The Dynamics of Intention in Collaborative Activity

Barbara J. Grosz [a] and Luke Hunsberger [b]

[a] *Division of Engineering and Applied Sciences*
Harvard University
grosz@eecs.harvard.edu

[b] *Department of Computer Science*
Vassar College
hunsberg@cs.vassar.edu

Abstract

An adequate formulation of collective intentionality is crucial for understanding group activity and for modeling the mental state of participants in such activities. Although work on collective intentionality in philosophy, artificial intelligence, and cognitive science has many points of agreement, several critical issues remain under debate. This paper argues that the dynamics of intention—in particular, the inter-related processes of decision making and intention updating—play crucial roles in an explanation of collective intentionality, and it is in these dynamic aspects that coordinated group activity differs most from individual activity. The paper specifies a model of the dynamics of agent intentions in the context of collaborative activity and defines an architecture for a collaboration-capable computer agent based on that model. This integrated treatment of group decision-making and coordinated updating of group-related intentions fills an important gap in prior accounts of collective intentionality and helps resolve a long-standing debate about the nature of intentions in collaborative activity.

1 Introduction

There is broad agreement in philosophy, artificial intelligence and cognitive science [3, 39, 13, 14, 47, 46, 17, 27, 24, 10] that the collective, joint activity of a group is more than the simple sum of the individual, domain-oriented actions

⋆ The research reported in this paper was supported in part by National Science Foundation grants IIS-0222892 and IIS-9978343 to Harvard University.

of the members of the group; coordinating activities, typically including some communicative actions, are required. There is also agreement that the plans that underlie the collective, collaborative activity of a group involve more than the simple sum of the individual plans of the members of the group and that the simple sum of the intentions of the group members toward their own actions does not fully capture collective intentionality. However, there is disagreement about what is required to fill the gap between the whole and the (simple) collection of individual actions, intentions and plans.

Bratman [3] claims that no new kind of intention is required for characterizing collective action and intentionality. He argues that an interlocking web of beliefs, mutual beliefs and ordinary intentions is sufficient and that the coordination and commitment needed for collective, cooperative activities may be accommodated through the content of each participant's group-related intention including "that we J in accordance with and because of [our intentions] and meshing subplans of [our intentions]." Some computational approaches to formalizing collaboration and constructing collaborative multi-agent systems also rely solely on ordinary, individual intentional attitudes [17, 13, 14, 24].

In contrast to such "individualistic" accounts, several philosophers and cognitive scientists have argued that collective intentionality requires a different kind of intentional attitude, one that although individually held is different from (and not reducible to) an ordinary intention. For instance, Searle [40] claims that: "In addition to singular intentionality there is also collective intentionality," where "the intentionality that exists in each individual head has the form 'we intend'." He argues that collective intentionality "is a biologically primitive phenomenon that cannot be reduced to or eliminated in favor of something else." Similarly, Tuomela [47] argues for individually held "we-intentions" that are "agreement-based social intentions that agents have in situations of joint action" with content that "can be taken to be something like 'to do X jointly' or 'we to do X jointly'." He also distinguishes intending in the *we-mode* from intending in the *I-mode*, and argues that "the we-mode is not reducible to the I-mode and these modes may be in conflict" [45]. Gilbert [10] presents a "plural subject" account of collective intentionality in which agents form a joint commitment to "intend as a body." She argues that the individual "commitment shares" are not, and do not require the existence of personal commitments. Several computational approaches employ notions of joint intention that, although reducible to sets of individually held attitudes, reduce to something other than ordinary intentions [27, 23, 43].

Several accounts of collective intentionality argue that agreements and their entailed obligations are central to distinguishing group activities from simple sums of individual activities. Obligations serve to bind the agents together in coordinating and pursuing their collaborative activity. For instance, in Tuomela's account of full-blown, agreement-based joint intention [47], agreements entail certain obligations on the participants to adopt relevant intentions. Similarly, in Gilbert's account [10], joint commitments have inherent obligations on agents to provide appropriate action. Castelfranchi [5] argues for "social commitments", which involve similar obligations, rights and entitlements.

To reconcile these different stances requires a deeper look at the dynamics of group activity. Collaborative, multi-agent activity has certain features in common with individual activity. For any complex activity to be done by resource-bounded agents, whether people or artificial systems, it must be possible to form initial, incomplete plans and revise them over time [4, 13]. The intentions in such plans are typically under-specified. If George has not yet decided whether to get tomatoes from the grocery store or the farmer's market, his plan to make Caprese salad will be incomplete and his intention "to make a Caprese salad" under-specified. Likewise, if we have not yet decided which movie to see tonight, our plan to go to the movies will be incomplete, and the intention we share "to go to the movies tonight" will be under-specified. As a result of such incompleteness, plans and intentions related to actions, whether of individuals or of groups, require that means-ends reasoning be done, that certain decisions be made, and that intentions be subsequently updated in accordance with those decisions.

Despite such similarities, individual and coordinated group activities differ substantially in certain aspects relating to the intentions of the participants. This paper argues that the locus of greatest contrast lies not in the particular types of mental attitudes required, but rather in the *dynamics* of the characteristic intentions in such activities. The most fundamental distinctions arise from the multi-agent nature of the group decision-making processes required for carrying out multi-agent actions and the ways they differ from purely individual decision-making, means-ends reasoning, and problem solving. Unlike the simpler case of single-agent activity, in which all means-ends reasoning, decision making and intention updating is under the control of a single agent, in collaborative activity, responsibility for decision making is distributed. Furthermore, and importantly, collaborating agents must coordinate the updating of their group-related intentions.

These features of collaborative, group activity make a model of the dynamics of intention that adequately treats the inter-related processes of group decision-making and intention updating crucial for modeling collective intentionality. Although research has addressed the dynamics of intentions in individual activity [36, 35, 38, 37, 25, 42, 30, 29], the inter-related dynamics of decision making and intention updating in collaborative, multi-agent activity has received scant attention in the literature. Some exceptions include implementations [44, 34, inter alia] and work addressing the initial formation of a collaborative team [6, 24, 22, 31].

The rest of this paper is organized as follows. Section 2 briefly describes our prior work on the SharedPlans formalization of collaborative activity [13, 14] and the coordinated cultivation of intentions [21]. Section 3 presents the "coordinated cultivation of SharedPlans" model of the dynamics of intention in collaborative activity and an architecture for collaboration-capable computer agents based on that model. Section 4 shows how our model meets the requirements for collective intentionality identified by a variety of researchers and discussed above.

2 Background: SharedPlans and Coordinated Cultivation

The SharedPlans formalization [1] specifies the mental-state requirements of participants in a collaborative, group activity. In describing the formalization, we use the term "recipe" [33] to refer to a way of doing an action—i.e., the recipe for an action α is a set of actions (or "sub-acts") and constraints such that doing those actions under those constraints constitutes doing α. The formalization deploys two individual intentional attitudes to represent the commitments participants have to a joint activity, to their own actions in service of that joint activity, and to the actions of their co-participants in the activity: intending to do an action (intention-to) and intending that a proposition hold (intention-that). Intentions-that play a central role in realizing the commitments required for collective group activity and the cooperation and coordination that ensue from such commitments. Just as intentions to do actions are associated with means-ends reasoning, intentions-that are associated with a *cultivation* process [14].

In informal terms, the main constituents of the specification are that for agents to have a SharedPlan to do α, they must have the beliefs and intentions listed in Fig. 1. Through the cultivation process, each of the intentions-that in this formulation generates a set of decision problems. These decision problems include reaching agreements on the recipe to be used, the agents or subgroups to do the various subacts, and values for any action parameters, such as the time that actions are to be done or the particular resources that are to be used. The formalization also requires that the group of agents be committed to solving these decision problems together. In addition, it includes a set of axioms that embody various properties of intentions and interactions among them.

[1] The SharedPlans formalization of collaborative activity was initially presented at a 1986 workshop and is reported in a paper [17] in the volume resulting from that workshop. The theory was significantly generalized and revised in a subsequent set of papers [12, 13, 14, 28, 19]. We use "SharedPlans" in this paper to refer to the theory as refined through this progression.

(1) individual intentions that the group perform α;
(2) mutual belief of a (partial) recipe for α;
(3) individual or group plans for the sub-acts in the (partial) recipe;
(4) intentions that the selected agents or subgroups succeed in doing their sub-acts (for all sub-acts that have been assigned to some agent or group); and
(5) [in the case of a partial SharedPlan] individual intentions that the group complete the plan.

Fig. 1. The beliefs and intentions required for a SharedPlan

The coordinated cultivation of group intentions (CCGI) model [21] specifies additional important constraints on the cultivation process. [2] In particular, it addresses the need to ensure that the results of decisions made by the group in expanding a partial plan to a more complete one are actually reflected in the intentions of individual members of the group.

3 The Coordinated Cultivation of SharedPlans

The Coordinated Cultivation of SharedPlans (CCSP) model of the dynamics of intention in collaborative, group activity integrates the SharedPlans formalization of collaborative activity and the CCGI model of the coordinated cultivation of group-related intentions to provide a more uniform treatment of group decision-making and intention updating than either of these formulations on its own. This section presents an overview of the CCSP model, provides a detailed characterization of group decisions and their crucial role in collaborative activity, and discusses the formal specification of group decision-making mechanisms in the context of the CCSP model. Finally, it presents an architecture for a collaboration-capable agent based on the CCSP model.

3.1 Commitment to Decision Making in the CCSP Model

Collaborative group activity, like individual activity, involves an interlocking web of plans and intentions. For example, our plan to make dinner together might involve my intention to buy some food and your intention to make a salad. And in the group case, as in the individual case, plans may be incomplete and intentions under-specified. For example, our dinner-making plan might not yet specify who will make the main course or what will be served for dessert;

[2] Pasula [32] and Hadad [18] highlighted the need for participants in the group activity to coordinate their decisions about parameters of sub-acts such as the resources to be used and the times sub-acts were done.

and my intention to buy some food might be satisfied either by my going to a farmer's market or my going to a grocery store. In such cases, the agents must be committed to making the decisions that will enable them to eventually complete their plans and achieve their goals. However, both the nature of this commitment and its source are very different in purely single-agent activity and collaborative, multi-agent activity.

In purely single-agent activity, an intending agent must be committed to engaging in certain practical planning activity. This commitment derives directly from an axiomatic requirement that its plan be means-end coherent [2]. In contrast, in collaborative, multi-agent activity, each agent must be committed to participating in certain group planning processes. This commitment derives from a characteristic requirement that collaborating agents coordinate their updating of certain intentions, as described below.

As shown in clause (1) of Fig. 1, each agent in a SharedPlan is required to hold an intention whose content may be glossed as "that we do α." We refer to such intentions as *group-activity-related* (GAR) intentions. As in purely single-agent activity, the demand for means-end coherence requires that certain planning problems be addressed so that agents can update their GAR intentions and eventually complete their partial plan. However, for a collaborative plan to be coherent, it is also necessary that the GAR intentions held by the different participants all have the same content (e.g., "that we make dinner tonight"). We will refer to this characteristic as the "common-content property." To maintain the common-content property as their plans evolve, agents must coordinate the updating of their GAR intentions. For instance, in the dinner-making example, we must coordinate to avoid a situation where I update the content of my GAR intention to be "that we make dinner tonight at Martha's house", while you update the content of your GAR intention to be "that we make dinner tonight at Larry's house." Thus, the CCSP model includes the *coordinated cultivation requirement* (CCR), which is characteristic of collaborative activity:

> (CCR): The participants in a collaborative group activity update their corresponding GAR intentions only in accordance with decisions of the group.

The CCR explicitly prohibits the unilateral updating of GAR intentions—with the term "unilateral" applying not only to individual participants, but also to proper subgroups. Thus, to satisfy the CCR, each participant in a collaborative group activity must be committed to updating its GAR intention only in accordance with decisions of the group.

Under the assumption that the only way to generate group decisions is by agents participating in group decision-making processes, each participant in a collaborative activity must therefore be committed to participating in such group decision-making processes. In this way, the commitment to decision making that is inherent in any intention is, in the case of GAR intentions, transformed into a commitment to participate in group decision-making processes.

3.2 The Role of Group Decisions in the CCSP Model

In the context of collaborative activity, a group decision is an agreement that entails certain obligations on the participants. For instance, if a group GR decides to form a new collaborative team for doing some activity α, then each member of GR becomes obliged to adopt a new GAR intention ("that GR does α") and, furthermore, to constrain the updating of that GAR intention as stipulated by the CCR.[3] Other kinds of group decisions (e.g., to select a recipe, to allocate a task, or to bind a parameter) occur in the context of existing GAR intentions. Each group decision involves the making of a choice and entails an obligation on all of the participants to update their GAR intentions to reflect that choice. Thus, participants in collaborative activity are obliged to update their GAR intentions in accordance with, and—given the CCR—*only* in accordance with decisions of the group.

Task-allocation decisions are a special case in that they require not only the updating of GAR intentions, but also the adoption of new, subsidiary intentions. For example, whether allocating a task β to an agent G or subgroup SG, the rest of the agents in the group must adopt subsidiary intentions whose content may be glossed as "that G [or SG] is *able* to do β." In addition, the agent or subgroup chosen to do β must adopt subsidiary intentions aimed at doing β. In the case of allocating β to an agent G, G must adopt an intention to do β. Crucially, that intention is not subject to the CCR since G is the only relevant decision maker. In the case of allocating β to a subgroup SG, the members of SG must form a collaborative team whose goal is that they do β. In particular, each member of SG must adopt a subsidiary GAR intention ("that we [i.e., SG] do β"). That GAR intention is subject to the CCR: it must be updated only in accordance with decisions made by the subgroup SG.

The interactions between recipe-selection and task-allocation decisions typically lead to a complex hierarchy of subsidiary intentions that, together with related mutual beliefs, constitute the group's SharedPlan. A participant's view

[3] Bratman [3] has argued that collaboration can occur without an initiating group decision and its entailed obligations. The CCSP model's characterization of collaborative activity in terms of GAR intentions and the coordinated cultivation requirement is consistent with Bratman's view. However, in this paper, we focus on collaborative activity initiated by an explicit group decision.

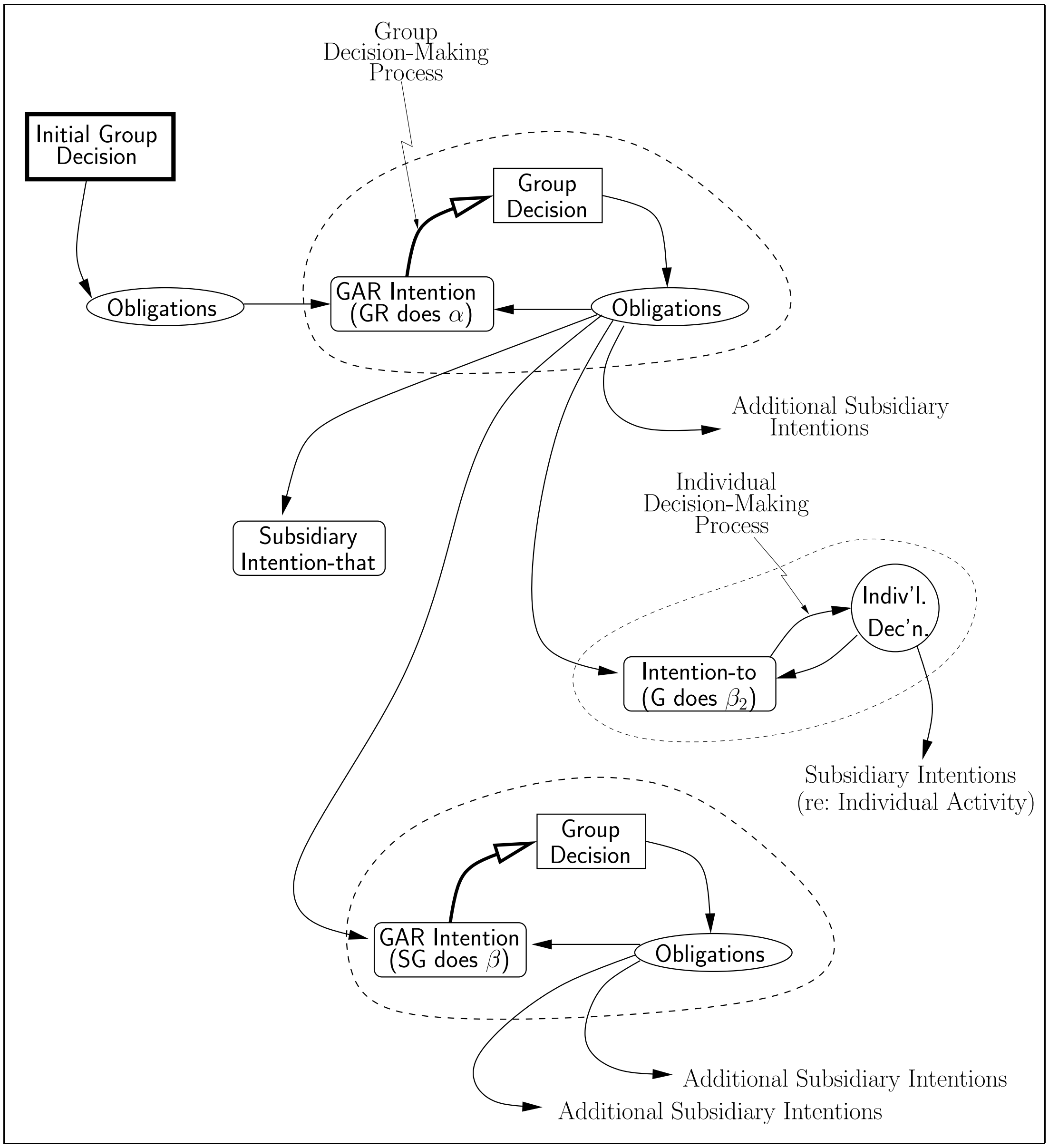

Fig. 2. A participant's view of the dynamics of a hierarchical SharedPlan

of the dynamics of a hierarchical SharedPlan is illustrated in Fig. 2. The process of incrementally completing that plan by moving from a set of GAR intentions, to a new group decision, to updated GAR intentions and newly adopted subsidiary intentions is recursive. The recursion stops when all intentions in the hierarchy have been fully specified and the single-agent actions associated with intentions at the bottom of the hierarchy have all been successfully executed (i.e., when the SharedPlan is complete).

As Fig. 2 shows, SharedPlans typically include subsidiary plans for constituent tasks that have been allocated to individual agents, in which case those agents are individually responsible for completing those sub-plans. We will not discuss further the incremental refinement of single-agent sub-plans because the focus of this paper is on the role of group decisions and the coordinated updating of GAR intentions in the incremental refinement of collaborative, multi-agent plans.

3.3 Characterizing Group Decisions

In the single-agent case, a decision is a mental action whereby an agent (internally) commits itself to some state of affairs. For example, my decision to see a movie tonight is a mental action whereby I commit myself to seeing a movie tonight. As a result of my decision, I now intend to see a movie tonight. Later on, I might decide to see a particular movie, in which case I update my intention. Thus, decisions typically involve adopting new intentions or modifying existing intentions.

The group case is analogous in that the end result of a group decision typically involves the adoption of new intentions or the modification of existing intentions by the participants. However, the analogy breaks down in that a group decision is not a mental action of the group—whatever that would mean.

Instead, a group decision has two aspects, one external (or social) and one internal (or mental).[4] In the CCSP model, a group decision-making mechanism generates a group decision in the social sense. It is in this sense that a group decision may be viewed simply as a proposition—such as, "The group has decided to see a movie tonight." Collaborating agents typically establish such group decisions through the use of a convention or mechanism. For example, you and I might establish our decision (or agreement) to see a movie tonight by making certain declarations and shaking hands (to "seal the deal"). Once established, such group decisions entail certain obligations which typically lead the participants to adopt new intentions or update their existing GAR intentions. The resulting, distributed network of internal commitments can then be viewed, in sum, as a group decision in the internal (or mental) sense.

The rest of this section focuses on the obligations entailed by group decisions. Figs. 3 and 4 identify the obligations entailed by each of a core class of group decisions that arise in collaborative group activity, including decisions to establish a new collaborative group or to coordinate the updating of related GAR intentions in an existing collaboration. The parameter-binding, recipe-selection and task-allocation decisions oblige the participants to update their relevant GAR intentions. In the figures, the particular updates are specified by giving the relevant portion of the content of those intentions both before

[4] Singh [41] has highlighted the importance of distinguishing the social and mental aspects of commitments.

- **Group Decision to Form a Collaboration:**

  Group Decision: Form team GR to do α collaboratively

  Obligations: New GAR int: that GR does α (constrained by CCR)

- **Group Decision to Bind a Parameter:**

  Prior GAR int: Unbound parameter p

  Group Decision: Bind parameter p to value v

  Obligations: Update GAR int:

  Replace all occurrences of p by v

- **Group Decision to Select a Recipe:**

  Prior GAR int: that GR does α

  Group Decision: Select recipe R_α (with sub-acts $\beta_1, \beta_2, \ldots, \beta_n$)

  Obligations: Update GAR int:

  that some agent or subgroup does β_1,

  that some agent or subgroup does β_2,

  ...

  that some agent or subgroup does β_n

Note 1: GAR int = group-activity-related intention

Note 2: Each agent in the group is obliged to perform the indicated updates to its GAR intention.

Fig. 3. The obligations entailed by group decisions (part 1)

and after the group decision. For task-allocation decisions, the subsidiary intentions that the participants are obliged to adopt are derived directly from the SharedPlans specifications.

It is important to note that the CCSP is not limited to the types of group decisions shown in Figs. 3 and 4, but rather it provides for the uniform treatment of a wide variety of group decisions. For example, agents seeking to coordinate their activities in the presence of temporal constraints may need to add new temporal constraints to their GAR intentions [20]. Such decisions would oblige the participants to update the content of their GAR intentions accordingly. Similarly, the obligations associated with a group's decision to delegate decision-making authority to an agent or subgroup may be handled by the same CCSP mechanism.

- **Group Decision to Allocate a Task to an Agent G:**

Prior GAR int:	that some agent or subgroup does β	
Group Decision:	Allocate sub-act β to agent G:	
Obligations:	Update GAR int:	that G does β
	Agent G:	Adopts int-to do β
	Other Agents:	Each adopts int-that G be able to do β

- **Group Decision to Allocate a Task to a Subgroup SG:**

Prior GAR int:	that some agent or subgroup does sub-act β	
Group Decision:	Allocate sub-act β to subgroup SG	
Obligations:	Update GAR int:	that SG does β
	Agents in SG:	Each adopts GAR int that SG does β (constrained by CCR)
	Other Agents:	Each adopts int-that SG be able to do β

Note 1: GAR int = group-activity-related intention

Note 2: In task-allocation scenarios, the intentions adopted by the agent(s) to whom the task is allocated and those adopted by the rest of the agents in the group are different.

Fig. 4. The obligations entailed by group decisions (part 2)

Finally, the CCSP model also accommodates another important aspect of group decision making, namely, that decisions about group activities are frequently interdependent and, thus, agents may need to combine multiple, related decisions into a single bundle. For example, the result of a lengthy multi-agent negotiation might be a bundle of interdependent task-allocation and parameter-binding decisions. In the CCSP model, the set of obligations entailed by a bundle of decisions is simply the union of the sets of obligations individually entailed by the decisions in that bundle.

3.4 Generating Group Decisions

People are quite adept at using informal mechanisms for establishing group decisions. For example, you and I might establish a group decision by nodding our heads and winking meaningfully. However, computer agents equipped with automated reasoning systems require formally specified mechanisms for establishing group decisions. Of course, even people employ formal mechanisms for generating group decisions when the stakes are high (e.g., as when a buyer, a seller, their attorneys, and a bank together agree to the conditions for the sale of a house).

In the CCSP model, group decision-making mechanisms (GDMMs) enable agents to reliably establish group decisions for the purpose of coordinating the updating of their GAR intentions. Different GDMMs may operate in very different ways. With some GDMMs, the parties to a group decision may always learn of the generation of a new decision simultaneously; with others, they may learn of a new decision at different times. However, all of the agents must eventually be able to verify that a group decision has in fact been established.

The CCSP model stipulates that the definition of a GDMM must specify: (1) the possible inputs an agent can make into the mechanism; (2) the conditions under which agents may make those inputs; (3) rules for determining which combinations of agent inputs establish group decisions; and (4) a method for making the new decision known to all the members of the group.

The original CCGI model [21] provides a general framework for formally specifying GDMMs satisfying the above criteria. The framework utilizes Dynamic Deontic Linear Time Temporal Logic (DDLTLB) [8, 9]. Agent inputs into a mechanism are in the form of declarative speech-acts [1, 40]. The conditions under which certain inputs are allowed, the rules for how certain combinations of inputs establish group decisions, and the resulting obligations are all easily expressed in DDLTLB. In that work, the use of the GDMM framework is illustrated by formally specifying a multi-agent voting protocol and proving its robustness under certain conditions.

In a related paper [22], we presented an auction-based mechanism that agents can use, when faced with some proposed group activity, to decide whether to initiate a collaborative effort. That mechanism, which can also be used to generate bundles of task-allocation decisions for a pre-existing collaborative effort, allows the participants to protect any private, pre-existing commitments they might have by including temporal constraints in their bids. Elsewhere [21], we have provided algorithms that agents can use to facilitate the bid-generation process needed to participate in such mechanisms.

The CCSP architecture for a collaboration-capable agent is based on the resource-bounded individual-agent architecture designed by Bratman, Israel, and Pollack [4]. This architecture, illustrated on the left side of Fig. 5 (i.e., all of the components lying outside the dashed line labeled "Components Related to Group Activities") addressed the ways in which means-ends reasoning and the weighing of alternative courses of action interact when agents are resource bounded. The architecture deals with the dynamics of intentions for individual activities, and the design embodies the constraining roles of plans and intentions. The means-ends reasoner, opportunity analyzer, filtering and deliberation components constitute the "practical reasoning system". Both the options generated by means-ends reasoning processes and new opportunities, whether internally or externally generated, are filtered for compatibility with existing plans and intentions. This filtering process eliminates options for future activity that would conflict with existing intentions, reflecting the focusing effect of plans (which is necessary because agents are resource-bounded), but it allows an agent to change its mind should a new, high-priority option arrive. As a result, an agent's partial plans are refined by means-ends reasoning, but in ways that are compatible with the full set of active intentions and plans. An agent is committed to what "it is doing" and the "characteristic roles" of those commitments are to drive means-ends reasoning and to constrain the set of options that the agent considers [2].

To extend this architecture to treat collaborative, group activities requires adding components to handle the interaction between group decision-making and intention updating in the group-activity context. The participants in a collaborative, group activity are not only resource-bounded, but also constrained in the decisions they are allowed to make unilaterally. Thus, corresponding to the single "deliberation" component on the "individual activity" side of the architecture, there is a set of components on the "group activity" side that handle the ways in which participants in the group activity deliberate together to augment incomplete plans and further specify GAR intentions. These components must embody the constraints imposed by the CCR in their handling of group decision making and the updating of GAR intentions. The components of Fig. 5 within the dashed line serve this purpose. To illustrate the roles of each component and their operation during the evolution of a typical SharedPlan, we will use the example of a group of musicians collaborating to perform a concert (or "gig").

The "Group-Activity Opportunity Analyzer" is responsible for generating new candidates for collaborative group activities by monitoring incoming communication and the database of beliefs (including those arising from perception of the external environment). For example, an agent might learn of a new

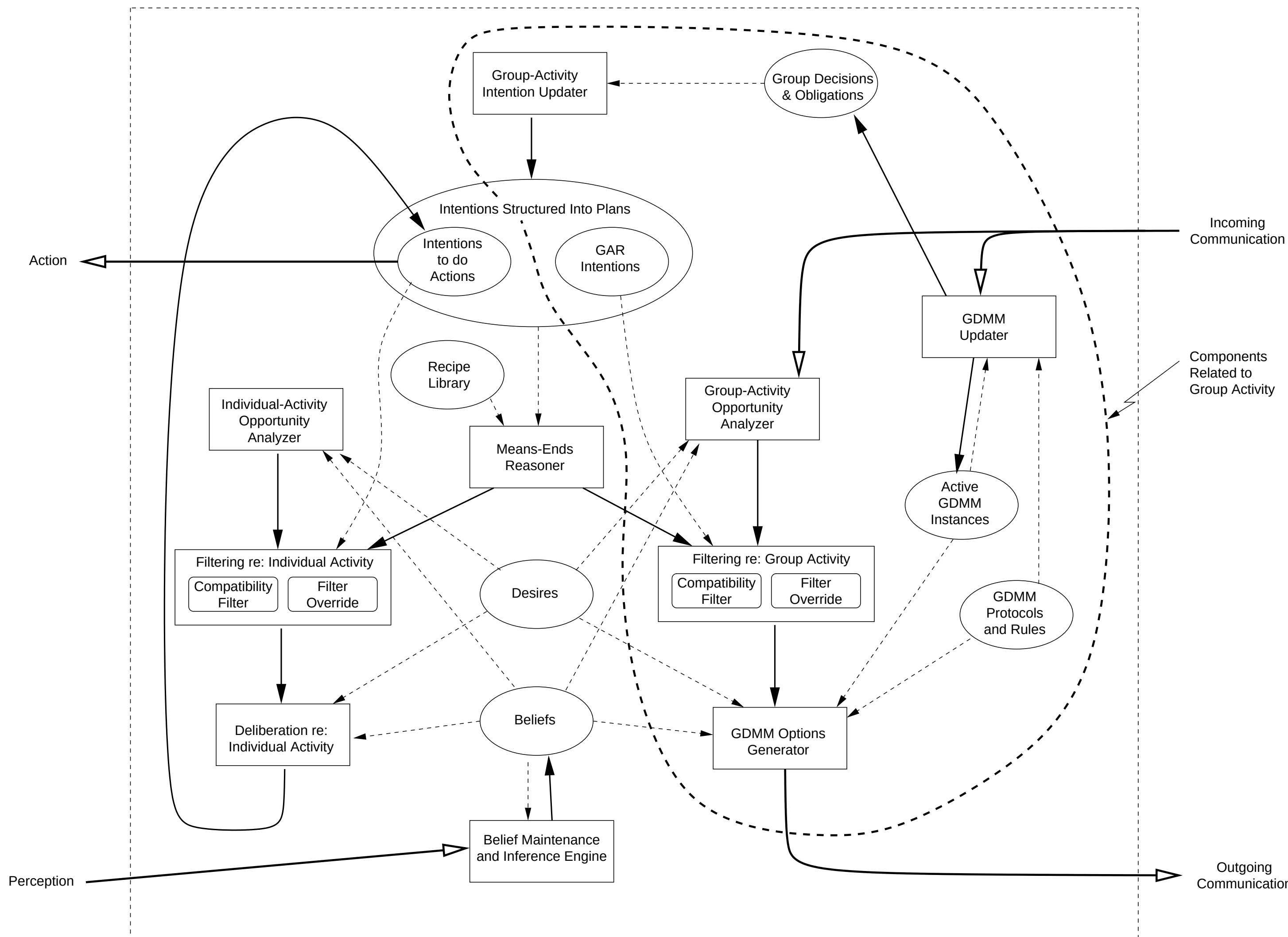

Fig. 5. The CCSP Architecture for a Collaboration-Capable Agent

opportunity for playing a gig from an incoming email message. When such an opportunity (or option) is generated, it is sent to the "Filtering re: Group Activities" module which determines whether that option would be compatible with the other activities that the agent is already engaged in (which are represented in the database labeled "Intentions Structured into Plans"). For instance, if the agent is already scheduled to play a gig on Friday night, an opportunity to play another gig that same night on the other side of the country would typically be discarded. However, the "Filter Override" module can allow certain high-priority opportunities to survive the filtering process even if they might conflict with an agent's pre-existing commitments.

The group compatibility and over-ride filters differ from those of individual activities in possibly taking into account the preferences and obligations not only of the agent, but also of other participants or the group as a whole. The process of intention reconciliation is more complex in the context of group activities because it requires the weighing of trade-offs between individual and group good [15, 16]. For instance, the agent might determine that the needs of the group are sufficiently great that it should reconsider the individual intention that led to a compatibility conflict.

If the new opportunity survives the filtering process, the "GDMM Options Generator" uses the "GDMM Protocols and Rules" database to select a particular GDMM and initiate a group decision-making process using that GDMM. For example, the agent might decide to use a *Unanimous Consent* GDMM and send an email message to its fellow musicians proposing that they agree to play this new gig on Friday night. The newly initiated GDMM instance is recorded in the "Active GDMM Instances" database.

Replies from other agents ("incoming communication") are processed by the "GDMM Updater" which updates the status of the GDMM instance in the "Active GDMM Instances" database. If all the other agents agree, then the agent who originated the proposal annouces the group decision ("outgoing communication") and records it in its own "Group Decisions and Obligations" database. Suggestions from other agents are handled in an analogous manner.

Group decisions to engage in new group activities lead the "Group-Activity Intention Updater" to create a new GAR intention and enter it into the database of "Intentions Structured Into Plans". When all agents in the group have adopted corresponding GAR intentions, each constrained by the CCR, the group's SharedPlan is initialized.

The new GAR intention requires that certain decisions be made (e.g., how to get to the gig, what equipment to bring). As in the single-agent case, the "Means-Ends Reasoner" generates potential solutions to such decision problems. However, in the multi-agent case, these potential solutions may involve actions to be performed with or by other agents (e.g., Bill and I borrow a car or Charlie brings the high-powered equipment). Each potential solution must be filtered to ensure that any action to be performed by this agent, whether individually or with others, is compatible with its existing intentions and plans. Each potential solution that survives this filter is sent to the "GDMM Options Generator", which determines possible communications the agent might make to the rest of the group (in the context of a GDMM) to get them to consider that potential solution.

When the agent learns of new group decisions, the "GDMM Updater" records the new decision (e.g., to rent a van) which leads the "Group Activity Intention Updater" to enter new intentions or update existing intentions in the database of "Intentions Structured into Plans". In the case of a task-allocation decision, say, "Bob to drive van", the updater enters new, subsidiary intentions concerning the driving of the van. If "Bob" refers to this agent, then the intention is an intention to drive the van, which may lead to subsidiary single-agent planning activity (on the left side of the diagram). If "Bob" refers to a different agent, then the new intention is an intention that "Bob" be able to drive the van. (Because the cultivation of this latter form of intention is not the focus of this paper, such intentions are not shown in the database of "Intentions Structured into Plans".)

Thus, the components in the CCSP agent architecture provide the functionality needed to facilitate an agent's participation in group decision-making processes and, more generally, collaborative activity.

4 Discussion

Tuomela [45] argues that collaborating agents employ a we-mode of intending that is different from and not reducible to the ordinary I-mode of intending. For example, he claims that in the group case, "the goal state or event comes about due to the collective effort by, or at least under the collective guidance of, the group members." Elsewhere [47] he stipulates that we-intenders must be disposed to engage in certain forms of practical reasoning, for example, to derive their individual contributions to the collaborative activity; and that collaborating agents require a "socially-existing authority system" (or group-will formation system) and must be motivated to use it. The CCSP's account of GAR intentions constrained by the CCR captures all of these requirements. Furthermore, it highlights the CCR as the source of an agent's motivation to participate in group decision-making mechanisms, provides criteria for formally specifying such mechanisms, and specifies the particular obligations that are entailed by the group decisions generated by such mechanisms.

Bratman [3] presents a model of Shared Cooperative Activity that employs ordinary I-mode intentions, but with a highly specialized content: "that we J in accordance with and because of [our intentions] and meshing subplans of [our intentions]." In contrast, the CCSP model employs GAR intentions with a much simpler content (e.g., "that we do α"), but with their updating constrained by the CCR. This separation of the content of an intention from constraints on how that content can be updated enables the CCSP agent architecture to employ essentially the same means-ends reasoner used in the single-agent case, while adding separate components (e.g., the "GDMM Options Generator" and the "GDMM Updater") to handle the deliberation and communication processing required to manage the agent's participation in group decision-making mechanisms.

Searle [39] argues that it is important to be able to distinguish, for example, the case of a set of agents each independently intending to run to a common point from the case of a group of agents collectively intending to run to that same point. The CCSP model makes this distinction by requiring, in the collective case, that each agent have a GAR intention ("that we run to the common point") and a derivative intention ("to run to the common point") that is correctly linked to the GAR intention. In addition, each agent's GAR intention is subject to the CCR, which ensures the coherence of the collective intention. Searle also argues that it is important to be able to distinguish cases where the content of the collective intentionality ("we are running a pass play") is different from the content of derivative intentions ("I am blocking the defensive end"), which the CCSP model also accommodates with the same mechanism: I intend "that we run a pass play" and I have a subsidiary intention "to block the defensive end."

Searle also presents two versions of a Business School scenario that he claims form a counter-example to any account of collective intentionality based solely on ordinary I-intentions. In the scenario, graduates of the Business School have learned that they may help humanity by pursuing their own selfish interests. In the first version, each graduate "intends to pursue his own selfish interests without reference to anybody else", whereas in the second version, the graduates "form a pact to the effect that they will all go out together and help humanity by way of each pursuing his own selfish interests." These cases are clearly distinguished in the CCSP model by the presence or lack of corresponding GAR intentions ("that we help humanity by each pursuing our own selfish interests"). In addition, the SharedPlans formalization has mechanisms that handle additional aspects of this distinction [11].

Finally, Searle argues that "the notion of a we-intention, of collective intentionality, implies the notion of *cooperation*", the following crucial elements of which are captured by the CCSP model: (1) an agent's motivation to participate in group decision-making processes; (2) the group's coordinated updating of corresponding GAR intentions; (3) the persistence of the common-content property and, thus, the coherence of a group's evolving SharedPlan; and (4) the commitment of agents to avoid interfering with the efforts of their fellow participants. In addition, hierarchical SharedPlans capture the relationships between GAR intentions and subsidiary plans, even if the higher-level intentions are cooperative while the lower-level intentions are competitive (as happens, for example, in organized sports).

Gilbert [10] argues that any model of collective intentionality (or "shared intention") must account for the following features: (1) that shared intentions entail certain obligations (e.g., "not to act contrary to the shared intention" or "to promote the fulfillment of the shared intention"), as well as corresponding rights and entitlements; (2) that participants in a shared intention "are not in a position to remove its constraints unilaterally"; and (3) "that there could be a shared intention to do such-and-such though none of the participants personally intend to conform their behavior to the shared intention." The CCSP model accounts for the first feature in that group decisions, like joint commitments, entail certain obligations (in particular, to update existing GAR intentions or to adopt related subsidiary intentions). The CCSP model's CCR constraint against the unilateral updating of GAR intentions captures the second feature. The third feature has debatable merit and we have not attempted to capture it in the CCSP model.

This section has examined the requirements for an account of collective intentionality as put forth by several researchers, who argue that their requirements can only be met by introducing a new kind of intention (e.g., a we-intention) or by adding complex constraints into the content of intentions. However, the CCSP model presented in this paper meets all of these requirements without introducing any new kind of intention or making the content of intentions more complex. Furthermore, by explicating the dynamic and interrelated processes of group decision making and intention updating, the CCSP model fills an important gap in existing accounts of collaborative intentionality.

References

[1] J.L. Austin. *How to do things with words.* Harvard University Press, Cambridge, MA, 1962.

[2] Michael E. Bratman. *Intentions, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, 1997.

[3] Michael E. Bratman. *Faces of Intention: Selected Essays on Intention and Agency.* Cambridge University Press, 1999.

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

[5] Cristiano Castelfranchi. Commitments: From individual intentions to groups and organizations. In Lesser [26], pages 41–48.

[6] Philip R. Cohen, Hector J. Levesque, and Ira Smith. On team formation. In G. Holmstrom-Hintikka and R. Tuomela, editors, *Contemporary Action Theory*, volume II, pages 87–114. Kluwer Academic Publishers, Netherlands, 1997.

[7] Philip R. Cohen, J. Morgan, and Martha E. Pollack, editors. *Intentions in Communication*. The MIT Press, Cambridge, Massachusetts, 1990.

[8] Frank Dignum and Hans Weigand. Communication and deontic logic. In R. Wieringa and R. Feenstra, editors, *Information Systems, Correctness and Reusability*, pages 242–260. World Scientific, 1995.

[9] Frank Dignum and Hans Weigand. Modelling communication between cooperative systems. In *Conference on Advanced Information Systems Engineering*, pages 140–153, 1995.

[10] Margaret Gilbert. *Sociality and Responsibility*. Rowman & Littlefield Publishers, Inc., New York, 2000.

[11] Barbara J. Grosz. The contexts of collaboration. *Cognition, Agency and Rationality: Proceedings of the Fifth International Colloquium on Cognitive Science (ICCS-97)*, pages 175–188, 1999.

[12] Barbara J. Grosz and Sarit Kraus. Collaborative plans for group activities. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI-93)*, San Mateo, CA, USA, 1993. Morgan Kaufmann.

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

[14] Barbara J. Grosz and Sarit Kraus. The evolution of SharedPlans. In Michael Wooldridge and Anand Rao, editors, *Foundations of Rational Agency*, number 14 in Applied Logic Series, pages 227–262. Kluwer Academic Publishers, The Netherlands, 1999.

[15] Barbara J. Grosz, Sarit Kraus, David G. Sullivan, and Sanmay Das. The influence of social norms and social consciousness on intention reconciliation. *Artificial Intelligence*, 142:147–177, 2002.

[16] Barbara J. Grosz, Sarit Kraus, and Shavit Talman. The influence of social dependencies on decision-making: Initial investigations with a new game. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2004)*. ACM Press, 2004.

[17] Barbara J. Grosz and Candace L. Sidner. Plans for discourse. In Cohen et al. [7], chapter 20.

[18] Meirav Hadad. *Combining Cooperative Planning and Temporal Reasoning in Dynamic Multi-Agent Systems*. PhD thesis, Bar Ilan University, Ramat Gan, Israel, 2002.

[19] Luke Hunsberger. Making SharedPlans more concise and easier to reason about. In *Agent Architectures, Theories and Languages V*, volume 1555 of *Lecture Notes in Artificial Intelligence*, pages 81–98. Springer-Verlag,

1999.

[20] Luke Hunsberger. Algorithms for a temporal decoupling problem in multi-agent planning. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-2002)*, 2002.

[21] Luke Hunsberger. *Group Decision Making and Temporal Reasoning*. PhD thesis, Harvard University, 2002. Available as Harvard Technical Report TR-05-02.

[22] Luke Hunsberger and Barbara J. Grosz. A combinatorial auction for collaborative planning. In *Fourth International Conference on MultiAgent Systems (ICMAS-2000)*, pages 151–158. IEEE Computer Society, 2000.

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

[24] D. Kinny, M. Ljungberg, A.S. Rao, E. Sonenberg, G. Tidhar, and E. Werner. Planned team activity. In C. Castelfranchi and E. Werner, editors, *Artificial Social Systems*, Lecture Notes in Artificial Intelligence. Springer Verlag, Amsterdam, 1994.

[25] Amy L. Lansky and Michael Georgeff. Procedural knowledge. Technical Report 411, SRI International, January 1987.

[26] Victor Lesser, editor. *First International Conference on Multi-Agent Systems (ICMAS-95)*. The MIT Press, 1995.

[27] Hector J. Levesque, Philip R. Cohen, and Jose H. T. Nunes. On acting together. In *Seventh National Conference on Artificial Intelligence*, volume 1, pages 94–99. AAAI Press/MIT Press, 1990.

[28] Karen E. Lochbaum. *Using Collaborative Plans to Model the Intentional Structure of Discourse*. PhD thesis, Harvard University, October 1994.

[29] J.-J. Ch. Meyer, W. van der Hoek, and B. van Linder. A logical approach to the dynamics of commitments. *Artificial Intelligence*, 113, 1999.

[30] Charles L. Ortiz, Jr. Introspective and elaborative processes in rational agents. *Journal of the Annals of Mathematics and Artificial Intelligence*, 1999.

[31] Pietro Panzarasa, Nicholas Jennings, and Timothy J. Norman. Formalising collaborative decision-making and practical reasoning inmulti-agent systems. *Journal of Logic and Computation*, 11(6):1–63, 2001.

[32] Hanna Pasula. Design of a collaborative planning system. Seniors Honors Thesis, Harvard University, 1996.

[33] Martha E. Pollack. Plans as complex mental attitudes. In Cohen et al. [7], chapter 5.

[34] David V. Pynadath, Milind Tambe, Nicolas Chauvat, and Lawrence Cavedon. Toward team-oriented programming. In Nicholas R. Jennings and Yves Lesperance, editors, *Intelligent Agents VI, Agent Theories, Architectures, and Languages*, volume 1757 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Berlin Heidelberg, 2000.

[35] Anand S. Rao and Michael P. Georgeff. Modeling rational agents within a BDI-architecture. In *Proceedings of the 2nd Conference on Principles of*

Knowledge Representation and Reasoning (KR'91), pages 473–484, San Mateo, CA, 1991. Kaufman.

[36] Anand S. Rao and Michael P. Georgeff. An abstract architecture for rational agents. In *Proceedings of the 3rd Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 439–449. Morgan Kaufman Publishers, Inc., 1992.

[37] Anand S. Rao and Michael P. Georgeff. BDI agents: From theory to practice. In Lesser [26], pages 312–319.

[38] Anand S. Rao and Michael P. Georgeff. The semantics of intention maintenance for rational agents. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)*, pages 704–710, 1995.

[39] John R. Searle. Collective intentions and action. In Cohen et al. [7].

[40] J.R. Searle. *The Construction of Social Reality*. Allen Lane, London, 1995.

[41] Munindar P. Singh. Agent communication languages: Rethinking the principles. *IEEE Computer*, 31(12):40–47, December 1998.

[42] SRI International. *Procedural Reasoning System User's Guide*.

[43] Milind Tambe. Agent architectures for flexible, practical teamwork. In *Fourteenth National Conference on Artificial Intelligence*, pages 22–28, 1997.

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

[45] Raimo Tuomela. Cooperation and the we-perspective. Paper for the conference *Rationality and Commitment*, St. Gallen, May 2004.

[46] Raimo Tuomela. Joint action. Invited paper for the Workshop on Holistic Epistemology and Theory of Action, University of Leipzig, June 24-27, 2004.

[47] Raimo Tuomela. *The Importance of Us: A Philosophical Study of Basic Social Notions*. Stanford University Press, Stanford, CA, 1995.