RESEARCH STATEMENT

Mahni Shayganfar (mshayganfar@wpi.edu)

Affect-Driven Cognitive Human Performance Evaluation & Adaptation

My current research spans the areas of computational collaboration theories, affective computing, human-robot collaboration, and cognitive robotics. Common threads in my research are the development of Affective Motivational Collaboration Theory, design of a domain-independent architecture, as well as the design of the framework which uses this architecture. The goal is to provide a collaborative behavior for robots or virtual agents based on this architecture. I have built on existing computational collaboration theories, i.e., SharedPlans theory, and computational models of emotions, i.e., cognitive appraisal theory to develop my own theory. Broadly speaking, my research belongs to the area of human-robot collaboration and its underlying processes, a growing field which has influence on different leading industries such as autonomous vehicles, space exploration, manufacturing, and any industry situation in which human-robot teamwork is required.

Improving the human-robot teamwork performance not only depends on the robot's computational and physical capabilities, but it also depends on the human's performance and the ability of the robot to maintain this performance according to human's cognitive state. It is crucial to understand that the human's performance dynamically changes with respect to the internal and external events. The internal events comprise changes in mental state, e.g., violation of a strong belief, occurance of ambivalent intentions, and the blockage of an urgent motive. The external events can also influence one's performance, e.g., abrupt changes of a task's timing, failure of high priority tasks, and delay in task duration. A collaborative robot requires to evaluate the existing data with respect to the human's cognitive model, adopt corresponding and appropriate behaviors, and adapt its own actions to maintain human's performance in task execution.

Background and Motivation

The construction of robots that are intelligent, collaborative problem-solving partners is important in robotics and applications of Artificial Intelligence. It has always been important for us to make robots better at helping us to do whatever they are designed for. To build collaborative robots, we need to identify the capabilities that must be added to them so that they can work with us or other agents. Collaboration must be designed into systems from the start; it cannot be patched on [9]. Collaboration is a type of coordinated activity in which the participants work together performing a task or carrying out the activities needed to satisfy a shared goal [12].

Collaboration Theories

Collaboration involves several key properties both a structural and functional levels. For instance, most collaborative situations involve participants who have different beliefs and capabilities; most of the time collaborators only have partial knowledge of the process of accomplishing the collaborative activities; collaborative plans are more than the sum of individual plans; collaborators are required to maintain mutual beliefs about their shared goal throughout the collaboration; they need to be able to communicate with others effectively; they need to commit to the group activities and to their role in it; collaborators need to commit to the success of others; they need to reconcile between commitments to the existing collaboration and their other activities; and they need to interpret others' actions and utterances in the collaboration context [10]. These collaboration properties are captured by the existing computational collaboration theories.

As I mentioned, to be collaborative, partners, e.g., a robot and a human, need to meet the specifications stipulated by collaboration theories. The prominent collaboration theories are mostly based on plans and joint intentions [3, 13, 17], and they were derived from the BDI paradigm developed by Bratman [1] which is fundamentally reliant on folk psychology [21]. The two theories, Joint Intentions [3] and SharedPlans [13], have been extensively used to examine and describe teamwork and collaboration. The SharedPlans model of collaborative action [11, 12, 13], aims to provide the theoretical foundations needed for building collaborative robots/agents [9]. SharedPlans is a general theory of collaborative planning that requires no notion of joint intentions, accommodates multi-level action decomposition hierarchies and allows the process of expanding and elaborating partial plans into full plans. SharedPlans theory explains how a group of agents can incrementally form and execute a shared plan that then guides and coordinates their activity towards the accomplishment of a shared goal.

The Joint Intentions theory of Cohen and Levesque [3, 4, 5, 6, 16] represents one of the first attempts to establish a formal theory of collaboration, and due to its clarity and expression, is a widely used teamwork theory. The basic idea of Joint Intentions theory is based on individual and joint intentions (as well as commitments) to act as a team member. Their notion of joint intention is viewed not only as a persistent commitment of the team to a shared goal, but also implies a commitment on part of all its members to a mutual belief about the state of the goal. In other words, Joint Intentions theory describes how a team of agents can jointly act together by sharing mental states about their actions while an intention is viewed as a commitment to perform an action [4].

There are also hybrid collaboration theories which borrow collaboration concepts from both SharedPlans and Joint Intentions theories. For instance, STEAM [24] (simply, a Shell TEAMwork) is a hybrid collaboration theory which its operationalization in complex, real-world domains is the key in its development to addressing important teamwork issues. STEAM is founded on the Joint Intentions theory and it uses joint intentions as the basic building block of teamwork while it is informed by key concepts from SharedPlans theory.

I believe the SharedPlans and Joint Intentions collaboration theories are the most well-defined and well-established theories in computer science. I found SharedPlans theory more convincing than the other major and subordinate approaches, with respect to its inclusive explanation of the collaboration structure and its association to discourse analysis which directly improves the communicative aspects of a collaboration theory. I also understand the value of Joint Intentions theory due to its clarity and closeness to the foundations of collaboration concepts. These specifications of the Joint Intentions theory can make it applicable in multi-agent system designs and human-robot collaboration. I also consider hybrid approaches valuable, such as STEAM, if they clearly understand drawbacks with existing theories and successfully achieve better collaborative agents by infusing different concepts from different theories.

Cognitive Appraisal Theory

The term affective computing encapsulates a new approach in Artificial Intelligence, to build computers that show human affection [20]. Studies show that the decision making of humans is not always logical [8], 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 [7]. Emotions impact fundamental parts of cognition including perception, memory, attention and reasoning [2]. This impact is caused by the information emotions carry about the environment and event values. If we want robots and virtual agents to be more believable and efficient partners for humans, we must consider the personal and social functionalities and characteristics of emotions; this will enable our robots to coexist with humans, who are emotional beings.

There are different types of computational theories of emotion. These theories differ in the type of relationships between their components and whether a particular component plays a crucial role in an individual emotion. 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 [23]. The appraisal procedure begins with the evaluation of the environment according to the internalized goals and is based on systematic assessment of several elements [22]. The outcome of this process triggers the appropriate emotions.

In my Ph.D thesis, I lay out a computational framework for the theory I have developed based on SharedPlans and Cognitive Appraisal theories. A great deal of my work has benefited from the integration of these well-established theories and their underlying structure.

Limitations of Existing Theories

There are several well-developed cognitive architectures, e.g., Soar [15] and ACT-R [14], each with different approaches to defining the basic cognitive and perceptual operations. There have also been efforts to integrate affect into these architectures [18]. In general, however, these cognitive architectures do not focus on processes to specifically produce affect-regulated goal-driven collaborative behaviors. At the same time, existing collaboration theories, e.g. SharedPlans theory [13], Joint Intentions [3], and STEAM [24], focus on describing the structure of a collaboration in terms of fundamental mental states, e.g., mutual beliefs or intentions. Although all the existing collaboration theories are well-defined and properly introduce collaboration concepts, they mostly explain the structure of a collaboration and they lack the underlying domain-independent processes with which collaborative procedures could be defined more systematically and effectively in different applications. Therefore, it is crucial to investigate the cognitive processes involved in a collaboration in the context of a cognitive architecture to be able to describe the underlying processes, their relationships, and their influences on each other.

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.

Current Research

In my Ph.D. thesis, I have developed Affective Motivational Collaboration Theory and the associated computational model that will enhance the performance and effectiveness of collaboration between robots and humans. This theory explains the functions of emotions in a dyadic collaboration and shows how affective mechanisms can coordinate social interactions by anticipating other's emotions, beliefs and intentions. This theory also specifies the influence of the underlying collaboration processes on appraisals. Affective Motivational Collaboration Theory elucidates 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. I am validating my theory using my computational framework in the context of a human-robot collaboration.

I have investigated how affect-driven mechanisms in my Affective Motivational Collaboration Theory can be involved to successfully maintain a dyadic collaboration. My research is based on two hypothetical collaboration examples of *task delegation* and *agreeing on a shared goal* for both success and failure of each scenario because of the emotion-awareness and emotion-ignorance behavior of the robot. I have implemented the rules associated with these examples using JESS (Java Expert System Shell) which is a rule engine for the Java platform. In my current implementation I have categorized the rules in different modules associated with the mechanisms and the underlying processes in Affective Motivational Collaboration Theory. The outcome of this part of my work is currently under review as a journal publication. I am implementing algorithms for each individual mechanism to be able to automatically generate the required facts within each mechanism to fire the existing rules. Ultimately, I am going to provide a platform which operates based on the collaboration structure, and employs affect-driven processes such as the appraisal processes in [19] to enable a robot to obtain and demonstrate collaborative behaviors.

As another crucial part of my research, I have developed domain-independent algorithms for the appraisal of a collaborative environment. These algorithms compute the value of the four appraisal variables: *relevance, desirability, expectedness* and *controllability.* These algorithms use a generalized collaboration structure to compute their outcome in a dyadic collaboration. The performance of individual algorithms will be verified based on the analysis of the answers human subjects provide to our questionnaire. All of the questions have been prepared based on the same task models provided to our framework. I will finalize my Ph.D. thesis by evaluating the overall architecture using a robot simulator to successfully and efficiently (in terms of time and number of errors) collaborate with human subjects.

Future Research Directions

Collaborative robots are becoming an integral part of humans' environment to accomplish their industrial and household tasks. In these environments humans are involved in robots' operations and decision making processes. This involvement of humans influences the efficiency of robots' interaction and performance, and makes them dependent on the humans' cognitive abilities and mental states. Similarly, humans' performance can be maintained using robot's adaptive behavior. The evaluative nature of affective processes can assist a robot to perform adaptive behaviors with respect to the current goal. Goal management, alarm mechanism and action selection are some examples of the functions that a system with affect-driven processes offers.

Performance evaluation of a human requires perception of the human's physical and mental states. This perception can be performed based on available bio-information, as well as verbal (e.g., utterances) and non-verbal (e.g., facial expressions) behaviors. However, a robot requires cognitive mechanisms to evaluate and adapt its own behavior to change the human's performance on the given task and environment. My Ph.D. thesis provides fundamental affect-driven mechanisms and defines their relationship within a cognitive architecture. These mechanisms can be involved from initial evaluation of the input data to higher level processes such as action selection or alarm mechanism. All these functions can improve a robot's awareness and self-synchronization capabilities during the interaction. For instance, if the appraisal of the input data reveals controllability (an appraisal variable) of an event, the robot can choose an appropriate action to make changes to the status of the interaction. Ultimately, these changes can improve human's performance in a task execution. I would like to apply my knowledge I attained in my research to develop new algorithms to evaluate performance of a human and adapt the robot's behavior respectively.

References

[1] Michael E. Bratman. *Intention, Plans, and Practical Reason.* Cambridge, Mass.: Harvard University Press, 1987.

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

[3] Philip Cohen and Hector J. Levesque. *Teamwork.* SRI International, 1991.

[4] Philip R. Cohen and Hector J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261, 1990.

[5] Philip R. Cohen and Hector J. Levesque. Persistence, intention, and commitment. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*, pages 33–69. MIT Press, Cambridge, MA, 1990.

[6] Philip R. Cohen, Jerry Morgan, and Martha E. Pollack. *Intentions in Communication.* A Bradford Book, 1990.

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

[8] 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.

[9] Barbara J. Grosz. AAAI-94 presidential address: Collaborative systems. *AI Magazine*, 17(2):67–85, 1996.

[10] Barbara J. Grosz. Beyond mice and menus. *Proceedings of the American Philosophical Society*, 149(4):523–543, 2005.

[11] Barbara J. Grosz, Luke Hunsberger, and Sarit Kraus. Planning and acting together. *AI Magazine*, 20(4):23–34, 1999.

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

[13] Barbara J. Grosz and Candace 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.

[14] Christian Lebiere John Robert Anderson. *The Atomic Components of Thought.* Lawrence Erlbaum Associates, 1998.

[15] John Laird. *The Soar Cognitive Architecture.* MIT Press, 2012.

[16] Hector J. Levesque, Philip R. Cohen, and Jos H. T. Nunes. On acting together. In *AAAI*, pages 94–99. AAAI Press / The MIT Press, 1990.

[17] 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.

[18] Robert P. Marinier III, John E. Laird, and Richard L. Lewis. A computational unification of cognitive behavior and emotion. *Cognitive System Research*, 10(1):48–69, March 2009.

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

[20] Rosalind W. Picard. *Affective Computing.* The MIT Press, 2000.

[21] Ian Ravenscroft. *Folk Psychology as a Theory.* Stanford Encyclopedia of Philosophy, 2004.

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

[23] Klaus R. Scherer, Angela Schorr, and Tom Johnstone. *Appraisal Processes in Emotion: Theory, Methods, Research.* Oxford University Press, 2001.

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