Computational Theories of Collaboration
Ph.D. Comprehensive Exam

Mohammad Shayganfar - mshayganfar@wpi.edu
May, 26 2015

1 Introduction to Collaboration Theories

The construction of computer systems and robots that are intelligent, collaborative problem-solving partners is important in Artificial Intelligence (AI) and its applications. It has always been important for us to make computer systems better at helping us to do whatever they are designed for. To build collaborative systems, we need to identify the capabilities that must be added to individual agents so that they can work with us or other agents. As Grosz says, collaboration must be designed into systems from the start; it cannot be patched on [23].

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 [27]. Collaboration involves several key properties both in 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 through out 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 [24]. These collaboration properties are captured by the existing computational collaboration theories.

As we mentioned, to be collaborative, partners, e.g., a robot and a human, need to meet the specifications stipulated by collaboration theories. These theories argue for an essential distinction between a collaboration and a simple interaction or even a coordination in terms of commitments [22, 45]. This document briefly provides descriptions of major computational collaboration theories, their similarities and differences, and their application in AI and robotics. It primarily focuses on Joint Intention, SharedPlans and hybrid theories of collaboration. In this document, we do not present the theroies in formal language, but simply describe their features in general terms.

2 Computational Theories of Collaboration

The prominent collaboration theories are mostly based on plans and joint intentions [15] [29] [44], and they were derived from the BDI paradigm developed by Bratman [4] which is fundamentally reliant on folk psychology [57]. The two theories, Joint Intentions [15] and SharedPlans [29], have been extensively used to examine and describe teamwork and collaboration.

The SharedPlans theory is based on the theories of Bratman and Pollack [7, 53, 54], who outline a mental-state view of plans in which having a plan is not just knowing how to do an action, but also having the intention to do the actions entailed. Bratman's views of intention goes back to the philosophical views of Anscombe [2] and *Castañeda* [10] about intention. Also, as Grosz and Sidner mention in [29] the natural segmentation of discourse reflects intentional behaviors in each segment. These intentions are designated as Discourse Segment Purposes (DSPs) which are the basic reasons for engaging in different segments of discourse. DSPs are a natural extension of Gricean intentions at the utterance level [51].

Cohen and Levesque also mention that in Joint Intentions theory their view of intention is primarily future-directed [16] which makes their view similar to Bratman's theory of intention [5], contra Searle [64].

Commitment – One of the most important concepts of teamwork and collaboration is the concept of commitment. Collaboration theories are required to meet the notion of commitment, otherwise the participants are just doing some coordinated works. Since the prominent computational collaboration theories, reviewed in this paper, are based on Bratman's view of intention, we briefly provide his view of commitment here before describing these theories. Bratman defines certain prerequisites for an activity to be considered shared and cooperative [6]. He stresses the importance of:

a) **Mutual commitment to joint activity** – which can be achieved by agreement on the joint activity, and prevention of abandoning the activity without involving teammates;

b) **Mutual suppport** – which can be achieved by team members if they actively try to help teammate activity;

c) **Mutual responsiveness** – which means team members should take over tasks from teammates if necessary.

In the following sections, we are also going to see how each collaboration theory addresses the notion of commitment.

2.1 SharedPlans Theory

The SharedPlans model of collaborative action, presented by Grosz and Sidner [26, 27, 29], aims to provide the theoretical foundations needed for building collaborative robots/agents [23]. SharedPlans is a general theory of collaborative planning that requires no notion of joint intentions (see Section 2.2), accommodates multi-level action decomposition hierarchies and allows the process of expanding and elaborating partial plans into full plans (see Section 2.1.3). 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. SharedPlans is rooted in the observation that collaborative plans are not simply a collection of individual plans, but rather a tight interleaving of mutual beliefs and intentions of different team members. In [27] Grosz and Kraus use first-order logic to present the formalization of SharedPlans.

Grosz and Sidner in [29] present a model of plans to account for how agents with partial knowledge collaborate in the construction of a domain plan. They are interested in the type of plans that underlie discourse in which the agents are collaborating in order to achieve a shared goal. They propose that agents are building a shared plan (see Section 2.1.2) in which participants have a collection of beliefs and intentions about the actions in the plan. Agents have a library of how to do their actions, i.e. recipes (see Section 2.1.1). These recipes might be partially specified as to how an action is executed, or contributes to a goal (see Section 2.1.3). Then, each agent communicates their beliefs and intentions by making utterances about what actions they can contribute to the shared plan. This communication leads to the construction of a shared plan, as well as termination of the collaboration with each agent mutually believing that there exists one agent who is going to execute an action in the plan, and the fact that that agent has intention to perform the action, and that each action in the plan contributes to the goal [29] [46].

Later in Section 2.1.2, we are going to see that to successfully complete a plan the collaborators must mutually believe that they have a common goal and have agreed on a sequence of actions for achieving that goal. They should believe that they are both capable of performing their own actions and intend to perform those actions while they are committed to the success of their plans.

2.1.1 Recipes

The SharedPlans theory differentiates between knowing how to accomplish a goal (a recipe) and having a plan, which includes intentions. The SharedPlans definition of mutual beliefs states that when agents have a shared plan for doing some action, they must hold mutual beliefs about the way in which they should perform that action [27, 29]. Following Pollack [54], the term recipe refers to what collaborators know when they know a way of doing an action. Recipes are specified at a particular level of detail. Although the agents need to have mutual beliefs about actions specified in the recipe, they do not need to have mutual beliefs about all levels of performing actions. Therefore, having mutual beliefs of the recipe means that the collaborators hold the same beliefs about the way in which an action should be accomplished. Consequently, the collaborators need to agree on how to execute an action. Recipes are aggregations of action-types and relations among them. Action-types, rather than actions, are the main elements in recipes. Grosz and Sidner in their earlier work [29] have considered only simple recipes in which each recipe consisted of only a single action-type relation [46]. Recipes can be partial, meaning they can expand and be modified over time.

2.1.2 Shared Plans

Grosz and Sidner propose that collaboration must have the following three elements, which also shows the importance of the shared plans:

1. the participants must have commitment to the shared activity;
2. there must be a process for reaching an agreement on a recipe for the group action;
3. there must be commitment to the constituent actions.

Shared plan is an essential concept in the collaboration context. The definition of the shared plan is derived from the definition of plans Pollack introduced in [53, 54] 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 [46]. 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 [27]. Grosz and Kraus in [27], reformulate Pollack's definition of the individual plans [54], and also revise and expand the SharedPlans to address these shortcomings.

Figure 1 shows what we need to add to individual plans in order to have plans for group actions. The top of the figure lists the main components for individual plans. First, an individual agent needs to know the recipe for an action, whereas agents in a group need to have a mutual belief of a recipe for an action (bottom of the figure). In the case of a group plan, having a mutual belief of a recipe, leads the agents to agree on how they are going to execute the action. Then, similar to individual agents that need to have the ability to perform the constituent actions in an individual plan and must have intentions to perform them, the participants in a group activity need to have individual or group plans for each of the constituent actions in the mutually agreed recipe [23, 29].

PLANS FOR COLLABORATIVE ACTION

- To have an individual plan for an act, need
  - knowledge of a recipe
  - ability to perform subacts in the recipe
  - intentions to do the subacts
- To have a group plan for an act, need
  - mutual belief of a recipe
  - individual or group plans for the subacts
  - intentions that group perform act
  - intentions that collaborators succeed

Figure 1: Plans for collaborative action [23].

As shown in Figure 1 (bottom), plans for group actions include two essential constituents that do not have correlates in the individual plan. First, the agents need to have a commitment to the group activity; All the agents need to intend that (see Section 2.1.5) the group will do the action. For instance, a robot and an astronaut need to have intentions that they install solar panels together. Among other things, these intentions will keep them both working on the panels until the panels are installed. Second, the participants need to have some commitment to the other agents to succeed in their own their actions. For instance, the robot must have an intention that the astronaut be able to measure the quality of installation successfully. This intention will prevent the robot from interrupting the astronaut's measurement action or prevent the robot from using the astronaut's measurement tool [23, 29].

2.1.3 Full Vs. Partial Shared Plan

The SharedPlans formalization distinguishes complete plans and partial plans. A shared plan can be either a *Full Shared Plan (FSP)* or a *Partial Shared Plan (PSP)*. An *FSP* is a complete plan in which agents have fully determined how they will perform an action. A *PSP* definition provides a specification of the minimal mental state requirements for collaboration to exist and gives criteria governing the process of completing the plan.

An *FSP* to do α represents a situation where every aspect of a joint activity α is fully determined. This includes mutual belief and agreement in the complete recipe to do α. A recipe is a specication of a set of actions A_i, which constitutes the performance of α when executed under specified constraints. *FSP(*$\boldsymbol{P}$*,* $\Theta, \alpha, T_p, T_\alpha, \boldsymbol{R}_\alpha$*)* denotes a group Θ's plan $\boldsymbol{P}$ at time T_p to do action α at time T_α using recipe $\boldsymbol{R}_\alpha$. In short, *FSP* holds if and only if the following conditions are satisfied:

1. All members of group Θ mutually believe that they intend to do α.
2. All members of group Θ mutually believe that $\boldsymbol{R}_\alpha$ is the recipe for α.
3. For each step A_i in recipe $\boldsymbol{R}_\alpha$:
    - A subgroup Θ_j has an *FSP* for A_i, using recipe $\boldsymbol{R}_{A_i}$.
    - Other members of group Θ believe that there exists a recipe such that subgroup Θ_j can bring about A_i and have an FSP for A_i.
    - Other members of group Θ intend that subgroup Θ_j can bring about A_i using some recipe.

Most of the times a team and its members do not possess an *FSP* to achieve their shared goal. In this case, the concept of *FSP* puts limits on the SharedPlans theory. However, SharedPlans uses the concept of *PSP* as a snapshot of the team's mental states in different situations, which further leads to communication and planning to fulfill the conditions of an *FSP*. The idea behind *PSP* is enabling the agents to modify the shared plan over the course of planning without impairing the achievement of the shared goals. Notice that for the same reason recipes also can be partial [27, 29].

2.1.4 Communicating Intentions

In SharedPlans theory Grosz and Sidner are interested in the type of plans that underlie a discourse in which the agents collaborate to achieve a shared goal. Here we present their view of discourse structure, since it is directly related to the intentions behind collaborators' actions. In [29], Grosz and Sidner argue that the SharedPlans theory recognises three interrelated levels of discourse structure, and 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 [29]. Using discourse plans can help to encode the knowledge about conversation.

2.1.5 Intention-to and Intention-that

In Grosz and Sidner's SharedPlans theory [29], two intentional attitudes are employed: *intending to* (do an action) and *intending that* (a proposition will hold). The notion of *intention to*, as an individual-oriented intention, models the intention of an agent to do any single-agent action while the agent not only believes that it is able to execute that action, but it also commits to doing so. In short, it is an intention to perform an action, similar to Bratman's view of intention. In contrast with *intention to*, an *intention that*, as an intention directed toward group activity, does not directly imply an action. In fact, an individual agent's *intention that* is directed towards its collaborators' action or towards a group's joint action. *Intention that* guides an agent to take actions (including communication), that enable or facilitate other collaborators to perform assigned tasks. This leads an agent to behave collaboratively. Therefore, agents will adopt intentions to communicate about the plan [27]. As another difference, *Intention to* commits an agent to means-end reasoning and acting [4] while *Intention that* does not necessarily entail this commitment. The key point about *Intention to* and *intention that* is that both commit an agent not to adopt conflicting intentions, and constrain replanning in case of failure. Further, an agent can *intention that* another agent achieve the specified proposition.

2.2 Joint Intentions Theory

Following Bratman's guidelines, Cohen and Levesque propose a formal approach to building artificial collaborative agents. The Joint Intentions theory of Cohen and Levesque [15, 16, 17, 18, 43] 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. A joint intention is a shared commitment to perform an action while in a group mental state [16].

In [15] Cohen and Levesque establish that joint intention cannot be defined simply as individual intention with the team regarded as an individual. The reason is that after the initial formation of an intention, team members may diverge in their beliefs and their attitudes towards the intention. Instead, Cohen and Levesque generalize their own definition of intention. First, they present a definition of individual persistent goal (see Section

2.2.1) and individual intention (see Section 2.2.2). Then, they define analogues of these concepts by presenting mutual belief in place of individual belief. The definition of joint persistent goal (see Section 2.2.3) requires team members to commit to informing other members, if it comes to believe that the shared goal is in its terminal status. As a result, in Cohen and Levesque's theory, a team with a joint intention is a group that shares a common objective and a certain shared mental state [34].

In this theory, once an agent entered into a joint commitment with other agents, the agent should communicate its private beliefs with other team members if the agent believes that the joint goal is in its terminal status, i.e., either the joint goal is achieved, or it is unachievable, or irrelevant [72]. Thus, as we mentioned above, team members are committed to inform other team members when they reach the conclusion that a goal is achievable, impossible, or irrelevant. For instance, if a robot and an astronaut are collaborating to install a solar panel, and the robot reaches the conclusion that the welding tool has deficiency, it is essential for the robot to have an intention to communicate with the astronaut and make this knowledge common. Therefore, according to this theory, in a collaboration, agents can count on the commitment of other members, first to the goal and then to the mutual belief of the status of the goal.

2.2.1 Individual Commitment

As we mentioned earlier, intentions and commitments are the basic ideas of Joint Intentions theory. Here, we provide the definition of "individual commitment" (also called *persistent goal*) by Cohen et. al. in [14]. According to their definition an agent has a persistent goal relative to q to achieve p only when:

1. agent believes that p is currently false;
2. agent wants p to be true;
3. it is true (and agent knows it) that (2) will continue to hold until the agent comes to believe either that p is true, or that it will never be true, or that q is false.

Note that the condition q is an "escape" clause, which can be omitted for brevity, or it can be used as a reason for the agent to drop a commitment, even though it could be quite vague.

2.2.2 Individual Intention

As we mentioned in this Section, Joint Intention theory adopts Bratman's view of future-directed properties of intention. In this theory, an intention is defined to be a commitment to act in a certain mental state. In other words, an agent intends relative to some condition to do an action just in case it has a persistent goal or commitment (relative to that condition) of having done the action and, moreover, believing throughout that it is doing that action [15].

Intention inherits all the properties of commitment (e.g., consistency with mental states). Typically, an agent uses an intention as a decision within a subgoal-supergoal hierarchy to do a particular action. For instance, initially, the agent commits to p becoming true without having any concern about who or how p is going to be accomplished. Then, the agent commits to x or y as a mean to accomplish p. Lastly, the agent selects one of the actions (e.g., x) and forms an intention to do it. This intention will be given up when for whatever reason p is accomplished.

2.2.3 Joint Commitment

Before talking about joint commitment, we provide the definition of the *Weak Achievement Goal* (WAG) concept in Joint Intentions theory which shows the state of a team member nominally working on a goal. The concept of WAG is used to provide the definition of the Joint Commitment in this theory.

An agent has a WAG relative to q and with respect to a team to bring about p if either of the following conditions holds:

- The agent has a normal achievement goal to bring about p; that is, the agent does not yet believe that p is true and wants p to be true as a goal.
- The agent believes that p is true, will never be true, or is irrelevant, but has as a goal that the status of p be mutually believed by all the team members.

Joint commitment – A joint intention of a team Θ is based on its joint commitment, which is defined as a *Joint Persistent Goal* (JPG). A JPG to achieve a team action p, denoted JPG(Θ, p) requires all team members to mutually believe that p is currently false and want p to eventually be true.

A JPG guarantees that team members cannot decommit until p is mutually known to be *achieved, unachievable* or *irrelevant*. Basically, JPG(Θ, p) requires team members to each hold p as a *Weak Achievement Goal* (WAG). WAG(μ, p, Θ), where μ is a team member in Θ, requires μ to achieve p if it is false. However, if μ privately believes that p is either achieved, unachievable or irrelevant, JPG(Θ, p) is dissolved, but p is left with a commitment to have this belief become Θ's mutual belief. Such a commitment is required to establish mutual belief in Θ; this commitment typically makes an agent communicate with its teammates [15].

An important consequence of achieving joint commitment in a team is that it predicts future communication which is critical within the course of a collaboration. Thus, this communication leads team members to attain mutual beliefs which is a fundamental concept in teamwork activities. Notice that the minimum mutual belief for team members to attain is the achievement or failure of the shared goal which terminates collaboration.

2.2.4 Joint Intention

Joint intention is defined to be a joint commitment to the team members trying to do a joint action. Based on Cohen and Levesque's definition of joint intention, a team of agents jointly intends (relative to some escape condition) to do an action if and only if the members have a JPG (relative to that condition) of their having done the action, and having done it mutually believing throughout that they were doing it (knowingly) [15].

2.2.5 Teamwork & Communication

In summary, according to Joint Intentions theory, the notion of teamwork is characterized by joint commitment, also known as joint persistent goal (see Section 2.2.3). The definition of JPG states that the agents mutually believe they have the appropriate goal, and that they mutually believe a persistent weak achievement goal (which represents the one-way commitment of one agent directed towards another) to achieve it persists until the agents mutually believe that the goal has either been achieved, or become impossible, or irrelevant.

Joint Intentions theory claims that an efficient collaboration requires communication. Sharing information through communication is critical given that collaborators have different capabilities, and each individual often has only partial knowledge relevant to solving the problem, and sometimes diverging beliefs about the state of the collaborative activity. Communication plays an important role in coordinating team members' roles and actions to accomplish their goal. For instance, it can help team members to establish and maintain a set of mutual beliefs regarding the current state of the collaboration, and the respective roles and capabilities of each member.

2.3 STEAM – A Hybrid Collaboration Approach

Tambe in [70] argues that teamwork in complex, dynamic, multi-agent domains requires the agents to obtain flexibility and reusability by using integrated capabilities. Tambe created STEAM (simply, a **S**hell **TEAM**work) based on this idea. STEAM's operationalization in complex, real-world domains is the key in its development to addressing important teamwork issues, some of which are discussed in Section 4. 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.

Building on the well developed theory of joint intentions [15] and sharedPlans [27, 29], the STEAM teamwork model [70] was operationalized as a set of domain-independent rules that describe how teams should work together. According to Tambe's claim, several advantages accrue due to this use of Joint Intentions theory, such as achieving a principled framework for reasoning about coordination and communication in a team, which the joint intention can provide. Another advantage is the guidance for monitoring and maintenance of a team activity which the joint commitment in joint intention again provides. And lastly, Tambe believes the joint intention in a team can facilitate reasoning about team activity and team members' contribution to that activity.

However, he also believes that for a high level team goal, one single joint intention is not sufficient to achieve all these advantages. Thus, STEAM borrows some of the concepts of SharedPlans theory. First, STEAM uses the concept of "intention that" (see Section 2.1) towards an activity as well as the fact that SharedPlans theory mandates team members' mutual belief in a common recipe and shared plans for individual steps in the common recipe. Thus, in this case, SharedPlans helps STEAM to achieve coherency within the teamwork. Besides, STEAM uses joint intentions to ensure the teamwork coherency to build the mental attitudes of team members. In other words, as the recipe evolves, STEAM requires all team members to agree on the execution of a step and form joint intentions to execute it while other joint intentions are formed, leading to a hierarchy. A second concept STEAM borrows from SharedPlans is the amount of information that a team member needs to know to perform an action. According to Shared-

Plans, team members require to know only that a recipe exists to enable them to perform actions (not recipe details – see Section 2.1.1). Similarly in STEAM, team members only track the responsible subteam or individual team member to perform a specific step while this tracking does not need detailed plan recognition. The third issue is parallel to what is called an unreconciled case in SharedPlans theory, which in STEAM is handled by replanning and communication between team members assigning the unassigned or unachieved task. The last issue is communication between team members which also borrows the concept of "intention that" from SharedPlans theory, to help the generalization of STEAM's communication capabilities beside what Joint Intentions theory offers.

In summary, STEAM builds on both Joint Intention Theory and SharedPlans theory and tries to overcome their shortcomings. Based on joint intentions, STEAM builds up hierarchical structures that parallel the SharedPlans theory. Hence, STEAM formalizes commitments by building and maintaining Joint Intentions, and uses SharedPlans to formulate the team's attitudes in complex tasks.

In [70] Tambe argues that the novel aspects of STEAM relate to its teamwork capabilities. The key novelty in STEAM has team operators beside individual team member operators. In STEAM when agents select a team operator for execution, they instantiate a team's joint intentions. Team operators explicitly express a team's joint activities, unlike the regular individual operators which express an agent's own activities. Hence, STEAM agents maintain their own private (to apply individual operators) and team states, e.g., mutual belief about the world (to apply team operators).

However, Tambe added more practical concepts into STEAM's architecture. For instance, STEAM has a team synchronization protocol to establish joint intention (see JPG in Section 2.2), or it has constructs for monitoring joint intentions which helps the agent to be able to monitor team performance. STEAM facilitates this monitoring by exploiting its explicit representation of team goals and plans. In particular, STEAM allows an explicit specification of monitoring conditions to determine achievement, unachievability or irrelevancy conditions of team operators. Finally, in STEAM, communication is driven by commitments embodied in the Joint Intentions theory, i.e., team members may communicate to obtain mutual belief while building and disbanding joint intentions. Thus, joint intentions provide STEAM with a principled framework for reasoning about communication. Also, STEAM addresses some practical issues, not addressed in other teamwork theories. One of these issues is STEAM's detailed attention to communication overheads and risks, which can be significant [71]. Furthermore, operationalization of STEAM is based on enhancements to the Soar architecture [42], plus a set of about 300 domain-independent Soar rules.

2.4 Other Approaches

There are other frameworks, approaches, and models focusing on teamwork and collaborative agents. For instance, Jennings provides the Joint Responsibility framework which is specified formally using modal, temporal logic. Joint Responsibility stresses the role of joint intentions (based on Joint Intentions theory) specifying how both individuals and teams should behave whilst engaged in collaborative problem solving [35, 36, 37, 38]. Jennings has developed *Generic Rules and Agent model Testbed Environment* (GRATE) as a prototype system based on the Joint Responsibility framework. In [40] Kinny et. al. elaborate the concept of Planned Team Activity and introduce a language for representing joint plans for teams of agents and describe how agents can organize the formation of a skilled team to achieve a joint goal. They use joint intentions to capture the mental properties which characterize team activity.

3 Similarities and Differences

There are some similarities between SharedPlans and Joint Intentions theories. Here, we specify some of these similarities:

1. Similar to SharedPlans theory, Joint Intentions theory specifies what it means for agents to execute actions as a team [69].

2. Both theories follow Bratman's basic ideas about intention's roles in relational actions which prevent the collaborative agents from adopting conflicting intentions. Besides, these two theories are also agreed and follow Bratman's BDI model.

3. Just as SharedPlans theory, Joint Intentions theory also states that a joint action could not be seen as a collection of individual actions but that agents working together need to share beliefs.

4. Both theories in their latest articles show that the agents are required to communicate to maintain collaboration. SharedPlans theory requires collaborators to communicate to establish and maintain the shared plan which is crucial especially when collaborators only have

partial shared plan. Similarly in Joint Intentions theory, communication is an explicit requirement of collaborative agents until the shared goal is achieved, unachievable or irrelevant.

5. Both Joint Intentions and SharedPlans theories are concerned about commitment to the joint activity. Although, these two theories use different concepts to fulfill the requirements of commitment during collaboration.

There are also differences between SharedPlans and Joint Intention theories; we address some of them here in this section:

1. Although the crucial components of the SharedPlans theory (see Section 2.1) lack the notion of a joint intention, which is the most significant notion within the Joint Intentions theory, Grosz and Sidner do not believe that such a phenomenon (joint intention) exists in a collaboration. They believe their notion of "intention that" and mutual beliefs about states of the collaboration can provide similar functionalities as described in Joint Intentions theory (see Section 2.2).

2. In SharedPlans theory teammates agree on the shared plan, whereas in Joint Intentions theory teammates agree on intentions.

3. In contrast to Joint Intentions, the SharedPlans theory employs hierarchical structures over intentions, thus it overcomes the shortcoming of a single joint intention for complex team tasks.

4. The SharedPlans theory describes the way to achieve a common goal through the hierarchy of plans, whereas the Joint Intentions theory describes only this common goal [67].

5. Joint Intentions theory assumes that knowledge about the teammates is always available, whereas SharedPlans theory uses the concept of partial plan/recipe to make the process of dynamically achieving information possible through out the collaboration.

6. Communication requirements are derived from "intention that" in SharedPlans theory, as opposed to being "hard-wired" in Joint Intentions theory.

A critique to Joint Intention theory – Castelfranchi criticizes the necessary and sufficient conditions (see Section 2.2) for the joint persistent goal which plays a crucial role in the Joint Intentions theory. According to his example, if a French scientist and an American scientist are both working on an AIDS vaccine and both have the final goal of p "vaccine anti-AIDS be found" relative to the belief q that "if vaccine is found, AIDS is wiped out", they both share the mental attitudes described in Joint Intentions theory. It means that they mutually believe that p is currently false, and they mutually know they both want p to be true, and it is true that until they come to believe either that p is true, that p will never be true, or that q is false, they will continue to mutually believe that they each have a weak achievement goal (see Section 2.2) relative to q and with respect to the team (i.e., the WAG with respect to the team has been defined as "a goal that the status of p be mutually believed by all the team members"). The problem is that we can not claim the French and American professors are working as a team. In fact, given their personal goals of finding the vaccine, they might come to strongly compete with each other [11].

4 Application in Human-Computer Collaboration

There are many research focusing on different aspects of collaboration based on different collaboration theories, i.e., SharedPlans, Joint Intentions, and hybrid theories of collaboration. In this section, we provide some examples of homogeneous and heterogeneous agent/robot and human collaborations.

There are some works focusing on the concepts of robot assistants [13], or teamwork and its challenges in cognitive and behavioral levels [52, 62]. Some researchers have an overall look at a collaboration concept at the architectural level. In [20] authors present a collaborative architecture, COCHI, to support the concept of emotional awareness. In [19] authors present the integration of emotional competence into a cognitive architecture which runs on a robot, MEXI. In [68] authors discuss the challenges of integrating natural language, gesture understanding and spatial reasoning of a collaborative humanoid robot situated in space. The importance of communication during collaboration has also been considered by some researchers from human-computer interaction and human-robot collaboration [12, 48, 60] to theories describing collaborative negotiation, and discourse planning and structures [1, 28, 65]. There are other concepts such as joint actions and commitments [25], dynamics of intentions during collaboration [43], and task-based planning providing more depth in the context of collaboration [9, 58]. The concept of collaboration has also received attention in the industry and in research in robotic laboratories [21].

Applications of SharedPlans Theory – COLLAGEN [59, 60] is the first implemented system based on the SharedPlans theory. It incorporates certain algorithms for discourse generation and interpretation, and is able to maintain a segmented interaction history, which facilitates the discourse between the human user and the intelligent agent. The model includes two main parts: (1) a representation of a discourse state and (2) a discourse interpretation algorithm for the utterances of the user and agent [61]. In [30] Heeman presents a computational model of how a conversational participant collaborates in order to make a referring action successful. The model is based on the view of language as goal-directed behaviour, and in his work, he refers to SharedPlans as part of the planning and conversation literature. In [46], Lochbaum and Sidner modify and expand the SharedPlan model of collaborative behavior [29]. They present an algorithm for updating an agents beliefs about a partial shared plan and describe an initial implementation of this algorithm in the domain of network management. Lochbaum, also in [45], provides a computational model (based on the collaborative planning framework of SharedPlans [27]) for recognizing intentional structure and utilizing it in discourse processing. In short, she presents a SharedPlans model for recognizing Discourse Segment Purposes (DSPs) [29] [65] and their interrelationships. CAST (Collaborative Agents for Simulating Teamwork) [73] [74] is a teamwork framework based on the SharedPlans Theory. CAST focuses on flexibility in dynamic environments and on proactive information exchange enabled by anticipating what information team members will need. Petri Nets are used to represent both the team structure and the teamwork process, i.e., the plans to be executed. Researchers in [32] discuss developing an ontology of microsocial concepts for use in an instructional system for teaching cross-cultural communication. They believe being acquainted with one another is not a strong enough relationship from which to create a society. Hence, there is a need for commitment and shared plans (as the basis of social life) to achieve a shared goal. In this work, Grosz and Sidner's SharedPlans theory [29] is used to explain the concept of shared plans within the interpersonal relationships of societies in an industrial environment. In [33] Hunsberger and Grosz discuss the idea of whether the rational, utility-maximizing agents should determine commitment to a group activity when there is an opportunity to collaborate. They call this problem the "initial-commitment decision problem" (ICDP) and provide a mechanism that agents can use to solve the ICDP. They use the representation of action, act-types and recipes in the SharedPlans theory. In [75] an integrated agent-based model for Group Decision Support Systems is proposed and discussed. The decisional model that authors outline in this paper is based on the SharedPlans theory. Rauenbusch and Grosz in [56] formally define a search problem with search operators that correspond to the team planning decisions. They provide an algorithm for making the three types of interrelated decisions by recasting the problem as a search problem. Their model respects the constraints on mental states specified by the SharedPlans theory of collaboration. Babaian et. al. in [3] describe Writer's Aid, a system that deploys AI planning techniques to enable it to serve as an author's collaborative assistant. While an author writes a document, Writer's Aid helps in identifying and inserting citation keys and by autonomously finding and caching potentially relevant papers and their associated bibliographic information from various on-line sources. They believe the underlying concepts of SharedPlans is relevent since in collaborative interfaces like Writers Aid, the users establish shared goals with the system and user and the system both take initiative in satisfying them. In [49] researchers address high-level robot planning issues for an interactive cognitive robot that acts in the presence of or in collaboration with a human partner. They describe a Human Aware Task Planner (HATP) which is designed to provide socially acceptable plans to achieve collaborative tasks. They use notions of plans based on SharedPlans theory. In [66] Sidner and Dzikovska argue that robots, in order to participate in conversations with humans, need to make use of conventions of conversation and the means to be connected to their human counterparts. They provide an initial research on engagement in human-human interaction and applications to stationary robots in hosting activities. They believe hosting activities are collaborative because neither party completely determines the goals to be undertaken nor the means of reaching the goal. To build a robot host, they rely on an agent built using COLLAGEN which is implemented based on the SharedPlans theory.

Applications of Joint Intentions Theory – In [40] authors introduce a language for representing joint plans for teams of agents. They describe how agents can organize the formation of a suitably skilled team to achieve a joint goal, and they explain how such a team can execute these plans to generate complex, synchronized team activity. In this paper, authors adopt the underlying concepts of the Joint Intentions theory as the structure of their collaborative agents. Breazeal et. al. in [8] present an overview of their work towards building socially intelligent, cooperative humanoid robots, Leonardo, that can collaborate and learn in partnership with humans. They employ the Joint Intentions theory of collaboration to implement the collaborative behaviors while performing a task in collaboration with humans.

In [69] the researchers' goal is to develop an architecture (based on the concepts of Joint Intentions theory) that can guide an agent during collaborative teamwork. They describe how a joint intention interpreter that is integrated with a reasoner over beliefs and communicative acts can form the core of a dialogue engine. Ultimately, the system engages in dialogue through the planning and execution of communicative acts necessary to attain the collaborative task at hand. Mutlu et. al. in [50] discuss key mechanisms for effective coordination toward informing the design of communication and coordination mechanisms for robots. They present two illustrative studies that explore how robot behavior might be designed to employ these mechanisms (particularly joint attention and action observation) to improve measure of task performance in human-robot collaboration. Their work uses Joint Intentions theory to develop shared task representations and strategies for task decomposition. The system GRATE* by Jennings [37] is based on the Joint Intention Theory. GRATE* provides a rule-based modelling approach to cooperation using the notion of Joint Responsibilities, which in turn is based on Join Intentions. GRATE* is geared towards industrial settings in which both agents and the communication between them can be considered to be reliable.

Applications of Hybrid Theories – This domain independent teamwork model, STEAM, has been successfully applied to a variety of domains. From combat air missions [31] to robot soccer [41] to teams supporting human organizations [55] to rescue response [63], applying the same set of STEAM rules has resulted in successful coordination between heterogeneous agents. The successful use of the same teamwork model in a wide variety of diverse domains provides compelling evidence that it is the principles of team- work, rather than exploitation of specific domain phenomena, that underlies the success of teamwork based approaches. In [47] authors provide their in RoboCup (robotics soccer testbed) in which their focus is on teamwork and learning challenges. Their research investigation in RobotCup is based on ISI Synthetic, a team of synthetic soccer-players. They also investigate the use of STEAM as their model of teamwork which is influenced by the Joint Intentions and SharedPlans theories. In [39] researchers propose a behavioral architecture C^2BDI that allows the enhancement of the knowledge sharing using natural language communication between team members. They define collaborative conversation protocols that provide proactive behavior to agents for the coordination between team members. Their agent architecture provides deliberative and conversational behaviors for collaboration, and it is based on both of the SharedPlans and Joint Intentions theories.

5 Conclusion

In this response, we started by defining the concept of collaboration based on Grosz and Sidner's work [29], and listed a number of collaboration properties. Then, we provided the background of two prominent collaboration theories which helped develop a better understanding of the actual theories and how they relate to each other. Next, we presented the SharedPlans theory and its major properties, e.g., partial shared plan, recipe, and two notions of intention. Afterwards, we delivered key concepts of the Joint Intentions theory including joint commitment and joint intention. Then, we continued with the hybrid approach of modeling collaboration and provided one of the most well-established models, STEAM. We also briefly mentioned some other approaches. Later, we presented two different lists to compare similarities and differences between SharedPlans and Joint Intentions collaboration theories. We ended this document with different categories of applications of these theories in agent/robot and human collaboration areas.

We believe the SharedPlans and Joint Intentions collaboration theories are the most well-defined and well-established theories in computer science. We 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. We 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. We 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. Although all these 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.

References

[1] Jerry Andriessen, Koenraad de Smedt, and Michael Zock. Discourse planning: Empirical research and computer models. In Anton Dijkstra and Koenraad de Smedt, editors, *Computational psycholinguistics: AI and connectionist models of human language processing*, pages 247–278. Taylor & Francis, 1996.

[2] Gertrude Elizabeth Margaret Anscombe. *Intention.* NY: Cornell University Press, 1963.

[3] Tamara Babaian, Barbara J. Grosz, and Stuart M. Shieber. A writer's collaborative assistant. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI2000)*, pages 7–14. ACM Press, 2002.

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

[5] Michael E. Bratman. What is intention? In *Intentions in Communication*, Cognitive Technologies, pages 15–32. The MIT Press, Cambridge, MA, June 1990.

[6] Michael E. Bratman. Shared cooperative activity. *Philosophical Review*, 101(2):327–341, 1992.

[7] Michael E. Bratman, David J. Israel, and Martha E. Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4(3):349–355, 1988.

[8] Cynthia Breazeal, Andrew Brooks, Jesse Gray, Guy Hoffman, Cory Kidd, Hans Lee, Jeff Lieberman, Andrea Lockerd, and David Mulanda. Humanoid robots as cooperative partners for people. *Journal of Humanoid Robots*, 1(2):1–34, 2004.

[9] Catherina Burghart, Ralf Mikut, Rainer Stiefelhagen, Tamim Asfour, Hartwig Holzapfel, Peter Steinhaus, and Ruediger Dillmann. A cognitive architecture for a humanoid robot: A first approach. In *5th IEEE-RAS International Conference on Humanoid Robots*, pages 357–362, 2005.

[10] Hector-Neri *Castañeda. Thinking and Doing.* Dordrecht, Holland: D. Riedel, 1975.

[11] Cristiano Castelfranchi. Commitments: From individual intentions to groups and organizations. In *Proceedings of the first international conference on multiagent systems*, pages 41–48, 1995.

[12] Aaron B. St. Clair and Maja 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.

[13] William 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.

[14] Philip Cohen, Hector Levesque, and Ira Smith. On team formation. In *Contemporary Action Theory. Synthese*, pages 87–114. Kluwer Academic Publishers, 1997.

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

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

[17] 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.

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

[19] Natascha Esau, Lisa Kleinjohann, and Bernd Kleinjohann. Integrating emotional competence into man-machine collaboration. In *Biologically-Inspired Collaborative Computing, September 8-9, Milano, Italy*, pages 187–198, 2008.

[20] Octavio García, Jessús Favela, Guillermo Licea, and Roberto Machorro. Extending a collaborative architecture to support emotional awareness. In *Emotion Based Agent Architectures (ebaa99*, pages 46–52, 1999.

[21] Scott A. Green, Mark Billinghurst, XiaoQi Chen, and J. Geoffrey Chase. Human-robot collaboration: A literature review and augmented reality approach in design. *International Journal of Advanced Robotic Systems*, 5(1):1–18, 2008.

[22] Barbara Grosz and Sarit Kraus. The evolution of shared plans. In *Foundations and Theories of Rational Agency*, pages 227–262, 1998.

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

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

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

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

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

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

[29] 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.

[30] Peter Anthony Heeman. *A Computational Model of Collaboration on Referring Expressions*. PhD thesis, University of Toronto, 1991.

[31] Randall W. Hill, Jr., Johnny Chen, Jonathan Gratch, Paul Rosenbloom, and Milind Tambe. Intelligent agents for the synthetic battlefield: A company of rotary wing aircraft. In *Innovative Applications of Artificial Inttigence (IAAI-97)*, pages 227–262, 1997.

[32] Jerry R. Hobbs, Alicia Sagae, and Suzanne Wertheim. Toward a commonsense theory of microsociology: Interpersonal relationships. In *Formal Ontology in Information Systems - Proceedings of the Seventh International Conference*, pages 249–262, 2012.

[33] Luke Hunsberger and Barbara J. Grosz. A combinatorial auction for collaborative planning. In *In Proceedings of ICMAS*, 2000.

[34] Bevan Jarvis, Dennis Jarvis, and Lakhmi Jain. Teams in multi-agent systems. In *Intelligent Information Processing III*, volume 228. Springer US, 2007.

[35] Nicholas R. Jennings. *Joint Intentions as a Model of Multi-Agent Cooperation in Complex Dynamic Environments*. PhD thesis, Department of Electronic Engineering, University of London, 1992.

[36] Nicholas R. Jennings. On being responsible. In E. Werner and Y. Demazeau, editors, *Proceedings of the Third European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, pages 93–102. North-Holland, 1992.

[37] Nicholas R. Jennings. Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence*, 75(2):195–240, 1995.

[38] Nicholas R. Jennings and E. H. Mamdani. Using joint responsibility to coordinate collaborative problem solving in dynamic environments. In *10th National Conference on Artificial Intelligence (AAAI-92)*, pages 269–275, 1992.

[39] Alexandre Kabil, Camille De Keukelaere, and Pierre Chavaillier. Coordination mechanisms in human-robot collaboration. In *Proceeding of the 7th International Conference on Advances in Computer-Human Interactions*, pages 389–394, 2014.

[40] David Kinny, Magnus Ljungberg, Anand Rao, Gil Tidhar, Eric Werner, and Elizabeth Sonenberg. Planned team activity. In *Lecture notes in artificial intelligence*. Springer-Verlag, 1992.

[41] Hiroaki Kitano, Minoru Asada, Yasuo Kuniyoshi, Itsuki Noda, Eiichi Osawai, and Hitoshi Matsubara. Robocup: A challenge problem for AI. *AI Magazine*, 18(1):73–85, 1997.

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

[43] 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.

[44] 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.

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

[46] Karen E. Lochbaum, Barbara J. Grosz, and Candace 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.

[47] Stacy Marsella, Jafar Adibi, Yaser Al-Onaizan, Ali Erdem, Randall Hill, Gal A. Kaminka, Zhun Qiu, and Milind Tambe. Using an explicit teamwork model and learning in robocup: An extended abstract. In *RoboCup-98: Robot Soccer World Cup II*, volume 1604, pages 237–245. Springer Berlin Heidelberg, 1999.

[48] Laetitia Matignon, Abir Beatrice Karami, and Abdel-Illah Mouaddib. A model for verbal and non-verbal human-robot collaboration. In *AAAI Fall Symposium Series*, pages 62–67, 2010.

[49] Vincent Montreuil, Aurélie Clodic, Maxime Ransan, and Rachid Alami. Planning human centered robot activities. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pages 2618–2623, 2007.

[50] Bilge Mutlu, Alison Terrell, and Chien-Ming Huang. Coordination mechanisms in human-robot collaboration. In *Proceedings of the HRI 2013 Workshop on Collaborative Manipulation*, 2013.

[51] Stephen Neale. Paul grice and the philosophy of language. *Linguistics and Philosophy*, 15(5):509–559, 1992.

[52] Stefanos Nikolaidis, Przemyslaw A. Lasota, Gregory F. Rossano, Carlos Martinez, Thomas A. Fuhlbrigge, and Julie A. Shah. Human-robot collaboration in manufacturing: Quantitative evaluation of predictable, convergent joint action. In *ISR*, pages 1–6, 2013.

[53] Martha 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.

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

[55] David V. Pynadath and Milind Tambe. An automated teamwork infrastructure for heterogeneous software agents and humans. *Journal of Autonomous Agents and Multi-Agent Systems, Special Issue on Infrastructure and Requirements for Building Research Grade Multi-Agent Systems,*, 7(1-2):71–100, 2003.

[56] Timothy W. Rauenbusch and Barbara J. Grosz. A decision making procedure for collaborative planning. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, 2003.

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

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

[59] Charles Rich and Candace L. Sidner. COLLAGEN: A collaboration manager for software interface agents. *User Modeling User-Adapted Interaction*, 8(3-4):315–350, 1998.

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

[61] Jeff Rickel, Neal Lesh, Charles Rich, Candace L. Sidner, and Abigail Gertner. Collaborative discourse theory as a foundation for tutorial dialogue. In *Proceedings Sixth International Conference on Intelligent Tutoring Systems*, 2002.

[62] Paul Scerri, David Pynadath, Lewis Johnson, Paul Rosenbloom, Mei Si, Nathan Schurr, and Milind 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.

[63] Paul Scerri, David Pynadath, Lewis Johnson, Paul Rosenbloom, Mei Si, Nathan Schurr, and Milind Tambe. A prototype infrastructure for distributed robot-agent-person teams. In *The Second International Joint Conference on Autonomous Agents and Multiagent Systems*, 2003.

[64] John R. Searle. Collective intentionality. In *Intentions in Communication*, pages 401–415. MIT Press, 1990.

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

[66] Candace L. Sidner and Myroslava Dzikovska. A first experiment in engagement for human-robot interaction in hosting activities. In *Advances in Natural Multimodal Dialogue Systems*, volume 30 of *Cognitive Technologies*, pages 55–76. Springer Netherlands, 2005.

[67] Hendrik Skubch. *Modelling and Controlling of Behaviour for Autonomous Mobile Robots*. Springer Science Business Media, 2012.

[68] Donald Sofge, Magdalena D. Bugajska, J. Gregory Trafton, Dennis Perzanowski, Scott Thomas, Marjorie Skubic, Samuel Blisard, Nicholas Cassimatis, Derek P. Brock, William Adams, and Alan C. Schultz. Collaborating with humanoid robots in space. *International Journal of Humanoid Robotics*, 2(2):181–201, 2005.

[69] Rajah Annamalai Subramanian, Sanjeev Kumar, and Philip Cohen. Integrating joint intention theory, belief reasoning, and communicative action for generating team-oriented dialogue. In *AAAI*, pages 1501–1507. AAAI Press, 2006.

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

[71] Milland Tambe. Agent architecture for flexible, practical teamwork. In *Proceedings of the National Conference on Artificial Intelligence*, pages 22–28, 1997.

[72] Burt Wilsker. A study of multi-agent collaboration theories. In *ISI Research Report*, pages 396–449, 1996.

[73] John Yen, Jianwen Yin, Thomas R. Ioerger, Michael S. Miller, Dianxiang Xu, and Richard A. Volz. Cast: Collaborative agents for simulating teamwork. In *Proceedings of IJCAI2001*, pages 1135–1142, 2001.

[74] Jianwen Yin, Michael S. Miller, Thomas R. Ioerger, John Yen, and Richard A. Volz. A knowledge-based approach for designing intelligent team training systems. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 427–434. ACM, 2000.

[75] Zamfirescu and Candea. On integrating agents into gdss. In *Preprints of the 9th IFAC / IFORS / IMACS / IFIP/ Symposium on Large Scale Systems: Theory and Applications*, 2001.