of the self's belief can bias (recency effect) the evaluation processes of the cognitive mechanism during collaboration. It can create a tendency to weight recent events more than earlier ones whenever it is required according to self's Mental States. The self can allow or hinder this tendency to adopt an appropriate Coping mechanism.

- ***Saliency:*** The saliency of a belief is a cognitive attribute that pertains to how easily the self becomes aware of a belief. This property of a belief has a prominent influence on the attention mechanism during collaboration. It directs the self's focus of attention to the most pertinent spatio-temporal salient internal or external event(s). Although belief saliency can determine the self's focus of attention, the self does not necessarily select an action based on the salient events.

- ***Persistence:*** It is argued that beliefs form and change due to cognitive and social considerations [26]. Persistent beliefs are very resistant to these changes. However, even persistent beliefs can change. Persistence of goal-related belief(s) influences the self's intentions and subsequently behaviors.

3.6.2 Attributes of Intentions

The attributes of an intention influence several processes in *Affective Motivational Collaboration Theory*. They are involved in mechanisms such as Appraisal and Motivation as well as other Mental States, e.g., goals. One of the most important uses of intention attributes is to moderate the intention-behavior relations [37]. Ultimately, the self can show more consistent behavior with respect to its own

preceding behaviors and current state of the collaboration. I decided to include the following five intention attributes extracted from the psychology literature in *Affective Motivational Collaboration Theory.*

- ***Temporal Status:*** The temporal status of an intention can be defined as the extent to which an intention remains consistent over time. The self needs to maintain the stability of its intentions as time passes until the task is performed. Temporally stable intentions helps the other to accurately predict the self's behavior. The anticipated cognitive load of perceiving the self's task by the other impacts the temporal stability of the self's intention. In other words, the temporal stability of an intention moderates the intention-behavior relation of the self during collaboration.

- ***Direct Experience:*** The direct experience of an intention refers to whether the self previously has performed a task based on the similar intention. The self can refer to the corresponding Mental States of the intention directly experienced in the past before taking a new action. The Mental States associated with the prior experience of an intention can influence the appraisal of a new event requiring the self to perform the same task. For instance, the existence of a direct experience of an intention impacts the degree of the *expectedness* and *controllability* of an event during the collaboration which ultimately guides the Coping mechanism to produce an appropriate behavior.

- ***Certainty:*** The certainty of an intention is determined by the quality of the underlying motive and the beliefs associated with that motive. The more *strong, accurate, frequent, recent, salient* and *persistent* a set of pertinent beliefs of the self are, the more chance the related motive has to be selected. Since the certainty of an intention depends on the associated motive, the nature of the pursued goal also implicitly impacts the certainty of that intention. A goal with higher *specificity* (see Section 3.6.3) value influences the Belief Formation process of a new motive by increasing the certainty of the affiliated

intention. The certainty of an intention is an important moderator of the self's intention-behavior consistency.

- ***Ambivalence:*** The Mental States of the self might contain contradictory intentions towards the pursuit of the same goal, which makes those intentions ambivalent. For instance, the self might already have an intention to perform a task according to the shared plan, while the Appraisal and the Motivation mechanisms dynamically provide a new motive forming a new opposing intention. Furthermore, ambivalent intentions can occur because of the contrast between the self's private goal and the shared goal during the collaboration. The ambivalence attribute of an intention is against the intention-behavior consistency of the self.

- ***Affective-Deliberative Consistency:*** The self's intentions possess an affective and a deliberative component. The affective component refers to the emotion instance and in general the affective evaluation of the self's intention towards its own behavior. However, the deliberate component refers to the self's actual intention which is formed either based on the existing shared plan or through a new motive generated by the Motivation mechanism. For instance, as an example of affective-deliberative inconsistency, the self can appraise the formation of the current intention as an *urgent* and *uncontrollable* one (which leads the self's emotion towards anger), despite the fact that performing the task associated with this intention is required for the satisfaction of the shared plan. In general, mutually consistent affective and deliberate components of an intention positively impacts the consistency of the self's intention and behavior.

3.6.3 Attributes of Goals

The attributes of a goal also impact the processes in *Affective Motivational Collaboration Theory*, especially the processes involved in Motivation and Appraisal

mechanisms. The attributes of a goal are important because the Motivation and the Appraisal mechanisms in this theory are goal-driven and attribution of the goals according to the self's standards provides coherency of the processes and their outcomes. I discuss the three most relevant goal attributes in this section.

- ***Proximity:*** Goals can be distinguished by how far they project into the future during the collaboration. Proximal (short-term) goals result in more related motives and subsequently better self and social-regulation than temporally distant goals. Proximal goals impact the self's behaviors by influencing the evaluation process of the Appraisal mechanism. As a result, the self can determine the collaboration progress towards the shared goal more accurately while operating based on proximal goals.

- ***Specificity:*** Goals incorporating specific performance standards are more likely to enhance the self's self-evaluation than general goals. Specific goals raise the self-evaluation performance because they provide a more accurate baseline for the mechanisms, e.g., Appraisal or Motivation, that the self needs for self-evaluation during collaboration. Consequently, by increasing the self-evaluation performance, the self can compute more accurate anticipated self-satisfaction. For instance, holding an object A in a particular position with respect to an object B for a certain amount of time and welding them with a material C is a more specific goal than a general goal of installing an object A on object B.

- ***Difficulty:*** The goals that are moderately difficult have the most impact on the Motivation mechanism, and ultimately the self and social regulation processes of the self. Conversely, overly easy or impossible goals usually do not motivate an individual to achieve the goal. Difficult goals increase the probability of a motive's *failure disruptiveness*, and overly easy goals decrease the *importance* of the related motive; in both cases the motives have less chances to form new intentions. The lower chance of new intention formation will disrupt

the self and social regulation since the self cannot regulate internal processes and influence the external world without forming appropriate intentions to take required actions. The existence of a partial shared plan, dependency on the other to perform a task, the failure of the same or similar task in the past and the conflict between the self's private goal and the shared goal all increase the difficulty level of a goal.

3.6.4 Attributes of Motives

According to Sloman, motives can be compared on various dimensions [160]. This comparison is based on motive attributes. In *Affective Motivational Collaboration Theory* motives are formed based on the self's existing Mental States under the influence of Appraisal mechanism. The existence of different Mental States, and the results of self appraisal as well as the reverse appraisal of the other can cause a variety of motives to be formed. The Motivation mechanism needs a set of attributes to compare newly generated motives and choose the one which is most related to the current state of the collaboration. I have chosen the following five motive attributes as most related to the collaboration context.

- ***Insistence:*** The insistence of a motive defines the "interrupt priority level" of the motive, and how much that motive can attract the self's focus of attention. This dimension of motive is associated with what the Appraisal mechanism considers as *relevance* and *desirability* when evaluating an event. Beliefs about successive subgoals and the other's anticipated Mental States influence the insistence attribute of a motive. Insistent motives have higher priority and are able to interrupt self's ongoing tasks. The insistence of a motive is a function of the importance, urgency and the elapsed time for that motive.

- ***Importance:*** The importance of a motive is determined by the corresponding beliefs about the effects of achieving or not achieving the associated goal. It is a function of belief attributes (including strength, accuracy, frequency, recency,

saliency and persistence) and the current goal. For instance, if a motive is supported by a belief about the current goal with relatively high attribute values, that motive will become important for the self.

- ***Urgency:*** The urgency of a motive defines how much time the self has to acknowledge and address that motive before it is too late. The urgency of a motive is a function of beliefs about the other's mental states as well as the required and the estimated time to fulfill the associated goal. For instance, the self responds to an urgent motive due to the existence of an important anticipated outcome for the other, and limited time to accomplish the corresponding tasks, even if those tasks are not important for the self.

- ***Intensity:*** The intensity of a motive determines how actively and vigorously that motive can help the self to pursue the goal if adopted. Motives with higher intensity will motivate the self to apply certain types of coping processes for an obstructed goal to avoid termination of the collaboration. Motives with higher intensity cause the self to find alternative solutions for the problem rather than abandoning the goal and ultimately the collaboration.

- ***Failure Disruptiveness:*** The failure disruptiveness attribute of a motive determines how disruptive failure is to achieving the corresponding goal. In other words, it gives the self a measure of the pleasantness of achieving a related goal. This attribute directs the self's behavior toward positive and negative outcomes during collaboration. It also plays a role in performance assessment processes when the self needs to compare its competence level on a given task relative to the other.

3.6.5 Emotion Instances

Each emotion has its own functionality in either intrapersonal or interpersonal level. These emotions not only regulate the self's internal processes, but also assist the

self to anticipate the other's Mental States. In this section, I provide the description of some of the emotions that can be elicited during collaboration, and are involved in our scenario (see Section 3.1.2). In this document, to avoid controversial issue of whether virtual agents or robots can feel emotions, I am going to use the convention of having emotions by the agent or the robot. The agent can also possess belief about an emotion instance which is similar to having belief about any other proposition.

- ***Joy:*** Joy is the state of an individual's well-being and is associated with the sense of successful achievement of a goal. Joy reveals one's sense of pleasure which implies an impending gain for the individual.

- ***Anger:*** Anger can be elicited by an unfair obstacle, hindering the individual's goal attainment and it is usually triggered by some external event (e.g., threat) which provokes a behavioral reaction. Anger functions to set boundaries or escape from dangerous situations, and implies an urgent desire for justice.

- ***Hope:*** Hope is the result of an optimistic evaluation of an event by an individual having expectations of positive and desirable future outcomes related to that event. It is usually a poignant assimilation of the present discontent and the future content implying an imagined or anticipated successful future goal state.

- ***Guilt:*** Guilt is based on self-condemnation in response to a negative outcome of one's self performance evaluation. It is caused by the violation of others' beliefs about the self, and others' standards and bearing significant responsibility due to that violation. The occurrence of guilt usually implies the desire to atone in social context.

- ***Pride:*** Pride is a product of the satisfied sense of one's own actions or decision outcomes. It implies the self-approval of the evaluation oucomes of one's own actions. Pride is associated with the achievement motivation (see Section

2.3.2) wherein succeeding at a particular goal motivates the corresponding action.

- ***Shame:*** Shame is produced when one evaluates one's own actions or behaviors and attributes failure to the outcome. The individual focuses on specific features of self which led to failure. Shame implies the existence of remorse.

- ***Worry:*** Worry is one's emotional attempt to avoid anticipated potential threats or unidentified undesirable events. The individual's concern can be about a real or an imagined issue. Worry implies a fear of a future failure about which one should make a decision or take an action at present.

CHAPTER 4

COMPUTATIONAL MODEL

In this chapter, I will provide a computational model for an autonomous agent to apply *Affective Motivational Collaboration Theory* to interaction with a human. It will help a virtual agent or a robot to effectively collaborate towards a shared goal. It derives and synthesizes concepts from the psychological and sociological phenomena described in Chapters 2 and 3, and implements this theory in the form of formulas, rules and algorithms. The mechanisms involved in this model are those I discussed in Chapter 3 (see Figure 3.1). This chapter provides more details on these mechanisms (see Figure 4.1). I will explain the input, output and function of each of these processes in Section 4.2. In this chapter, I will use the word "agent" to refer to either a robot or a virtual agent, since this computational model is applicable to both of them. I will also use the word "mechanism" for all the boxes depicted in the Figure 4.1 (except Mental States). Each mechanism can contain one or more major processes.

4.1 Four Computational Examples Based on the Scenario

In this section, I provide four computational examples based on the scenario in Section 3.1.2. In each example, I show how the various mechanisms discussed in this chapter and Chapter 3 are involved in the collaborative interactions between the Robot and the Astronaut. These examples help the reader to understand how different mechanisms, e.g., Appraisal and Collaboration, function in a dyadic inter-

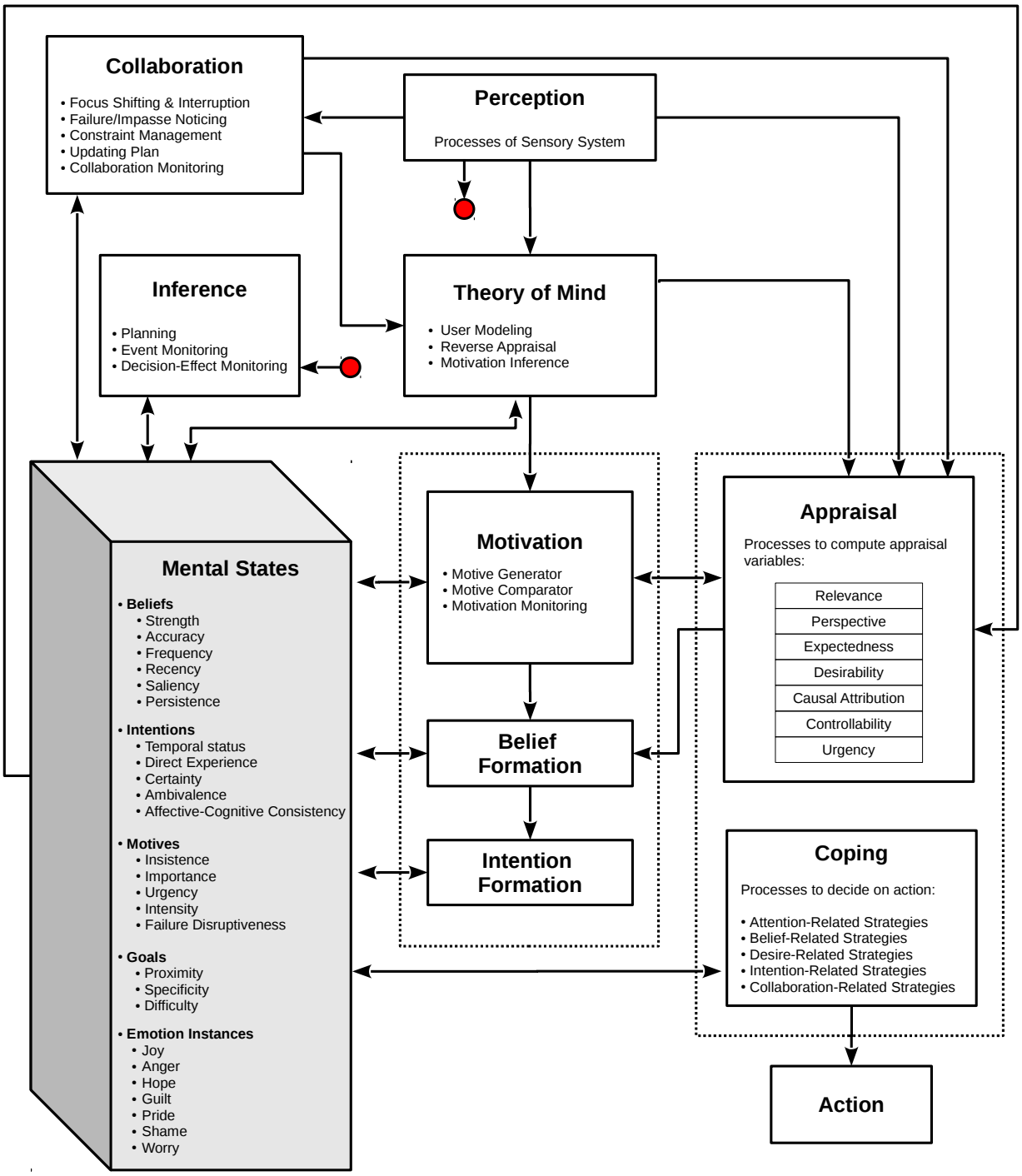

Figure 4.1: Computational model.

action. The arrows (see → below) indicate how do the mechanisms relate to each other in each step (turn) of the collaboration. Consequently, the arrows reveal the necessity of the involved mechanism in the computational model of the *Affective Motivational Collaboration Thoery*.

4.1.1 Interruption and acceptance of a new task

The Robot's current task is calculating the predicted level of energy consumption for the next task. The Astronaut interrupts the Robot with another task, i.e., holding the panel in a certain way. Interrupting a collaborator typically has negative effects on task success, completion time, and the collaborator's affective state. These negative effects can be mitigated if the interuptee considers the necessity and importance of the new task, its relation to the status of the ongoing private and the shared goals, and in general, the cognitive equilibrium of the interrupted collaborator. Since handling interruptions is important for the collaborative agents, I provide the following example to show: a) how this common issue of collaborative agents is addressed by my proposed computational model, and b) how the appraisal of the environment makes changes in collaboration structure. The agent should be able to accept or reject the interruption; in this example it accepts.

(2) *Astronaut*: At this point you should be careful how you hold the panel. Turn the right side 45 degrees towards me.

- **(Perception):** Astronaut interrupts the Robot's current task.
- **(Appraisal):** Robot evaluates Astronaut's request as a new event. Robot appraises the event as unexpected (low *Expectedness*) and controllable (high *Controllability*).
- **(Mental States → Motivation: Motivation Monitoring):** Robot finds an already existing motive to accept Astronaut's interruption.

- **(Belief Formation → Mental States):** Robot updates current beliefs and intentions to postpone the current task as requested by the Astronaut.
- **(Collaboration: Collaboration Monitoring + Focus Shifting):** Robot infers that its Mental States are updated; consequently it decides to update its plan and shift its focus of attention to the new task (holding the panel).
- **(Mental States → Coping: Intention-Related Strategies: Planning):** Robot updates the plan with the new task given by the Astronaut.
- **(Mental States → Coping: Attention-Related Strategies: Seek Information):** Since the Robot appraised the event as unexpected, it forms a new intention to ask the Astronaut whether the panel is in the right position.
- **(Action):** Robot asks this question.

(3) *Robot*: Is this what you want?

4.1.2 Motivation to minimize the risk of impasse

The Robot is holding the panel for the Astronaut. The Astronaut realizes there is a problem with the connectors on the panel and perceives it as a failure for the overall task. The Astronaut informs the Robot about this impasse. I chose this example because applying motivation to reduce the possibility of impasse during collaboration is one of the crucial roles of emotions. This example shows how the reverse appraisal of the Astronaut's verbal behavior informs the Robot about the possible existence of an impasse, and how the reappraisal of the alternative tasks helps the Robot to choose and suggest the best alternative option to the blocked task in the collaboration structure.

(5) *Astronaut*: The connectors on this panel have problems and we might not be able to finish this task.

- **(Perception):** Astronaut says we have a problem.
- **(Appraisal):** Robot evaluates what Astronaut said. Robot finds Astronaut's appraisals about the panel's issue as uncontrollable, undesirable (low *Desirability*) and unexpected.
- **(Theory of Mind → Mental States → Collaboration: Failure Impasse Noticing + Collaboration Monitoring → Mental States):** Robot thinks that the Astronaut believes the ongoing task is blocked.
- **(Mental States → Motivation: Motive Generator → Belief Formation → Mental States):** Robot thinks this impasse should be solved. Robot tries to find alternatives to motivate the Astronaut to continue collaboration.
- **(Mental States → Inference: Planning → Collaboration: Updating Plan):** Robot comes up with two alternatives for the blocked task: (1) replacing the connectors on the panel, and (2) fetching a new panel from the base.
- **(Theory of Mind → Appraisal):** Robot appraises the two alternatives before discussing them with the Astronaut. Replacing the connectors on the panel, the first alternative, is evaluated by the robot as relevant (high *Relevance*), highly expected, highly likely to be desirable, and highly controllable for the Astronaut. Fetching a new panel from the base, the second alternative, is relevant, less expected than the first alternative, unlikely to be desirable but is controllable for the Astronaut.
- **(Motivation: Motive Comparator):** Because of these appraisals, Robot chooses the first alternative to suggest to the Astronaut.
- **(Belief Formation → Mental States):** Robot believes the first alternative should be discussed with the Astronaut.

- **(Mental States → Theory of Mind → Appraisal → Belief Formation → Mental States):** Robot believes that since the Astronaut is facing a negative emotion as a result of an impasse in the current task, a positive valence emotion (e.g., hope) should be expressed while talking about the solution.

- **(Mental States → Coping: Belief-Related Strategies: Wishful Thinking + Desire-Related Strategies: Positive Reinterpretation):** Robot decides to tell the Astronaut that they can continue this task since it is possible to replace the defective connector.

- **(Action):** Robot talks about the solution in a positive way.

(6) *Robot*: Don't worry! I can replace the connectors in 4 minutes. We definitely can finish this task after that.

4.1.3 Action selection and updating plan

The Robot waits for the Astronaut to give it a new task. The Astronaut asks the Robot to do two different tasks. Action selection is what the Robot needs to do in this situation and is another essential function of emotions. Emotions, in this case, strongly affect the outcome of other processes involved in action selection. The influence of affect on selecting the appropriate action makes the whole process of action selection coherent with the human collaborator's state. This is another example showing the influence of appraisal on the Robot's decisions and the related underlying beliefs which ultimately result in updating the collaboration plan.

(8) *Astronaut*: I need you to begin welding this panel and also to prepare the measurement tool for me.

- **(Perception):** Astronaut asks the Robot both to begin welding and to prepare a tool for him.

- **(Collaboration → Inference: Event-Monitoring → Mental States):** Robot checks whether these two tasks are already in the current plan. Robot finds both tasks in the plan.

- **(Theory of Mind → Mental States):** The welding and the measurement tasks are not constrained in the agent's plan. However, the Robot thinks about whether Astronaut has a preference of doing these tasks in a particular order. Robot does not find any preference.

- **(Appraisal → Belief Formation → Mental States):** Robot evaluates Astronaut's request as two separate events. Robot appraises welding task as unexpected, relevant and uncontrollable. It appraises the preparing measurement tool task as expected, relevant and controllable.

- **(Mental States → Coping: Intention-Related Strategies: Planning/Action Selection):** Robot decides to prepare tool for the Astronaut first and to begin to weld next.

- **(Theory of Mind → Appraisal → Belief Formation → Mental States):** Robot evaluates this decision based on the anticipated emotion(s) of Astronaut.

- **(Theory of Mind → Mental States):** Robot thinks about whether this order of tasks is controllable from Astronaut's point of view.

- **(Mental States → Coping: Attention-Related Strategies: Seek Information):** Robot decides to ask the Astronaut whether he agrees with this order of the tasks.

- **(Action):** Robot shares this idea of preparing measurement tool first, with Astronaut to see whether he agrees with it.

  (9) *Robot*: Do you want me to prepare the measurement tool first? Then, I can begin welding afterwards.

- **(Perception):** Astronaut agrees with the Robot on its proposed order.

  (10) *Astronaut*: Yes, that's fine!

- **(Collaboration: Updating Plan):** Robot updates the plan.

4.1.4 Goal management (reprioritization of existing goals)

The Robot has welded the panel and prepared the measurement tool for the Astronaut. But, the Astronaut faces a problem with a private goal which could hinder the achievement of their shared goal. The Robot evaluates this situation and uses its Motivation mechanism to revise the collaboration plan. This illustrates goal management, which is another important function of emotions. Emotions can be involved to create a new goal or reprioritize an existing one, whether it is a private or a shared goal. The Motivation mechanism forms a set of beliefs and intentions under the influence of the Appraisal mechanism. It also evaluates different motives to determine the one with higher priority for the current state of collaboration. Consequently, the collaboration plan will be updated with the preferred collaborative goal.

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

- **(Perception):** Astronaut tells the Robot how the last part of the installation phase is frustrating for him because of a problem in the measurement tool.

- **(Appraisal):** Robot evaluates the new situation and finds it unexpected, undesirable, urgent (high *Urgency*) and controllable.

- **(Mental States):** Robot's private goal is to go back and fetch the next panel and prepare it for installation. Also, Robot has shared goals which are to install the solar panels successfully, and to measure the quality of installation.

- **(Mental States → Motivation: Motivation Generator):** Robot creates three alternative new motives: (1) achieving to its own prior private goal, (2) prioritizing the shared goal, and (3) canceling the current task.

- **(Motivation: Motivation Comparator → Belief Formation → Mental States):** According to the theory of mind, appraisal and the current Mental States, Robot finds the motive related to the first (private) goal less *Insistent*, less *Urgent*, but *Important* with a medium *Intensity* level and a very low *Failure Disruptiveness* level compared to other motives. Robot evaluates the motive related to second (shared) goal as a highly *Insistent*, very *Important* and *Urgent* with a high *Intensity* and very high *Failure Disruptiveness* value. The motive related to the third (Astronaut's proposed) goal gets lowest *Insistence*, *Importance* and *Urgency* value as well as medium *Intensity* and low *Failure Disruptiveness*. So the Robot forms the belief that reaching the second goal has a higher priority than the first and the first has a higher priority than the third.

- **(Mental State → Coping: Belief-Related Coping: Wishful Thinking):** Robot makes a decision about preventing the shared goal from being canceled, since its private goal can be postponed, and the Robot can also help the Astronaut to overcome the existing problem.

- **(Action):** Robot tells Astronaut about this decision.

  (12) *Robot*: I see! But, I can help you with the measurement and finishing up the task as it was originally planned.

- **(Perception):** Astronaut accepts the plan.

  (13) *Astronaut*: That would be great!

- **(Appraisal):** Robot evaluates Astronaut's response and finds it expected, desirable and relevant.

- **(Belief Formation → Mental States):** Robot believes goals can be reprioritized to achieve a new plan.

- **(Mental State → Coping: Intention-Related Coping: Procrastination + Planning):** Robot decides to reprioritize the three goals considered earlier and also decides to change the plan.

- **(Mental State → Planning → Collaboration: Updating Plan):** Robot updates the plan based on the new goal (postponing private goal and replacing it with shared goal).

4.2 Components of the Architecture

The *Affective Motivational Collaboration Model* consists of ten mechanisms (see Figure 4.1) most of which directly store and fetch the data in the Mental States. The Mental States will be represented using a logical formalization, e.g. [1], and will keep all the required data about the self (agent), other (human) and the environment (including events). I will use a blackboard [52] architecture to implement this model (see Section 5.1). The rest of this chapter continues with eleven sections which first explain the Mental States and then each of the ten mechanisms in detail.

4.2.1 Mental States

The Mental States shown in Figure 4.1 comprise the knowledge base required for all the mechanisms in the overall model.

Beliefs are a crucial part of the Mental States. I have two different perspectives on categorization of beliefs. In one perspective, I categorize beliefs based on whether they are shared or not between the collaborators. The SharedPlans [74] theory is the foundation of this categorization in which for any given proposition the agent may have: a) private beliefs (the agent believes the human does not know these), b) the inferred beliefs of the human (the agent believes the human collaborator has

these beliefs), and c) mutual beliefs (the agent believes both the self and the human have these same beliefs and both of them believe that). From another perspective, I categorize beliefs based on who or what they are about. In this categorization, beliefs can be about the self, the other, or they can be about the environment. Beliefs about the environment can be about internal events, such as outcomes of a new appraisal or a new motivate, or external events such as the human's offer, question or request, and general beliefs about the environment in which the agent is situated. Beliefs can be created and updated by different processes. They also affect how these processes function as time passes.

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 human collaborator, c) formation of beliefs about self and anticipated beliefs about the other, and d) behavior selection in the Coping mechanism. First, taking actions means that the agent will intend to take an action for primitive tasks that have gained the focus of attention, possess active motives, 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 agent 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. And lastly, intentions are involved in selecting intention-related strategies, e.g., planning, seeking instrumental support and procrastination, which these strategies are an essential category of the strategies in the Coping mechanism. Intentions possess a set of attributes, e.g. *Involvement, Certainty, Ambivalence* (see Section 3.6.2) 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.

Motives are 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 agent according to: a) its own emotional states (how the agent feels about something), b) its own private goal (how an action helps the agent to make progress), c) the collaboration goal (how an action helps to achieve the shared goal), and d) other's anticipated beliefs (how an action helps the other). Motives also possess a set of attributes, e.g., *Insistence* or *Failure Disruptiveness* (see Section 3.6.4). These attributes are involved in comparison of newly generated motives based on the current state of the collaboration. Ultimately, the agent forms or updates a belief about the winning motive in the Mental States.

Goals help the agent to create and update its collaboration plan according to the current private and shared goal content and structure, i.e., the *Specificity, Proximity* and *Difficulty* of the goal. 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. The *Specificity* of goals has two functions for the agent. First, it defines the performance standard for evaluating the progress and quality of the collaboration. Second, it serves the agent 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 people motivated to attempt what they believe are impossible goals.

Emotions in Mental States are emotion instances that are elicited by the Ap-

praisal mechanism (see Section 3.6.5 for list of emotion types used in this model). The agent also keeps beliefs about these emotion instances in the Mental States. The Belief Formation mechanism creates or updates these beliefs about emotions. These emotion instances include the agent's own emotions as well as the anticipated emotions of the other which are created with the help of the processes in the Theory of Mind mechanism.

4.2.2 Collaboration

- **Input:** The input to the *Collaboration* mechanism includes all the data that affects the execution of individual tasks in the collaboration plan. This data will be provided via the different elements of Mental States including beliefs, intentions and goals. These Mental States will establish agent's initial plan and will be updated throughout the collaboration process by the Perception and Inference mechanisms.

- **Output:** The output of *Collaboration* includes all the data that is modified or created during execution of a plan in the form of Mental States. These Mental States will be used by the internal processes in the Theory of Mind mechanism. Additionally, the Appraisal mechanism will use these Mental States to evaluate the events during collaboration. These Mental States also will be used by the Inference mechanism for the purpose of maintaining the collaboration.

- **Function:** The *Collaboration* mechanism will construct a hierarchy of tasks and also manage and maintain the constraints and other required details of the collaboration specified by the plan. These details include the inputs and outputs of individual tasks, the *preconditions* specifying whether it is appropriate to perform a task, and the *postconditions* specifying whether a just-completed task was successful (which can be used as an indication of an impasse or failure). *Collaboration* also keeps track of the focus of attention, which determines the salient objects, properties and relations at each point of the collaboration.

Moreover, *Collaboration* has the ability to shift the focus of attention during the collaboration. All the other mechanisms in the overall *Affective Motivational Collaboration Model* are influenced by changes in the collaboration plan.

4.2.3 Appraisal

- **Input:** The most significant part of *Appraisal's* input data is based on the activity of the Collaboration mechanism. This data includes all the required Mental States associated with the Collaboration mechanism. For instance, the beliefs and their Strengths will be used by algorithms inside of *Appraisal* to compute the value of the appraisal variables. *Appraisal* also receives data from the Theory of Mind mechanism. This data helps the agent use *Appraisal* for inferring the human's intentions and motives based on a reverse appraisal procedure.

  The input data from the Perception is generally needed to support the evaluation of the events. The Motivation mechanism sends data to *Appraisal* whenever, for example, a new set of motives is being generated or the motive comparator wants to evaluate different motives for an arbitration procedure.

- **Output:** The output of *Appraisal* can directly and indirectly impact other mechanisms. The Motivation mechanism uses this data to generate, compare and monitor motives based on the current internal appraisal of the agent as well as the appraisal of the environment. This output data also assists the Belief Formation mechanism to create new beliefs about the current emotional state of the agent or the human as well as the corresponding appraisal of individual motives.

- **Function:** *Appraisal* is a subjective evaluation mechanism based on individual processes each of which computes the value of the appraisal variables used

in my computational model. The Collaboration mechanism needs the evaluative assistance of the *Appraisal* mechanism for various reasons. The course of a collaboration is based on a full or a partial plan 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 events helps the agent to update the collaboration plan and avoid further critical failures during collaboration. *Appraisal* also helps the agent to have a better understanding of the human's behavior by making inferences based on appraisal variables. 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 States. For instance, there are beliefs about the appraisal of the self and the other that augment the model of what collaborators have done, and what and how they are planning to achieve the current shared goal based on their emotional states. This process will be discussed in more detail under the Motivation mechanism (see Section 4.2.5). Additionally, the beliefs formed based on the appraisals can impact other mechanisms such as the Theory of Mind, Motivation and Coping, essentially including the whole computational model.

4.2.4 Coping

- **Input:** The *Coping* mechanism operates based on the data stored in different aspects of the Mental States. This data includes changes in the agent's beliefs as well as the intentions of the agent (whether they are created or altered during the collaboration), and the private or shared goals.

- **Output:** The output of the *Coping* mechanism is the data specifying the behavior which the agent needs to perform based on the current state of the collaboration.

- **Function:** The *Coping* mechanism is responsible for interpreting ongoing changes in the Mental States and adopting the appropriate behavior with respect to these changes. This component includes rules categorized into four coping strategies which are *Belief-related, Intention-related, Attention-related* and *Desires-related* strategies [109]. These rules will apply to the attributes and structures of the Mental States to cope with the internal changes as well as the demands of the environment. For example, the *Coping* mechanism will utilize certain beliefs about the self to regulate the agent's internal states, while using mutual beliefs to maintain progress in the existing collaboration. As another example, attributes of intentions can guide the *Coping* mechanism to sustain the consistency between intention and behavior.

4.2.5 Motivation

- **Input:** The most essential part of the input to *Motivation* is the Mental States, and more specifically the private and shared goals associated with the collaboration. *Motivation* also uses data from two other mechanisms, namely, Theory of Mind and Appraisal. Input from Theory of Mind is used by *Motivation* whenever new motives need to be generated or compared according to the shared goal. Input from Appraisal is used whenever the motive attributes are involved in the internal processes of the *Motivation*.

- **Output:** The output of *Motivation* includes the data required to form a new set of beliefs about the private and shared goals. These beliefs can ultimately lead to the required updates in the collaboration plan. The motives which are the output of the *Motivation* mechanism are also used by the Coping mechanism, i.e., by the desire-related strategies to choose appropriate behavior according to the goals of the collaboration plan.

- **Function:** The *Motivation* mechanism closely work with the Appraisal mechanism. The purpose of this component is to generate new motives which will

be added to the Mental States in the form of new beliefs. These motives are generated based on what the agent believes about the environment including self and the other collaborator and the corresponding appraisals. The agent uses these motives to achieve a private or shared goal according to new conditions, to interact better with a human who needs social interactions, or to evaluate the success of task performances. The *Motivation* mechanism consists of a pipeline of distinct processes. First, several motives can be generated with respect to the current Mental States. Then, a comparator decides which motive is more likely to be consistent with the current state based on the values of the motive attributes. Finally, the result will be given to the Belief Formation mechanism, and this result will be monitored until the occurrence of a new state.

4.2.6 Belief Formation

- **Input:** The input to *Belief Formation* comes from the two main affect-related components of my computational model, namely Motivation and Appraisal. *Belief Formation* uses input from Motivation whenever a motive is required for internal or collaborative reasons (e.g., the human needs positive expressions of affect) or whenever a failed or a blocked task delays the collaboration. Input data from Appraisal is used to form beliefs about the self and other's emotional states. This input data also provides the outcome of appraisals as well as emotion instances to the Motivation mechanism to create new motives.

- **Output:** The output of *Belief Formation* will either be stored as a new belief in the Mental States, or will be passed to the Intention Formation mechanism whenever generation of a new intention is required based on the corresponding belief.

- **Function:** The agent uses the *Belief Formation* mechanism to create new beliefs about the appraisal of the self and the environment including the hu-

man and the events influencing the collaboration. It also generates new beliefs about the motives and the associated goals which consequently can help the agent to maintain its cognitive equilibrium according to the collaboration progress. These beliefs also can be used by the *Intention Formation* mechanism to form new intentions.

4.2.7 Intention Formation

- **Input:** The input of *Intention Formation* is limited to the beliefs that have been formed by the Belief Formation process.

- **Output:** The main output of *Intention Formation* is the new intentions which will be stored in Mental States. This process can also update the existing intentions in the Mental States.

- **Function:** The agent uses the *Intention Formation* to form new intentions whenever they are required according to the underlying beliefs. This process is also responsible for updating existing intentions, since intentions must evolve as time passes. This process can also remove an intention because a conflict between existing intentions has occurred, or it is not required anymore because the environment has changed.

4.2.8 Theory of Mind

- **Input:** *Theory of Mind* receives its input from the Mental States as well as the Collaboration and Perception mechanisms. This mechanism uses the current Mental States to update the user model (see Section 2.5) and infer the human's Mental States (which are simpler than the Mental States associated to self). The Collaboration mechanism provides the structure of the collaboration plan, including the constraints which can be used in the internal inference processes of the *Theory of Mind*, such as reverse appraisal. The Perception mechanism

also helps *Theory of Mind* with the input data from the sensory system. I will decide later exactly what sensor information I am going to use and how.

- **Output:** The output of *Theory of Mind* will be stored in the Mental States. It can be appraised by the Appraisal mechanism as part of the evaluation of a new event. The Motivation mechanism also uses this output to generate new motives according to the current state of the user model.

- **Function:** The agent 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 the course of the collaboration. I will use one of the existing user modeling techniques in dyadic interaction setting, e.g., Bayesian user model, to build human's user model for the agent. The agent can also infer the Mental States of the human based on the reverse appraisals of the human's behavior. Another internal process of the *Theory of Mind* is inferring the human's motives on the basis of his behavior.

4.2.9 Inference

- **Input:** The *Inference* mechanism gets its input data from the Mental States and the Collaboration mechanism. Thus, in general, the input data of this mechanism is influenced by other mechanisms that can update the Mental States.

- **Output:** The output data of the *Inference* mechanism impacts the internal processes of the Collaboration, e.g., updating the plan. It also updates the Mental States used by the Collaboration mechanism.

- **Function:** The *Inference* mechanism in general performs various logical deductions required by other processes in my computational model. It is designed to ameliorate the shortcomings of the existing Collaboration theories (see Section 2.4) by providing required inferences such as dynamic planning based on

the recent changes in the collaboration environment and the internal changes in the agent's Mental States. For instance, in our scenario (see Section 3.1.2), when the Astronaut interrupts the Robot asking for a new and urgent task, the Robot needs to alter the collaboration plan to continue. *Inference* also supports essential monitoring processes during the collaboration such as event monitoring.

4.2.10 Perception

I consider the *Perception* mechanism only as a source of data to my computational model (see Figure 4.1). Thus, my computational model starts with high-level semantic representation of events (including utterances), i.e., Natural Language Processing is out of the scope of this work.

- **Output:** I will support the human side of the dialogues using predefined utterances for verbal communication with the agent. These utterances will be a part of the output data of the *Perception* mechanism. The output of the *Perception* mechanism will be given to the Collaboration, Theory of Mind and Appraisal mechanisms. I will provide a unified perception representation across all of these mechanisms.

- **Function:** The *Perception* mechanism is responsible for producing the sensory information used by other processes in my model.

4.2.11 Action

I consider the *Action* mechanism only as a sink of data in my computational model (see Figure 4.1).

- **Input:** The input to the *Action* mechanism is provided by the Coping mechanism. This data will cause the *Action* mechanism to execute an appropriate

behavior of the agent. This data has the same level of abstraction as the output of the Perception mechanism, i.e., it includes agent's utterances, primitive actions and emotional expressions.

- **Function:** The *Action* mechanism functions whenever the agent needs to show a proper behavior according to the result of the internal processes of the collaboration procedure.

CHAPTER 5

IMPLEMENTATION

In this chapter, I am going to sketch my implementation approach. I will discuss the blackboard architecture which will be used to implement *Affective Motivational Collaboration Theory*. The mechanisms and processes discussed in Chapters 3 and 4 will appear as knowledge sources in this model. Also, the blackboard will contain all of the Mental States involved in my theory.

5.1 Blackboard Architecture

There are several criteria that system designers use to consider a particular architecture for a specific application. I decided to apply the blackboard architecture based on some of these criteria including configurability, robustness and simplicity. In my work, there are multiple processes involved in each individual operation which make the configurability of the whole system challenging. Such systems, once configured, work effectively, but the job of adding or modifying components is complex and prone to breakdown. The structure of the blackboard architecture gives me the flexibility of gradually adding new components (processes) and only making changes to the control component whenever it is required. The blackboard architecture also aids the goal of having a robust overall system in the face of a failure or a malfunction of any component, since the components do not directly influence each other. Finally, the simplicity of implementing and maintaining the blackboard architecture is another reason for me to apply this architecture.

A blackboard architecture provides a computational framework for the design and implementation of systems that need to integrate large and diverse specialized modules, which require complex, nondeterministic control strategies. The blackboard paradigm defines heterogeneous problem solving representations as independent modules called *knowledge sources*. This architecture allows the knowledge sources to remain simple, while the complexity resides mostly within the centralized control component. Blackboard architectures have been used in many applications of artificial intelligence such as speech understanding systems, e.g., Hearsay-II [52], and other interpretation problems such as signal understanding [49, 97, 105], planning and scheduling [77] and machine translation [122].

A blackboard architecture consists of three components:

1. **Knowledge sources** are independent modules that contain the knowledge needed for problem solving with their own representations. They do not need to know about the existence of the other knowledge sources, but they have to understand the state of the problem-solving process and the representation of relevant information on the blackboard. Knowledge sources can be seen as specialists in sub-fields and are only able to solve sub-problems. They read and write relevant data in a blackboard and when a knowledge source produces a significant change in the blackboard, it generates an event.

   In my computational model, the mechanisms that I discussed earlier are our knowledge sources. For instance, the Appraisal mechanism is the one which evaluates the environment based on the appraisal variables and creates or updates the corresponding beliefs. The Collaboration mechanism, as another example of a knowledge source, provides the structure of the collaboration plan in the blackboard, and refers to the Mental States to update the plan.

2. **Blackboard** is used as a global memory containing objects from the solution space for sharing information input data, partial solutions, alternatives and final solutions. The blackboard tends to have a complex structure with multiple

levels of abstraction. It is a structured global memory where a solution to the problem under consideration is incrementally constructed. The blackboard in my computational model contains all the details of the Mental States including beliefs, intentions, goals and Emotion instances. These details dynamically change as time passes.

3. **Control component** makes runtime decisions about which knowledge sources to execute next for a problem solution. At any time, many competing knowledge sources may be enabled. It is the purpose of the control component to select one for immediate execution. Thus, in general, the control component selects, configures and executes knowledge sources.

   The control component will make higher level decisions about which knowledge source should be executed in my computational model at each turn. The Appraisal mechanism for instance will evaluate the environment every time a new belief is formed for internal or external reasons. The Belief Formation process will be executed whenever the Motivation mechanism generates a new motive. Furthermore, the Theory of Mind mechanism updates the user model of the human whenever the human creates a new event, or the Inference mechanism updates the existing plan in the Collaboration mechanism based on the most recent changes of the Mental States in the blackboard. These are all various examples showing how the control component of the blackboard architecture directs the knowledge sources in my computational model.

5.2 ROS Packages

The Robot Operating System (ROS) is a flexible framework for writing robot software. It is a collection of tools, libraries, and conventions that aim to simplify the task of creating complex and robust robot behavior across a wide variety of robotic platforms (see ros.org). I am going to develop my own ROS package for the *Affective Motivational Collaboration Model*. The motivation for this ROS packaging

is to allow other researchers to easily connect my model to other robots or virtual agents. This ROS package will be a wrapper (see Figure 5.1) for the whole *Affective Motivational Collaboration Model* which is implemented within the structure of a blackboard architecture. The ROS wrapper package will include everything but the *Perception* and *Action* mechanisms. This approach will enable us to conduct my studies based on a robot simulator (e.g. Baxter on Gazebo) and to apply the same code to an actual robot. This ROS package will transfer the output of the *Action* mechanism in my computational model to the appropriate action(s) on the robot or its simulator to generate the proper behavior.

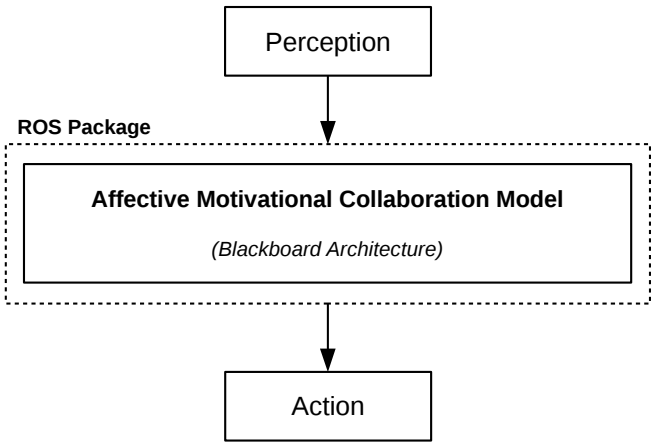

Figure 5.1: ROS wrapper for complete blackboard architecture.

CHAPTER 6

EVALUATION

The evaluation of my thesis consists of two phases. The first phase is an incremental procedure based on online questionnaires which I will use while I am developing different mechanisms in my model. The numbers in Figure 6.1 show the order in which studies based on incremental evaluation will be run. As it is shown, I will run my first study on the Appraisal mechanism to evaluate and refine the associated algorithms, and to test this evaluation methodology. Due to the overall thesis' time schedule the other online studies will remain optional (see more details in Section 6.1). In the second phase, which will be at the end of my thesis, I will conduct an end-to-end system evaluation to study the effectiveness of the affective components during collaboration with human participants, and the generalizability of the overall model (see more details in Section 6.2).

6.1 Incremental Evaluations

To evaluate the individual parts of my computational model, I will use online crowd-sourcing resources (e.g., Amazon Mechanical Turk) to collect data and use it in two ways. First, I will use one part of the data to understand how humans think and evaluate events during collaboration with respect to particular details in my computational model (see Figure 4.1), e.g., appraisal variables. Then, after implementing the corresponding computational mechanisms, I will use the remaining part of data to compare the output of the implemented mechanisms with the data collected from

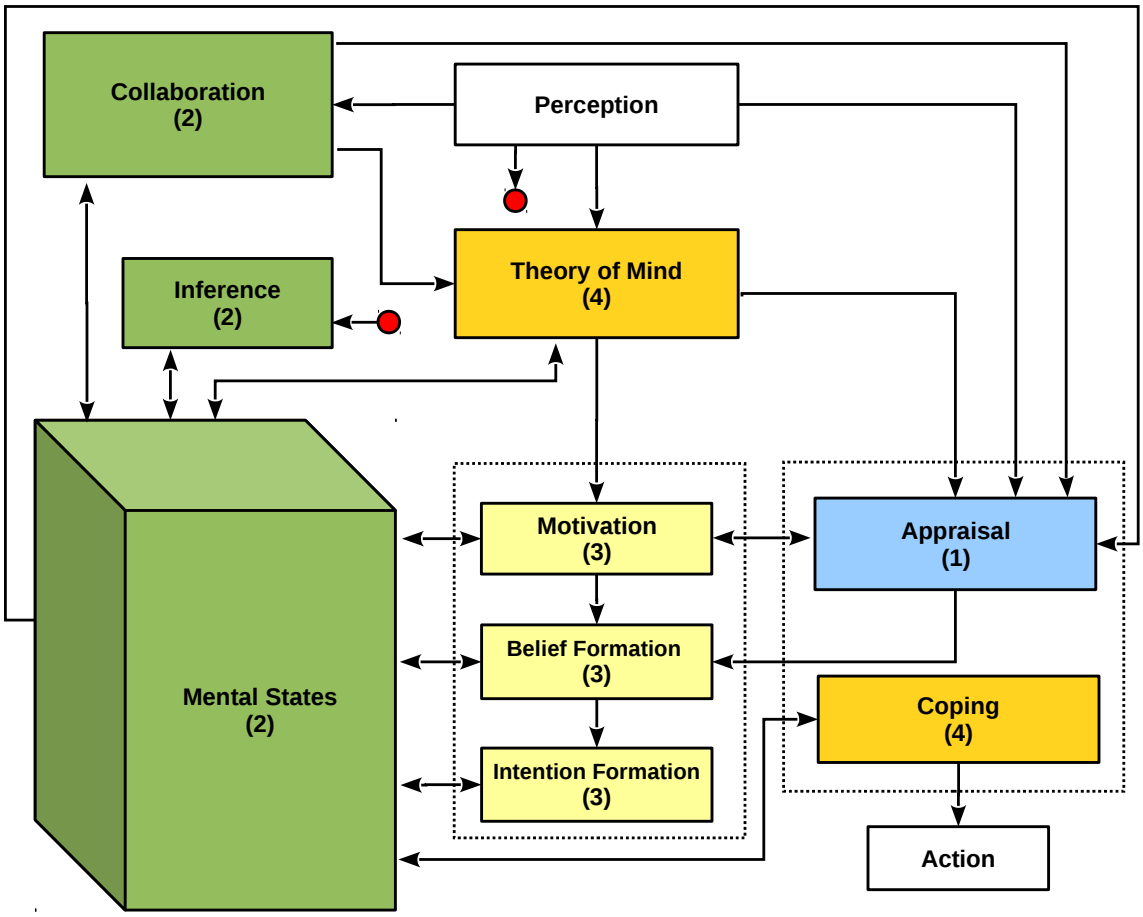

Figure 6.1: Order of incremental evaluations.

online users' responses. For instance, I will initially collect data using the online questionnaire (see Figure 6.2). I will use part of this data to investigate humans' appraisals. This analysis will help me to develop appropriate algorithms to compute the value of the appraisal variables in my implementation. The computational models, formulas, rules and algorithms developed for computing the value of appraisal variables will be evaluated after implementation. I will use rest of the data to assess the performance of these algorithms.

These incremental evaluations will be done in four different steps. For each major mechanism in my computational model (shown in Figure 4.1), I will try to see whether it can serve as an accurate model of the actual phenomena. I will determine

whether the implemented mechanisms can contribute to explaining the role of the corresponding cognitive processes. Most importantly, I want to investigate whether these implemented mechanisms are a good approximation of the underlying cognitive processes.

Figure 6.2 shows an example of the kind of online questionnaire I am planning to use to evaluate the appraisal processes (*Example 1*) and Motive attributes (*Example 2*). In it we ask subjects to rank different variables using a Likert 1-5 scale. Before answering the questionnaire participants will be given the backstory shown in Section 3.1.1. All the questions will be based on this scenario. This procedure will help me to develop, evaluate, compare and improve my computational model.

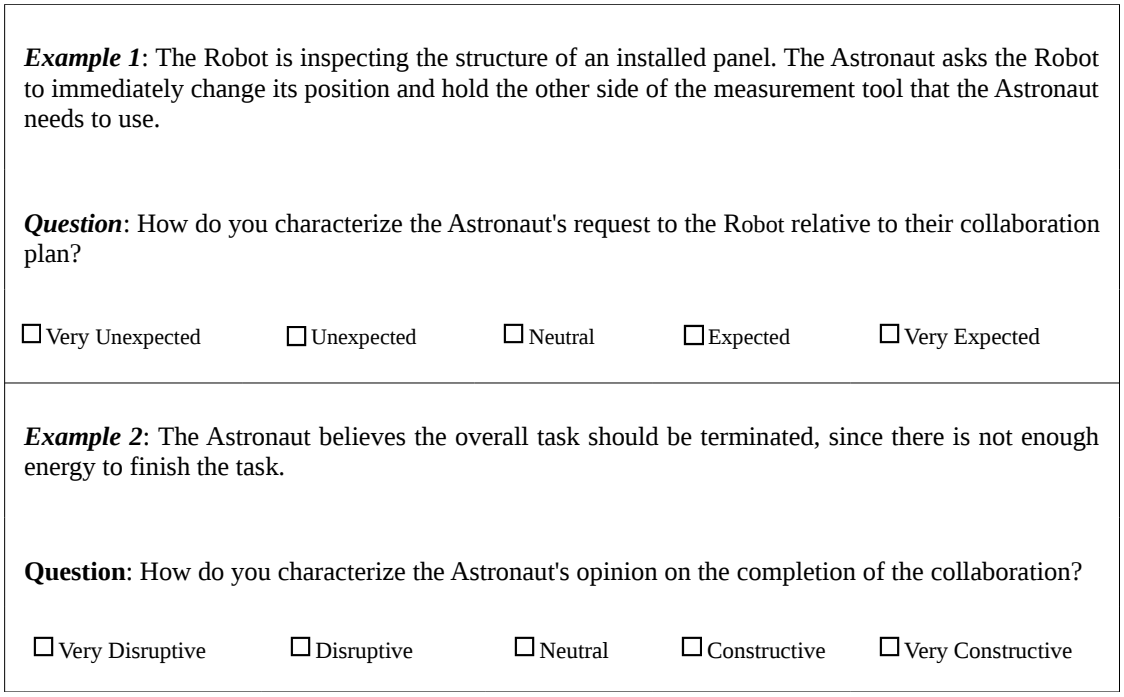

Figure 6.2: Example questions from an online questionnaire for evaluating appraisal processes (*Example 1*) and motive attributes (*Example 2*).

I planned these incremental evaluations in a particular order (shown by numbers in Figure 6.1) beginning with the evaluation of the appraisal processes. As mentioned above, other components of the overall model might be evaluated with applying the

same methodology due to the thesis schedule. This will help me to identify any potential complication in individual formulas or algorithms designed to compute appraisal variables. I will compare the output of these algorithms to what online participants provide as their own opinions about the same well-defined situation in a collaborative context. Next, I will assess the functionality of the Appraisal mechanism connected to the Collaboration mechanism in a minimal implementation of the Affective Motivational Collaboration Model. I will once again use crowdsourcing questionnaires for this evaluation (see numbers 1 and 2 in Figure 6.1). Then, after implementing the Motivation mechanism as well as Belief and Intention Formation mechanisms, I will evaluate the outputs including the motives, beliefs and intentions separately, as a distinct module (see number 3 in Figure 6.1). This step will help me to investigate the creation of the new motives and corresponding beliefs and intentions in different emotional and cognitive situations. The next and the last study based on crowdsourcing evaluations will include the Coping mechanism and the Theory of Mind (see number 4 in Figure 6.1). This step will incorporate all other mechanisms that have been implemented in the overall accomplished model. It will help me to speculate on how the interaction unfolds during the collaboration according to the internal processes of the robot and the behaviors of the human collaborator.

6.2 End-to-End System Evaluation

I will ultimately conduct a study to apply the complete *Affective Motivational Collaboration Model* on a simulated robot (e.g., Baxter in Gazebo – see Figure 6.4). In this stage, I would like to study the generalizability of the overall model in various scenarios of collaboration. I will employ a robot simulator and ROS packages (discussed earlier) to provide a collaborative environment for the user study. A mock-up user interface of my user study software is shown in Figure 6.3.

I will evaluate the overall computational model conducting a user study with

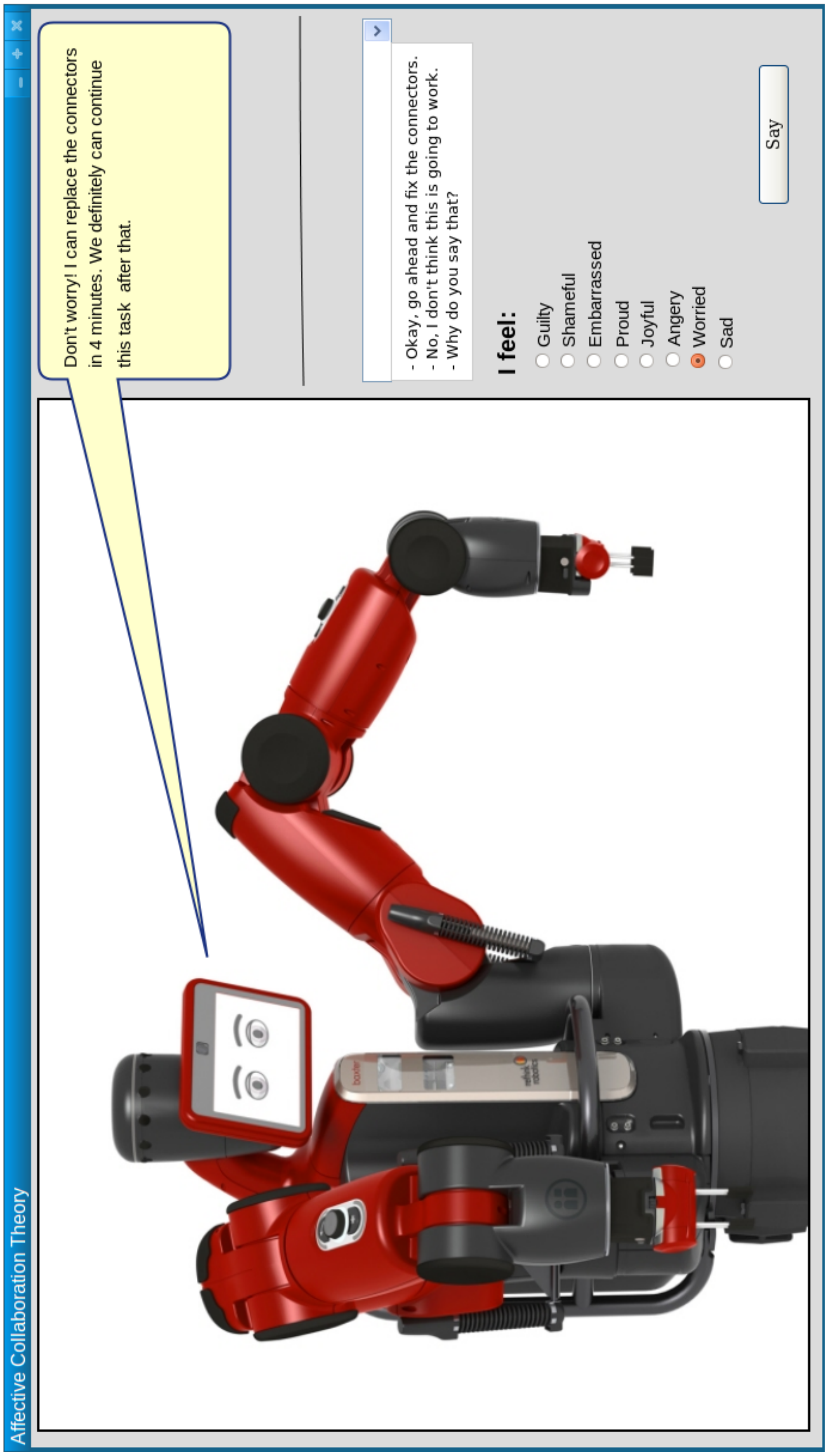

Figure 6.3: GUI mock-up for end-to-end system evaluation.

two different conditions. In one condition, the computational model will include all the mechanisms in Figure 4.1. In the second condition, the participants will collaborate with the same robot, except that the computational model that the robot uses will not include the affect-related mechanisms, i.e., *Appraisal & Coping* and *Motivation*. In both conditions, participants will collaborate with the robot through the same GUI.

I am planning to make the robot communicate with the human collaborator through predefined utterances. These utterances will communicate not only the robot's cognitive states, but also the outcome of the appraisals leading to its emotions. The upper right corner box in Figure 6.3 shows the robot's current utterance generated according to its internal state. In addition to using utterances, the robot will reveal its emotions through the facial expressions (I will try to make the Baxter's face more expressive). I will apply the knowledge and the methodology of emotional facial expression in robots that I developed earlier [154].

Figure 6.4: Simulated Baxter.

The dropdown menu in the GUI provides some (optional) utterances for the human collaborator which relate to the robot's most recent behavior. The human collaborator can choose utterance the most consistent with the state of the collaboration and his own internal states. These utterances will be also predefined by my software. Furthermore, I am going to annotate all these predefined utterances with the appropriate corresponding Mental States so that the most relevant ones at each point in time are shown in the GUI. Also in the GUI, the human collaborator can self-report the emotion instance nearest to his own. Finally, I am planning to experiment with available vision-based emotion recognition software to compare the results with self-reports.

BIBLIOGRAPHY

[1] C. Adam, A. Herzig, and D. Longin. A logical formalization of the OCC theory of emotions. *Synthese*, 168(2):201–248, 2009.

[2] 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.

[3] J. Andriessen, K. de Smedt, and M. Zock. Discourse planning: Empirical research and computer models. In A. Dijkstra and K. de Smedt, editors, *Computational psycholinguistics: AI and connectionist models of human language processing*, pages 247–278. Taylor & Francis, 1996.

[4] J. J. Appleton, S. L. Christenson, and M. J. Furlong. Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in The Schools*, 45(5):369–386, 2008.

[5] Aristotle. *The Nicomachean Ethics*. George Bell and Sons, 2009.

[6] M. Asada, K. F. Macdorman, H. Ishiguro, and Y. Kuniyoshi. Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Robotics and Autonomous Systems*, 37:185–193, 2001.

[7] J. W. Atkinson. *Motives in fantasy, action, and society: a method of assessment and study*. Van Nostrand, 1958.

[8] J. Bach. The MicroPsi Agent Architecture. *In Proceeding of ICCM-5*, pages 15–20, 2003.

[9] J. Bach. *Principles of Synthetic Intelligence PSI: An Architecture of Motivated Cognition*. Oxford University Press, Inc., 2009.

[10] J. Bach. A motivational system for cognitive ai. In J. Schmidhuber, K. R. Thórisson, and M. Looks, editors, *Artificial General Intelligence*, volume 6830 of *Lecture Notes in Computer Science*, pages 232–242. Springer Berlin Heidelberg, 2011.

[11] J. Bach. Micropsi 2: The next generation of the micropsi framework. In *Proceedings of the 5th International Conference on Artificial General Intelligence*, AGI'12, pages 11–20. Springer-Verlag, 2012.

[12] J. Bates. The role of emotion in believable agents. *Commun. ACM*, 37(7):122–125, 1994.

[13] P. Baxter, J. de Greeff, and T. Belpaeme. Cognitive Architecture for Human-Robot Interaction: Towards Behavioural Alignment. *Journal of Biologically Inspired Cognitive Architectures*, 6:30–39, 2013.

[14] R. Beck. *Motivation: Theories and Principles*. Pearson, 2003.

[15] M. Beheshtifar and F. Roasaei. Emotional intelligence: a review and appraisal. *The Journal of Public Management*, 28(2):197–203, 2012.

[16] K. C. Berridge. Motivation concepts in behavioral neuroscience. *Physiology & Behavior*, 81(2):179–209, 2004.

[17] L. M. Botelho and H. Coelho. Machinery for artificial emotions. *Cybernetics and Systems*, 32(5):465–506, 2001.

[18] T. R. Bradberry and L. D. Su. Ability-versus skill-based assessment of emotional intelligence. *Psicothema*, 18(Suppl):59–66, 2006.

[19] 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.

[20] C. Breazeal. A motivational system for regulating human-robot interaction. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference, AAAI 98, IAAI 98, July 26-30, 1998, Madison, Wisconsin, USA.*, pages 54–61, 1998.

[21] I. Bretherton, J. Fritz, C. Zahn-Waxler, and D. Ridgeway. Learning to talk about emotions: A functionalist perspective. *Child Development*, 57(3):529–548, 1986.

[22] N. Brody. *Human motivation: commentary on goal-directed action.* Academic Press New York, 1983.

[23] J. Broekens, W. A. Kosters, and F. J. Verbeek. On affect and self-adaptation: Potential benefits of valence-controlled action-selection. *In: J. Mira and J.R. lvarez (Eds.), IWINAC 2007, Part I, LNCS*, pages 357–366, 2007.

[24] C. Burghart, RalfMikut, 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.

[25] Z. Cai, B. Goertzel, C. Zhou, Y. Zhang, M. Jiang, and G. Yu. Dynamics of a computational affective model inspired by dörners PSI theory. *Cognitive Systems Research*, 17–18:63–80, 2012.

[26] K. Carley. On the persistence of beliefs. working paper. *Department of Social and Decision Sciences, Carnegie Mellon University*, 1990.

[27] D. Cañamero and W. V. de Velde. Socially emotional: Using emotions to ground social interaction. In *In Working Notes of the AAAI'97 Fall Symposium on Socially Intelligent Agents*, 1997.

[28] J. charles Pomerol. Artificial intelligence and human decision making. *European Journal of Operational Research*, 99:3–25, 1997.

[29] H. Q. Chong, A. H. Tan, and G. W. Ng. Integrated cognitive architectures: A survey. *Artificial Intelligence Review*, 28(2):103–130, 2007.

[30] 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.

[31] 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.

[32] G. L. Clore and J. R. Huntsinger. How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, 11(9):393–399, 2007.

[33] G. L. Clore and J. Palmer. Affective guidance of intelligent agents: How emotion controls cognition. *Cognitive System Research*, 10(1):21–30, 2009.

[34] J. A. Coan and J. J. Allen. *Handbook of Emotion Elicitation and Assessment*. Oxford University Press, USA, 2007.

[35] P. Cohen and H. J. Levesque. *Teamwork*. SRI International, 1991.

[36] T. Cole and B. Teboul. Non-zero-sum collaboration, reciprocity, and the preference for similarity: Developing an adaptive model of close relational functioning. *Personal Relationships*, 11(2):135–160, 2004.

[37] R. Cooke and P. Sheeran. Moderation of cognition-intention and cognition-behaviour relations: A meta-analysis of properties of variables from the theory of planned behaviour. *British Journal of Social Psychology*, 43:159–186, 2004.

[38] R. Cowie, N. Sussman, and A. Ben-Ze'ev. Emotion: Concepts and definitions. In *Emotion-Oriented Systems*, Cognitive Technologies, pages 9–30. Springer-Verlag Berlin Heidelberg, London, New York, March 2011.

[39] C. L. Dancy. ACT-Rϕ: A cognitive architecture with physiology and affect. *Biologically Inspired Cognitive Architectures*, 6:40–45, 2013.

[40] C. Darwin. *The expression of the emotions in man and animals*. New York; D. Appleton and Co., 1916.

[41] K. Dautenhahn. Getting to know each other - artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, 16:333–356, 1995.

[42] J. S. de Freitas and J. Queiroz. Artificial emotions: Are we ready for them? In F. A. e Costa, L. M. Rocha, E. Costa, I. Harvey, and A. Coutinho, editors, *ECAL*, volume 4648, pages 223–232. Springer, 2007.

[43] 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.

[44] E. de Sevin and D. Thalmann. A motivational model of action selection for virtual humans. In *Proceedings of the Computer Graphics International 2005*, CGI '05, pages 213–220, Washington, DC, USA, 2005. IEEE Computer Society.

[45] B. M. DePaulo. Nonverbal behavior and self-presentation. *Psychological Bulletin*, 111(2):203–243, 1992.

[46] S. D'Mello and S. Franklin. Computational modeling/cognitive robotics complements functional modeling/experimental psychology. *New Ideas in Psychology*, 29(3):217–227, 2011.

[47] P. F. Dominey and F. Warneken. The basis of shared intentions in human and robot cognition. *New Ideas in Psychology*, 29(3):260–274, 2011.

[48] H. L. Dreyfus. *What Computers Still Can'T Do: A Critique of Artificial Reason*. MIT Press, 1992.

[49] E. Durfee and V. Lesser. Incremental planning to control a blackboard-based problem solver. *Proceedings of the Fifth National Conference on Artificial Intelligence*, 1:58–64, August 1986.

[50] P. Ekman. An argument for basic emotions. *Cognition & Emotion*, 6(3-4):169–200, 1992.

[51] U. X. Eligio, S. E. Ainsworth, and C. K. Crook. Emotion understanding and performance during computer-supported collaboration. *Computers in Human Behavavior*, 28(6):2046–2054, 2012.

[52] L. D. Erman, F. Hayes-Roth, V. R. Lesser, and D. R. Reddy. The Hearsay-II speech-understanding system: Integrating knowledge to resolve uncertainty. *ACM Computing Surveys*, 12(2):213–253, 1980.

[53] 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.

[54] L. Faucher and C. Tappolet. Fear and the focus of attention. *Consciousness and Emotion*, 3(2):105–144, 2002.

[55] S. Folkman and J. T. Moskowitz. Coping: Pitfalls and promise. *Annual Review of Psychology*, 55:745–774, 2004.

[56] B. L. Fredrickson and C. Branigan. Positive emotions broaden the scope of attention and thoughtaction repertoires. *Cognition and Emotion*, 19(3):313–332, 2005.

[57] N. H. Frijda. *The Emotions*. Cambridge University Press, 1986.

[58] S. R. Fussell and R. M. Krauss. Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of personality and social psychology*, 62(1):378–391, 1992.

[59] S. C. Gadanho and J. Hallam. Robot learning driven by emotions. *AI Magazine*, 9(1):42–64, 2001.

[60] 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.

[61] P. J. Gmytrasiewicz and C. L. Lisetti. Emotions and personality in agent design and modeling. In J.-J. C. Meyer and M. Tambe, editors, *ATAL*, volume 2333 of *Lecture Notes in Computer Science*, pages 21–31. Springer, 2001.

[62] E. Goffman. *The Presentation of Self in Everyday Life*. Anchor, 1959.

[63] S. Graham and B. Weiner. *Theories and Principles of Motivation*. Prentice Hall, 1996.

[64] J. Grant, S. Kraus, and D. Perlis. A logic-based model of intention formation and action for multi-agent subcontracting. *Artificial Intelligence*, 163(2):163–201, 2005.

[65] J. Gratch and S. C.Marsella. A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4):269–306, 2004.

[66] J. Gratch and S. Marsella. Evaluating a computational model of emotion. *Journal of Autonomous Agents and Multi-Agent Systems. Special Issue on the Best of AAMAS 2004*, 11(1):23–43, July 2005.

[67] 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.

[68] J. J. Gross, G. Sheppes, and H. L. Urry. Emotion generation and emotion regulation: A distinction we should make (carefully). *Cognition and Emotion*, 25(5):765–781, 2011.

[69] S. Grossberg and W. E. Gutowski. Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review*, 94(3):300–318, 1987.

[70] B. Grosz and S. Kraus. The evolution of shared plans. In *Foundations and Theories of Rational Agency*, pages 227–262, 1998.

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

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

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

[74] 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.

[75] N. Guiraud, D. Longin, E. Lorini, S. Pesty, and J. Rivière. The face of emotions: A logical formalization of expressive speech acts. In *The 10th Interna-*

tional Conference on Autonomous Agents and Multiagent Systems - Volume 3, pages 1031–1038. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[76] S. Hareli and U. Hess. What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition & Emotion*, 24(1):128–140, 2009.

[77] B. Hayes-Roth, F. Hayes-Roth, S. Rosenschein, and S. Cammarata. Modeling planning as an incremental, opportunistic process. In *Proceedings of the 6th International Joint Conference on Artificial Intelligence - Volume 1*, pages 375–383. Morgan Kaufmann Publishers Inc., 1979.

[78] C. Hazan and P. Shaver. Romantic love conceptualized as an attachment process. *Personality & Social Psychology*, 52(3):511–524, 1987.

[79] U. Hess and P. Thibault. Darwin and emotion expression. *American Psychologist*, 64(2):120–128, 2009.

[80] L. M. Hiatt and J. G. Trafton. A cognitive model of theory of mind. In *In Proceedings of the International Conference on Cognitive Modeling*, pages 91–96, 2010.

[81] E. Hudlicka. To feel or not to feel: The role of affect in human-computer interaction. *International Journal of Human-Computer Studies*, 59(1-2):1–32, 2003.

[82] E. Hudlicka. What are we modeling when we model emotion? In *AAAI Spring Symposium: Emotion, Personality, and Social Behavior*, pages 52–59. AAAI, 2008.

[83] C. L. John Robert Anderson. *The Atomic Components of Thought*. Lawrence Erlbaum Associates, 1998.

[84] P. N. Johnson-laird and K. Oatley. Basic emotions, rationality, and folk theory. *Cognition & Emotion*, 6(3):201–223, 1992.

[85] D. P. Keating. A search for social intelligence. *Journal of Educational psychology*, 70:218–223, 1978.

[86] T. D. Kelley and D. N. Cassenti. Theoretical explorations of cognitive robotics using developmental psychology. *New Ideas in Psychology*, 29(3):228–234, 2011.

[87] D. Keltner and J. Haidt. Social functions of emotions at four levels of analysis. *Cognition and Emotion*, 13(5):505–521, 1999.

[88] 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.

[89] R. Kirby, J. Forlizzi, and R. Simmons. Affective social robots. *Robotics and Autonomous Systems*, 58:322, 2010.

[90] G. A. V. Kleef. How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science*, 18(3):184–188, 2009.

[91] A. C. Kruger. Peer collaboration: conflict, cooperation, or both? *Social Development*, 2(3):165–182, 1993.

[92] J. Laird. *The Soar Cognitive Architecture*. MIT Press, 2012.

[93] D. Laming. *Understanding Human Motivation: What Makes People Tick*. Wiley, 2003.

[94] P. Langley, J. E. Laird, and S. Rogers. Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, 10(2):141–160, 2009.

[95] R. S. Lazarus. *Emotion and Adaptation*. OXFORD University Press, 1991.

[96] R. S. Lazarus and S. Folkman. *Stress, Appraisal, and Coping*. Springer Publishing Company, 1984.

[97] V. R. Lesser and D. Corkill. The distributed vehicle monitoring testbed: A tool for investigating distributed problem solving networks. *AI Magazine*, 4(3):15–33, 1983.

[98] R. W. Levenson and J. M. Gottman. Martial interaction: Physiological linkage and affective exchnage. *Personality and Social Psychology*, 45(3):587–597, 1983.

[99] 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.

[100] J. Lin, M. Spraragen, and M. Zyda. Computational models of emotion and cognition. *Advances in Cognitive Systems*, 2:59–76, 2012.

[101] 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.

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

[103] 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.

[104] A. Luneski and R. K. Moore. Affective computing and collaborative networks: Towards emotion-aware interaction. In *Pervasive Collaborative Networks*, volume 283, pages 315–322. Springer US, 2008.

[105] J. N. Maksym, A. J. Bonner, C. A. Dent, and G. L. Hemphill. Machine analysis of acoustical signals. *Pattern Recognition*, 16(6):615–625, 1983.

[106] M. Manso and J. Moffat. Defining and measuring cognitive-entropy and cognitive self-synchronization. In *ICCRTS*, 2011.

[107] 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.

[108] R. P. MarinierIII, J. E. Laird, and R. L. Lewis. A computational unification of cognitive behavior and emotion. *Cognitive System Research*, 10(1):48–69, March 2009.

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

[110] S. C. Marsella, J. Gratch, and P. Petta. Computational models of emotion. In *Blueprint for Affective Computing (Series in Affective Science)*. Oxford University Press, Oxford, 2010.

[111] J. Martínez-Miranda and A. Aldea. Emotions in human and artificial intelligence. *Computers in Human Behavior*, 21(2):323–341, 2005.

[112] 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.

[113] D. C. McClelland. *Human Motivation*. Scott, Foresman, 1985.

[114] K. E. Merrick and K. Shafi. Achievement, affiliation, and power: Motive profiles for artificial agents. *Adaptive Behavior*, 19(1):40–62, 2011.

[115] M. Minsky. *The Society of Mind*. Simon & Schuster, Inc., 1986.

[116] M. Mirolli and D. Parisi. Towards a vygotskyan cognitive robotics: The role of language as a cognitive tool. *New Ideas in Psychology*, 29(3):298–311, 2011.

[117] L. P. Morency. Computational study of human communication dynamic. In *Proceedings of the 2011 Joint ACM Workshop on Human Gesture and Behavior Understanding*, pages 13–18. ACM, 2011.

[118] A. F. Morse, C. Herrera, R. Clowes, A. Montebelli, and T. Ziemke. The role of robotic modelling in cognitive science. *New ideas in psychology*, 29(3):312–324, 2011.

[119] H. A. Murray. *Explorations of Personality*. Oxford University Press, New York, 1938.

[120] L. C. namero. Emotion understanding from the perspective of autonomous robots research. *Neural Networks*, 18(4):445–455, 2005.

[121] 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.

[122] S. Nirenburg, V. Lesser, and E. H. Nyberg. Controlling a language generation planner. In *IJCAI-89*, pages 1524–1530, 1989.

[123] A. Ortony, G. L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[124] B. Parkinson. Emotions are social. *British Journal of Psychology*, 87(4):663–683, 1996.

[125] B. Parkinson. Do facial movements express emotions or communicate motives? *Personality and Social Psychology Review*, 9(4):278–311, 2005.

[126] B. Parkinson. What holds emotions together? meaning and response coordination. *Cognitive System Research*, 10(1):31–47, 2009.

[127] W. G. Parrott. Appraisal, emotion words, and the social nature of self-conscious emotions. *Psychological Inquiry*, 15(2):136–138, 2004.

[128] M. P. Paulus and A. J. Yu. Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in cognitive sciences*, 16(9):476–483, 2012.

[129] R. Pfeifer, H. Ishiguro, Y. Anzai, and N. Miyake. Robotics and emotion. In N. Miyake, D. Peebles, and R. P. Cooper, editors, *Building Bridges Across Cognitive Sciences Around the World: Proceedings of the 34th Annual Meeting of the Cognitive Science Society; Sapporo, Japan, August 1-4, 2012*, pages 30–31. Cognitive Science Society, 2012.

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

[131] M. E. Pollack. A model of plan inference that distinguishes between the beliefs of actors and observers. In *Proceedings of the 24th Annual Meeting on Association for Computational Linguistics*, pages 207–214. Association for Computational Linguistics, 1986.

[132] M. E. Pollack. Plans as complex mental attitudes. In *Intentions in Communication*, pages 77–103. MIT Press, 1990.

[133] D. V. Pynadath, M. Si, and S. C. Marsella. Modeling theory of mind and cognitive appraisal with decision-theoretic agents. In *Appraisal*, pages 1–30. 2011.

[134] A. K. Ray, G. Leng, T. McGinnity, S. Coleman, and L. Maguire. Development of a self sustaining cognitive architecture. *Biologically Inspired Cognitive Architectures*, 6:96–108, 2013.

[135] C. Rich. Building task-based user interfaces with ANSI/CEA-2018. *IEEE Computer*, 42(8):20–27, July 2009.

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

[137] T. H. Rumbell, J. A. Barnden, S. L. Denham, and T. Wennekers. Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Autonomous Agents and Multi-Agent Systems*, 25(1):1–45, 2012.

[138] R. M. Ryan and E. L. Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1):54–67, 2000.

[139] A. V. Samsonovich. Emotional biologically inspired cognitive architecture. *Biologically Inspired Cognitive Architectures*, 6:109–125, 2013.

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

[141] B. Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, 12(1):13–24, 2002.

[142] 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.

[143] 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.

[144] K. R. Scherer. Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99(2):143–165, 1986.

[145] K. R. Scherer. On the sequential nature of appraisal processes: Indirect evidence from a recognition task. *Cognition & Emotion*, 13(6):763–793, 1999.

[146] K. R. Scherer and H. Eligring. Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion*, 7(1):113–130, 2007.

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

[148] M. Scheutz. Affective action selection and behavior arbitration for autonomous robots. In H. Arabnia, editor, *Proceedings of the 2002 International Conference on Artificial Intelligence*, pages 334–340. CSREA Press, 2002.

[149] M. Scheutz. Useful roles of emotions in artificial agents: A case study from artificial life. In *Proceedings of AAAI*, pages 31–40. AAAI Press, 2004.

[150] M. Scheutz, A. Sloman, and B. Logan. Emotional states and realistic agent behaviour. In *SOCIETY FOR COMPUTER SIMULATION*, pages 81–88, 2000.

[151] D. A. Schon. *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*. Wiley, 1987.

[152] O. Schultheiss and J. Brunstein. *Implicit Motives*. Oxford University Press, 2010.

[153] M. Sellers. Toward a comprehensive theory of emotion for biological and artificial agents. *Biologically Inspired Cognitive Architectures*, 4:3–26, 2013.

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

[155] S. Shott. Emotion and social life: A symbolic interactionist analysis. *The American Journal of Sociology*, 84(6):1317–1334, 1979.

[156] M. Si, S. C. Marsella, and D. V. Pynadath. Modeling appraisal in theory of mind reasoning. *Journal of Autonomous Agents and Multi-Agent Systems*, 20(1):14–31, 2010.

[157] M. Si, S. C. Marsella, and D. V. Pynadath. Modeling appraisal in theory of mind reasoning. *Autonomous Agents and Multi-Agent Systems*, 20(1):14–31, 2010.

[158] C. Sidner. An artificial discourse language for collaborative negotiation. In *In Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 814–819. MIT Press, 1994.

[159] H. A. Simon. Motivational and emotional controls of cognition. *Psychology Review*, 74(1):29–39, 1967.

[160] A. Sloman. Motives, mechanisms, and emotions. *Cognition & Emotion*, 1(3):217–233, 1987.

[161] A. Sloman. Beyond shallow models of emotion. *Cognitive Processing: International Quarterly of Cognitive Science*, 2(1):177–198, 2001.

[162] A. Sloman and M. Croucher. Why robots will have emotions. In *In Proceeding 7th International Joint Conference on AI*, pages 197–202, 1981.

[163] 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.

[164] R. D. Sousa. *The Rationality of Emotion*. MIT Press, 1990.

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

[166] P. A. Thoits. The sociology of emotion. *Anuual Review of Sociology*, 15:317–342, 1989.

[167] K. R. Thórisson and H. P. Helgason. Cognitive architectures and autonomy: A comparative review. *Journal of Artificial General Intelligence*, 3(2):1–30, 2012.

[168] L. Z. Tiedens and C. W. Leach. *The Social Life of Emotions (Studies in Emotion and Social Interaction)*. Cambridge University Press, 2004.

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

[170] 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.

[171] M. N. Wexler. Emotional intelligence: a review and appraisal. *The Journal of Public Management*, 30(2):1–8, 2000.

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

[173] R. P. Worden. Primate social intelligence. *Cognitive Science*, 20(4):579–616, 1996.

[174] I. Wright and A. Sloman. *MINDER1: An Implementation of a Protoemotional Agent Architecture*. Cognitive science research papers. University of Birmingham, Cognitive Science Research Centre, 1997.

[175] R. B. Zajonc. Social facilitation. *Science*, 149(3681):269–274, 1965.

[176] J. Zhu and P. Thagard. Emotion and action. *Journal of Philosophical Psychology*, 15(1):19–36, 2002.

[177] E. L. Zurbriggen and T. S. Sturman. Linking motives and emotions: A test of mcclelland's hypotheses. *Personality and Social Psychology*, 28(4):521–535, 2002.