Collaborative Plans for Group Activities*

Barbara Grosz
Division of Applied Sciences
Harvard University
Cambridge, MA 02138 USA
grosz@das.harvard.edu

Sarit Kraus
Dept. of Mathematics and Computer Science
Bar Ilan University,
Ramat Gan, 52900 Israel
sarit@bimacs.cs.biu.ac.il

Abstract

The original formulation of SharedPlans [Grosz and Sidner, 1990] was developed to provide a model of collaborative planning in which it was not necessary for one agent to have intentions toward an act of a different agent. This formulation provided for two agents to coordinate their activities without introducing any notion of jointly held intentions (or, 'we-intentions'). However, it only treated activities that directly decomposed into single agents actions. In this paper we provide a revised and expanded version of SharedPlans that accommodates actions involving groups of agents as well as complex actions that decompose into multi-agent actions. The new definitions also allow for contracting out certain actions, and provide a model with the features required in Bratman's account of shared cooperative activity [Bratman, 1992]. A reformulation of the model of individual plans that meshes with the definition of SharedPlans is also provided.

1 Introduction

Collaboration in planning and acting is an essential ingredient of multi-agent cooperative problem solving. In this paper we present a model of collaborative planning that supports cooperative problem solving by teams consisting of humans and computer systems. The model deals more completely with collaboration than previous theories did in meeting two criteria. First, collaborative planning and activity cannot be analyzed simply in terms of the plans of individual agents, but require an integrated treatment of the beliefs and intentions of the different collaborating agents. Second, collaborative planning is a refinement process; a partial plan description is modified over the course of planning by the multiple collaborating agents. This model grew out of an attempt to provide an adequate treatment of the intentional component of discourse structure [Grosz and Sidner, 1986]. However, many multi-agent situations require that agents have an ability to plan and act together; merely avoiding conflicting actions or situations is not sufficient. Thus, the model is applicable not only to natural language processing, but also to the general problem of the design of computer-based collaborating agents.

The original formulation of the SharedPlan model of collaborative planning [Grosz and Sidner, 1990] extended Pollack's mental state model of plans [Pollack, 1990] to the situation in which two agents jointly have a plan to perform some action requiring actions by both agents. Pollack s definition of the individual plan of an individual agent to do an action a includes four constituent mental attitudes: (1) belief that performance of certain actions β_i would entail performance of α; we will refer to the β_i as constituting "a recipe for α;" (2) belief that the agent could perform each of the β_i; (3) intentions to do each of the β_i; (4) an intention to do *a* by doing the β_i. To define SharedPlans, Grosz and Sidner modified these components to provide a specification of the set of beliefs and intentions required for collaborative action. In subsequent work [Lochbaum *et* al., 1990; Lochbaum, 1991], algorithms were provided for constructing and augmenting SharedPlans in the context of a dialogue.

Although this formulation overcame several problems of previous models of planning for discourse (e.g. the treatment of intentions of one agent toward another agents actions in applications of speech act theory [Allen and Perrault, 1980]), it had several problems that emerged when we attempted to apply it [Lochbaum *et* al., 1990; Lochbaum, 1991]. First, the original model presumed that every multi-agent action decomposed directly into single agent actions, a similar assumption underlies several alternative models (e.g. [Cohen and Levesque, 1990]). As a result, the model did not adequately provide for complex activities involving joint activity at multiple levels or for meshing of individual plans for individual action with collaborative plans for joint activity. Second, the model did not account for the commitment of an agent to the success of a collaborative partner's actions. This omission combined with the first

---

* We thank Joyce Friedman for many provoking questions, and Karen Lochbaum for the same and for helpful comments on many drafts. This research was initiated when the first author was a Harold Perlman Visiting Professor, Hebrew University, Jerusalem. Partial support for the first author was provided by U S WEST Advanced Technologies. The second author was supported by NSF Grant No. IRI-9123460.

so that the model accepted some plans as collaborative that were not.1 Third, the agents who undertake the development of a collaborative plan often do not know a complete recipe for accomplishing their joint action; the model did not provide a sufficient means of describing the mental state of agents in this situation.2 Each of these problems is addressed in this paper.

Collaborative activity must rest eventually on the actions of individual agents; thus, SharedPlans must include as constituents at some levels the individual plans of individual agents. But these individual plans may be more complex than those accounted for by Pollack's formulation in two ways: the recipes require different types of relations [Balkanski, 1990]3, and an agent may not initially know a complete recipe. Hence, we also provide a revised definition of the plans of an individual agent.

We begin the description of the revised model with an overview of different types of intentional attitudes that play a role in collaborative planning. Then we provide a definition of individual plans that accommodates more complex recipes but still requires complete knowledge of the recipe. To describe these full individual plans requires a specification of certain properties of intention, so we define the Int.To operator at this point. Next, we show how to ease the complete knowledge requirement, yielding a definition of partial individual plan. We then define SharedPlans recursively in terms of full and partial SharedPlans. A full SharedPlan is the collaborative correlate of a full individual plan and includes full individual plans as constituents. To describe interactions among the intentions of the different agents requires that we introduce the notion of an agent intending that some proposition hold so we explain the Int. Th operator in this section. Finally, we provide a definition of partial ShartdPlan. At each stage we discuss those aspects of the resulting theory that address the problems described above to provide a more adequate model of collaborative activity.

2 Attitudes of Intention

The definitions of individual and SharedPlans will use a first-order logic augmented with several modal operators. We introduce four different intention operators. Two of these, Int. To and Int. Th, represent intentions that have been adopted by an agent. The other two, Pot.Int. To and Pot. Int. Th, are variations of the first two that are used to represent potential intentions, i.e. intentions an agent is considering adopting but to which it is not yet committed. Int. To and Pot.Int. To are action-directed whereas Int. Th and Pot.Int. Th are proposition-directed.

An Int. To commits an agent to means-ends reasoning [Bratman, 1987] whereas an Int.Th does not directly engender such behavior. Int.Th's form the basis for meshing subplans, helping one's collaborator, and coordinating status updates [Cohen and Leveque, 1991]. An Int. Th may, however, lead to adoption of an Int. To and thus indirectly to means-ends reasoning. Potential intentions are used to account for an agent's need to weigh different possible courses of actions [Bratman et al/., 1988]; they typically arise in the course of means-ends reasoning. Attitudes of Pot.Int. To stem from an agent's deliberations about how to do some action it is committed to perform. Pot.Int.Th's arise in the course of collaborative planning and are needed to insure that agents' individual plans mesh correctly [Bratman, 1992].

The difference among these operators can be illustrated with an example we will use throughout the paper. Two agents, Jan and Sandy, have agreed to make dinner together. Their collaborative plan consists of Jan making an appetizer, Sandy the main course, and the two of them together making the dessert. Their SharedPlan to make dinner includes Jan having an intention to [Int. To] make the appetizer (and an individual plan for doing so), Sandy having an intention to [Int.To] make the main course (and an individual plan for doing so), and their having a SharedPlan to make the dessert. The SharedPlan for making dinner also includes Sandy's intention that [Int.Th] Jan 4can make' the appetizer, and Jan's intention that [Int.Th] Sandy 'can make' the main course.

Jan may have decided to make cheese puffs for the appetizer, but not yet have chosen a recipe for doing so. If so, his individual plan will be partial. It will include an Int. To get a recipe for cheese puffs and a full individual plan for doing so.4 In addition, he believes that he can perform all of the actions in the recipe once he gets it. As he determines the recipe and thus the actions he needs to perform (according to that recipe), he adopts potential intentions to [Pot.Int. To] perform these actions. The potential intentions will become part of a deliberation process [Bratman et al/., 1988] and through that process may become Int. TVs.

In the definitions that follow, *Int.To*$(G,\alpha,T_i,T_\alpha,C_\alpha)$ represents the agent G's intention at time T_i to do action α at time T_α in the context C_α, (the role of context will be discussed later); *Int.Th*$(G,prop,T_i,T_{prop},C_{prop})$ represents an agent G's intention at time T_i that a certain proposition *prop* hold at time T_{prop} in the context C_{prop}. Adoption of an intention of either sort commits an agent to not adopting conflicting intentions, and may constrain replanning in case of failure [Bratman, 1987].

The definitions below make use of several operators that we can only define informally in this paper. GTD$(G_1,\gamma,G_2,\alpha,T_\alpha,T_\gamma)$ (read "get to do") holds if G_1's doing γ at T_γ will cause G_2 to intend to do α at T_α or (if G_2 is a group of agents) have a SharedPlan to do α. The operator CBA (can bring about), associated with an agent, an action, a recipe, and a time, means

---

1For example, joint activity like that in Searle's MBA counterexample [Searle, 1990], but involving actions that decomposed at multiple levels would have been inaccurately characterized.

2The notion of a partial SharedPlan, SharedPlan*, was intended to represent this kind of partiality, but was never specified in any detail.

3In Pollack's simple plans the β, and a were related only by the action relation of generation.

[4] The requirement that the agent have a full plan for getting the recipe may seem too strong, but it is necessary to avoid an infinite recursion.

that the agent either can do the action at the time using the recipe[5] or can get another agent to do it; CBAG is the analogous group operator. $Done(G, \alpha, T_\alpha, R_\alpha)$ holds when G (either a group or a single agent) has done a over time interval T_α using the recipe R_α.

In addition, we will use R_α to denote a recipe for α, i.e. a specification of a group of actions, which we will refer to as β_i, the doing of which under appropriate constraints, ρ_i, will constitute performance of α [Pollack, 1990; Balkanski, 1990; Lochbaum *et* a/., 1990]. A recipe may include uninstantiated variables (e.g. for the agent or time of an action) and constraints on these variables. We assume each agent has a library of recipes that it collects and updates over time. Agents' libraries may differ, and the successful completion of a SharedPlan may require integrating recipes from different libraries, i.e. from different agents.

3 Individual Plans and Intending To

3.1 Full Individual Plans

The definition of a full individual plan, FIP, is given in Figure 1. It specifies those conditions under which an individual agent G can be said to have a plan P, at time T_p to do action α at time T_α using recipe R_α in the context C_α. Full plans are distinguished by the requirement that the agent know a complete recipe for doing the action; as a result R_α is a parameter of the operator. A full individual plan for a represents the mental state of an agent after he has completely determined the means by which he will do α and has full-fledged intentions to do the actions in R_α. Most, typically an agent will not have a full plan until after he has done some or all of the actions in R_α; thus, most often agents have only partial plans. However, it is useful to understand the limiting case of the full plan before examining the partial version.

We will illustrate the FIP by showing its use in describing Sandy's individual plan for making the main course in the meals example. According to Clause (0), Sandy believes that a particular recipe, say his mother's recipe for lasagne, is a recipe for making a main course The remaining clauses provide a specification of certain attitudes Sandy must hold with respect to the individual constituents of this recipe. In particular, for each action β_i in the recipe (e.g. making noodles, preparing sauce), he must either intend to do the action (1) or believe that he can get someone else to do the action (2). We will refer to the first case as the "core case1' of the individual plan, and the second case as the "contracting case."

In the core case, the agent must believe either that the action is basic level and he can bring it about (la), or that he has a recipe that will enable him to do the action and a full individual plan to carry out the actions in that recipe (lb). The *Int. To* operator in Clause (1) includes a context parameter, $C_{\beta_i/\alpha}$, that is used in any replanning involving β_i. For the purposes of this paper, the important element of the context encoded in this parameter is a representation (using the Contributes relation [Lochbaum *et al.*, 1990]) of the fact that G is doing β_i at time T_{β_i} as part of G's performance of action α at time T_α; this element is constructed recursively as an agent chooses recipes and constructs plans for the actions in them.

$\text{FIP}(P, G, \alpha, T_p, T_\alpha, R_\alpha, C_\alpha)$

(0) $R_\alpha = \{\beta_i, \rho_i\} \wedge \text{BEL}(G, R_\alpha \in Recipes(\alpha), T_p) \wedge$
$(\forall \beta_i \exists T_{\beta_i})\{$

(1) $[\text{Int.To}(G, \beta_i, T_p, T_{\beta_i}, C_{\beta_i/\alpha}) \wedge$

(1a) $([basic.level(\beta_i) \wedge$
$\text{BEL}(G, \text{CBA}(G, \beta_i, R_{Empty}, T_{\beta_i}), T_p)] \vee$

(1b) $[\neg basic.level(\beta_i) \wedge (\exists P_{\beta_i}, R_{\beta_i})$
$(\text{BEL}(G, \text{CBA}(G, \beta_i, R_{\beta_i}, T_{\beta_i}), T_p) \wedge$
$\text{FIP}(P_{\beta_i}, G, \beta_i, T_p, T_{\beta_i}, R_{\beta_i}, C_{\beta_i/\alpha}))])] \quad \bigvee$

(2) $[\exists (G_1, \gamma, T_\gamma)$

(2a) $[\text{BEL}(G, \text{GTD}(G, \gamma, G_1, \beta_i, T_{\beta_i}, T_\gamma), T_p) \wedge$

(2b) $\text{Int.To}(G, \gamma, T_p, T_\gamma, C_{\gamma/\beta_i/\alpha}) \wedge$

(2c) $\text{Int.Th}(G, (\exists R_{\beta_i})$
$\text{CBA}(G_1, \beta_i, R_{\beta_i}, T_{\beta_i}), T_p, T_{\beta_i}, C_{\beta_i/\alpha}) \wedge$

(2d1) $([basic.level(\gamma) \wedge$
$\text{BEL}(G, \text{CBA}(G, \gamma, R_{Empty}, T_\gamma), T_p)] \vee$

(2d2) $[\neg basic.level(\gamma) \wedge \exists (R_\gamma, P_\gamma)$
$(\text{BEL}(G, \text{CBA}(G, \gamma, R_\gamma, T_\gamma), T_p) \wedge$
$\text{FIP}(P_\gamma, G, \gamma, T_p, T_\gamma, R_\gamma, C_{\gamma/\beta_i/\alpha}))])]]\}$

Figure 1: Full Individual Plan

$\text{Int.To}(G, \alpha, T_i, T_\alpha, C_\alpha)$

(1) $[basic.level(\alpha)] \vee$

(2) $[\neg basic.level(\alpha) \wedge$

(2a) $[(\exists P, R_\alpha)[\text{FIP}(P, G, \alpha, T_i, T_\alpha, R_\alpha, C_\alpha) \wedge$
$\text{BEL}(G, \text{CBA}(G, \alpha, R_\alpha, T_\alpha), T_i)]] \quad \bigvee$

(2b) $[(\exists P, T_{ach})[\text{PIP}(P, G, \alpha, T_p, T_\alpha, C_\alpha) \wedge$
$\text{Int.To}(G, Achieve((\exists R_\alpha)\text{FIP}(P, G, \alpha, T_p, T_\alpha,$
$R_\alpha, C_\alpha)), T_i, T_{ach}, C_{ach/\alpha})]]]$

Figure 2: The definition of *Int. To*

The definition of intending to do an action is given in Figure 2.[6] Within a full individual plan, Clause (2b) of this definition, does not apply. An *Int. To* do an action requires either (1) that the action be basic (in which case the agent is committed to doing the act [Bratman, 1987], something that we do not formalize here) or (2a) that the agent have a full individual plan to do the action. Thus, a full individual plan to do α requires intentions to do each of the actions, β_i, involved in a recipe for α which in turn require full individual plans for each of these β_i, until we get to basic level acts.[7]

The contracting case for accomplishing β_i has not been discussed in previous work on multi-agent plans, but clearly arises in many cases. For example, someone who has a plan to renovate his house might include in that plan contracting out certain subtasks, e.g. the refinishing of the floors. To contract out an action, the agent must believe there is some action γ that he can use to get

---

*If the action is basic level, this reduces to Pollack's EXEC operator. The connective $\vee$ that we use in all the definitions is really exclusive or (XOR).

[6] Space constraints preclude inclusion of the basic axiom for avoiding the adoption of conflicting intentions as well as other related axioms.

[7] This is the case that Pollack's definition covers, but it does so only for recipes in which actions are related by generation.

another agent to do β_i (Clause (2a) of Figure 1). Furthermore, he must intend to do γ (2b), believe he can bring about γ, and if γ is not basic-level (2d2), have a full plan for doing γ. Clause (2c) represents agent G's commitment to the success of the contractor in doing β_i. Notice that contracting is not in and of itself collaborative [Kraus, 1993]; Clause (2c) is one way: the contractor does not have the kinds of Int.Th's we will require in the SharedPlan definitions later; only the contracting agent does. *Int.Th* will be discussed in more detail under SharedPlans (Section 4.1), where there is such reciprocity.

In the meals example, Sandy may decide that the most efficient way to make the lasagne is to get Tony to make the sauce and sell it to him. In this case, Sandy's individual plan would include an action γ (e.g. making a sales agreement) that resulted in Tony's providing the sauce. For Sandy's plan to be complete, Sandy must believe that the action γ that he will do either is a basic-level action that he is able to do or is an action for which he knows a recipe and for which he has a full individual plan. As a consequence of Clause (2c), Sandy must be willing to assist in Tony's success, e.g. helping him find some of the raw ingredients if necessary. As a consequence of this clause and an intention-conflict axiom, Sandy cannot intend to do anything that would make it impossible for Tony to make the sauce.

3.2 Partial Individual Plans

Individual plans may be partial in several ways. Each of these will lead to some *Int. To* in which Clause (2b) of the definition in Figure 2 applies The *Achieve* function in this clause maps a proposition to a generalized action [Pollack, 1990].[8] A typical way in which partial individual plans differ from full plans is in allowing an agent to have only a partial recipe for an action. As can be seen in the definition in Figure 3, the minimal requirement for the partial plan is that the agent believe there is a recipe for *a* (Clause (0)), believe that it is able to determine any constituents of that recipe not already known (Ic), intend to obtain that recipe (Ib) and have a full individual plan for doing so (Id); a procedure associated with the GET operator will add potential intentions to do all the actions in this plan.[9] For example, if an agent is assembling a bicycle from a kit, he must believe that the accompanying instructions are complete, that he can read the instructions, and that he can perform each of them at the requisite time. While reading the instructions, the agent will adopt potential intentions to do the actions.

$\mathrm{PIP}(P, G, \alpha, T_p, T_\alpha, C_\alpha)$

$(\exists\{\beta_i, \rho_i\})[$
(0) $[\mathrm{BEL}(G, (\exists R_\alpha)[R_\alpha \in Recipes(\alpha) \ \wedge \ \{\beta_i, \rho_i\} \subseteq R_\alpha \wedge$
(1) $(\{\beta_i, \rho_i\} \subset R_\alpha \rightarrow$
(1a) $(\exists \delta_j, \rho_j)[(\{\beta_i, \rho_i\} \cup \{\delta_j, \rho_j\} = R_\alpha) \wedge$
(1b) $(\exists T_{get}, P_{get}, R_{get})$
$[\mathrm{Int.To}(G, \mathrm{GET}(G, \{\delta_j, \rho_j\}), T_p, T_{get}, C_{get/\alpha}) \wedge$
(1c) $\mathrm{CBA}(G, \mathrm{GET}(G, \{\delta_j, \rho_j\}), R_{get}, T_{get}) \wedge$
(1d) $\mathrm{FIP}(P_{get}, G, \mathrm{GET}(G, \{\delta_j, \rho_j\}),$
$T_p, T_{get}, R_{get}, C_{get/\alpha})] \wedge$
(1e) $(\forall \delta_j \in \{\delta_j, \rho_j\} \exists T_{\delta_j}) \mathrm{WCBA}(G, \delta_j, T_{\delta_j})])], T_p)\} \ \bigwedge$

(2) $set(\beta_i) = (set[\beta_k] \cup set[\beta_r]) \wedge$
(2a) $(\forall \beta_r \in set[\beta_r] \ \exists T_{\beta_r})\{$
(2a1a) $\{\mathrm{Int.To}(G, \beta_r, T_p, T_{\beta_r}, C_{\beta_r/\alpha}) \wedge$
(2a1b) $\mathrm{BEL}(G, \mathrm{WCBA}(G, \beta_r, T_{\beta_r}), T_p)\} \quad \bigvee$

(2a2) $\{(\exists G_1, \gamma, T_\gamma)$
(2a2a) $\{\mathrm{BEL}(G, \mathrm{GTD}(G, \gamma, G_1, \beta_r, T_{\beta_r}, T_\gamma), T_p) \wedge$
(2a2b) $\mathrm{Int.To}(G, \gamma, T_p, T_\gamma, C_{\gamma/\beta_r/\alpha}) \wedge$
(2a2c) $\mathrm{Int.Th}(G, (\exists R_{\beta_r})$
$\mathrm{CBA}(G_1, \beta_r, R_{\beta_r}, T_{\beta_r}), T_p, T_{\beta_r}, C_{\beta_r/\alpha}) \wedge$
(2a2d) $\mathrm{BEL}(G, \mathrm{WCBA}(G, \gamma, T_\gamma), T_p)]]\} \ \bigwedge$

(2b) $(\forall \beta_k \in set[\beta_k] \ \exists T_{\beta_k})$
(2b1) $[\mathrm{Pot.Int.To}(G, \beta_k, T_p, T_{\beta_k}, C_{\beta_k/\alpha}) \wedge$
(2b2) $\mathrm{BEL}(G, \mathrm{WCBA}(G, \beta_k, T_{\beta_k}), T_p)]]$

Figure 3: Core of Definition of Partial Individual Plan

Clause (2) represents the attitudes of the agent towards those portions of the recipe for α that he has already obtained (i.e., $\{\beta_i, \rho_i\}$)[10]. This part of his plan may still be partial in several ways; we discuss this case only informally here and give a more extensive discussion for the SharedPlan analogue. First, an agent may only have *Int. To*'s for some of the β_i's in his recipe for α (2a1a). In addition, for some of those β_i's that he intends, he may not have recipes. In this case, the agent must have on its agenda the task of finding a recipe for any β_i for which it does not have a recipe. In the contracting case (2a2), a plan is partial if the agent only has a partial plan for γ. Finally (2b), an individual plan may be partial because the agent has potential intentions that still need to be reconciled. In particular, when an agent selects a recipe for α, it directly adopts *Pot.Int.To* do the actions β_i in that recipe. However, these potential intentions must be reconciled with other intentions. If the agent discovers a conflict between adopting the intentions required for that recipe and intentions it already has, the reconciliation process may lead either to its dropping previous intentions or to its looking for a different recipe for the current action. The choice will depend on contextual factors captured in the context argument of the intention operator.

When an agent with only a partial individual plan needs to find a recipe for an action, there are two options: research (e.g. looking in an instruction manual, asking an expert) and contracting out the doing of the action. In contrast, agents engaged in SharedPlans can enlist the aid of their collaborators in such situations.

This function may be seen as connected to an agenda of tasks maintained by the agent. Discussion of this component of the model is beyond the scope of this paper. Briefly, because the agent must be committed to completing his partial plan for a for the *Int. To* operator to hold, the agenda must include tasks for establishing any of the beliefs and intentions needed in a full plan but absent in the partial plan.

[9] WCBA is a weaker version of the can bring about operator, one in which the agent may only believe it can find a recipe that it can use to do the action; WCBAG is a corresponding group operator.

[10] This clause is not applicable if the agent has not obtained any portion of the recipe, i.e. $\{\beta_i, \rho_i\} = \emptyset$.

4 SharedPlans and Intending That

Both the belief and the intention components of collaborative plans are more complex than those of individual plans. The collaborating agents must establish mutual belief of the ways in which they will perform their joint activity and must agree on the agent or agents who will do each action. Actions requiring multiple agents engender subsidiary SharedPlans of groups of agents; those requiring only a single agent lead to subsidiary individual plans. The agents also need to establish mutual belief of their individual intentions to act.

There are several important properties of these belief and intention components that are captured in the definitions that follow. First, agents do not need to know recipes for any actions that they are not personally committed to doing. In our meals example, Jan and Sandy need to establish mutual belief of the recipe for making dinner, namely that this will comprise Jan's making the appetizer, Sandy the main course, and the two of them together making the dessert. Only Jan needs to know the recipe for the appetizer; but Sandy needs to share mutual belief that Jan has such a recipe and can carry it out. The analogous case holds for Sandy and a recipe for the main course. In contrast, Sandy and Jan need mutual belief of the recipe for making dessert. Second, an agent only has *Int. 7Vs* to acts of which it is the agent. However, it has *Int.Th's* that the actions of other agents be successfully done. More generally, the ways in which the belief and intention operators are used differ. In the following definitions we presume the usual definition of mutual belief [Kraus and Lehmann, 1988] which requires infinite nestings of individual beliefs, but utilizes only a single belief operator, *BEL*. In contrast, to handle the intentions that arise in SharedPlans, we need two operators *Int. To* and *Int. Th* but there is no need for infinite embeddings of these operators (either in themselves or within one another). However, both operators may be embedded within the mutual belief operator, *MB*.

The SharedPlan operator, representing that a group of agents G has a plan to collaboratively perform some action *a*, is defined recursively in terms of full and partial SharedPlans as follows:

$$\begin{array}{ll}
\multicolumn{2}{l}{\mathrm{SP}(P, GR, \alpha, T_p, T_\alpha, C_\alpha)} \\
(1) & [(\exists R_\alpha)\mathrm{FSP}(P, GR, \alpha, T_p, T_\alpha, R_\alpha, C_\alpha)] \quad \bigvee \\
(2) & [\mathrm{PSP}(P, GR, \alpha, T_p, T_\alpha, C_\alpha) \wedge (\exists P_{ach}, T_{ach}) \\
 & \quad \mathrm{SP}(P_{ach}, GR, \mathit{Achieve}[(\exists R_\alpha)\mathrm{FSP}(P, GR, \alpha, T_p, \\
 & \qquad T_\alpha, R_\alpha, C_\alpha)], T_p, T_{ach}, C_{ach/\alpha})]
\end{array}$$

A group of agents will be said to have a SharedPlan just in case either (1) they have a full SharedPlan for doing *a* or (2) they have a partial SharedPlan, and a SharedPlan to complete that partial plan As will be seen from the definitions of these two types of SharedPlans, each of these possibilities leads eventually to individual intentions to do actions, including actions of elaborating or extending partial plans.

4.1 Full SharedPlans

We will use the formula $\mathrm{FSP}(P,G,\alpha,T_p,T_\alpha,R_\alpha,C_\alpha)$ to represent the situation in which a group of agents *G* has a full shared plan P at time T_p to do action *a* at time T_α using recipe R_α in context C_α.[11] Figure 4 gives the clauses in the core portion of the definition of this operator.[12] As in the case of full individual plans, a full SharedPlan represents the mental state of agents when they have completely determined their course of action, a state which usually does not obtain until many actions in R_α have been done. Clause (1) states that all agents in the group collaborating in the SharedPlan have an intention that they succeed in performing α. This *Int.Th* corresponds to Bratman's requirement that the agents each "intend that we" α. The other portions of the definition address the need for the agents' subplans to mesh.

Clause (2) of the definition covers those situations in which a single agent is needed to perform an action β_i in the recipe for α. In this case, the group's SharedPlan includes (2a) mutual belief by all members in the group that one member, G_k, has an individual intention to do β_i; from this mutual belief, it can be inferred[13]

$$\begin{array}{ll}
\multicolumn{2}{l}{\mathrm{FSP}(P, G, \alpha, T_p, T_\alpha, R_\alpha, C_\alpha)} \\
 & R_\alpha = \{\beta_i, \rho_i\} \wedge \mathrm{MB}(G, R_\alpha \in \mathit{Recipes}(\alpha), T_p) \wedge \\
(1) & \mathrm{MB}(G, [(\forall G_j \in G) \\
 & \quad \mathrm{Int.Th}(G_j, \mathit{Done}(G, \alpha, T_\alpha, R_\alpha), T_p, T_\alpha, C_\alpha)], T_p) \wedge \\
 & \\
(2) & [(\forall \beta_i \ \mathit{s.t.}\ \mathit{singleagent}(\beta_i)\ \exists G_k \in G, T_{\beta_i})[\\
(2a) & \quad \mathrm{MB}(G, \mathrm{Int.To}(G_k, \beta_i, T_p, T_{\beta_i}, C_{\beta_i/\alpha}), T_p) \wedge \\
(2b) & \quad ([\mathit{basic.level}(\beta_i) \wedge \\
 & \qquad \mathrm{MB}(G, \mathrm{CBA}(G_k, \beta_i, R_{Empty}, T_{\beta_i}), T_p)] \vee \\
(2c1) & \quad [\neg \mathit{basic.level}(\beta_i) \wedge \\
 & \qquad \mathrm{MB}(G, (\exists P_{\beta_i}, R_{\beta_i}) \\
(2c1a) & \qquad\qquad [\mathrm{CBA}(G_k, \beta_i, R_{\beta_i}, T_{\beta_i}) \wedge \\
(2c1b) & \qquad\qquad \mathrm{FIP}(P_{\beta_i}, G_k, \beta_i, T_p, T_{\beta_i}, R_{\beta_i}, C_{\beta_i/\alpha})], T_p) \wedge \\
(2c2) & \qquad (\exists P_{\beta_i}, R_{\beta_i}) \\
(2c2a) & \qquad\qquad [\mathrm{BEL}(G_k, \mathrm{CBA}(G_k, \beta_i, R_{\beta_i}, T_{\beta_i}), T_p) \wedge \\
(2c2b) & \qquad\qquad \mathrm{FIP}(P_{\beta_i}, G_k, \beta_i, T_p, T_{\beta_i}, R_{\beta_i}, C_{\beta_i/\alpha})]]) \wedge \\
(2d) & \quad \mathrm{MB}(G, (\forall G_j \in G, G_j \neq G_k)\mathrm{Int.Th}(G_j, (\exists R_{\beta_i}) \\
 & \qquad \mathrm{CBA}(G_k, \beta_i, R_{\beta_i}, T_{\beta_i}), T_p, T_{\beta_i}, C_{cba/\beta_i/\alpha}), T_p)]] \bigwedge \\
 & \\
(3) & [(\forall \beta_i \ \mathit{s.t.}\ \mathit{multiagent}(\beta_i)\ \exists G_k \subset G)[\\
(3a1) & \quad \mathrm{MB}(G, (\exists P_{\beta_i}, R_{\beta_i}) \\
 & \qquad\qquad [\mathrm{FSP}(P_{\beta_i}, G_k, \beta_i, T_p, T_{\beta_i}, R_{\beta_i}, C_{\beta_i/\alpha}) \wedge \\
 & \qquad\qquad \mathrm{CBAG}(G_k, \beta_i, R_{\beta_i}, T_{\beta_i})], T_p) \wedge \\
(3a2) & \quad (\exists P_{\beta_i}, R_{\beta_i}) \\
 & \qquad [\mathrm{FSP}(P_{\beta_i}, G_k, \beta_i, T_p, T_{\beta_i}, R_{\beta_i}, C_{\beta_i/\alpha}) \wedge \\
 & \qquad \mathrm{CBAG}(G_k, \beta_i, R_{\beta_i}, T_{\beta_i})] \wedge \\
(3b) & \quad \mathrm{MB}(G, (\forall G_j \in G \setminus Gk)\mathrm{Int.Th}(G_j, (\exists R_{\beta_i}) \\
 & \qquad \mathrm{CBAG}(G_k, \beta_i, R_{\beta_i}, T_{\beta_i}), T_p, T_{\beta_i}, C_{cba/\beta_i/\alpha}), T_p)]]
\end{array}$$

Figure 4: Core of Definition of Full SharedPlan

---

[11] The different group members may have different reasons for engaging in the collaborative activity. For example, hunger might underlie Jan's making dinner with Sandy, whereas a desire for social interaction underlies Sandy's making dinner with Jan. (cf. [Bratman, 1992])

[12] For space reasons, we have omitted the contracting out cases of FSP and PSP; they are similar to those for individual plans but with modifications for group knowledge and action.

[13] This inference is based on the theorem that for every α, $(*)\mathrm{BEL}(G, \mathrm{Int.To}(G, \alpha, T_p, T_\alpha, C_\alpha), T_p) \rightarrow \mathit{Int.To}(G, \alpha, T_p, T_\alpha, C_\alpha)$. The proof is based on (1) mutual belief axioms [Kraus and Lehmann, 1988] and on (2) the axioms that if α is a basic level action then $(*)$ is true, (3) similar axioms as $(*)$ for *Pot.Int.To*, and for *Int.Th* and

that G_k actually does have an *Int.To*. If β_i is basic level then (2b) the group must believe that G_k is able to bring about the action. Otherwise, the SharedPlan has embedded within it an individual plan. The group members must mutually believe (2c1) that there is some recipe which G_k knows and can use to bring about β_i. Clause (2c2) represents the fact that G_k does know a recipe that he believes will enable him to bring about β_i, and has a FIP for doing β_i using this recipe. Thus, the different scopings of (2c1) and (2c2) accurately capture an important distinction. G_k must know the recipe he will use to do β_i. The other members of the group do not need to know this recipe; however, all members of the group need to mutually believe that there is some recipe that G_k can use.

The *Int.Th* in Clause (2d) represents commitment on the part of all other agents in the group to ensuring that G_k can do β_i.[14] Although the preceding clauses have some analogue in the definition of FIP, this does not. This clause, along with Clauses (1) and (2a), distinguishes this situation from the contracting situation. All of the agents in G are committed to the performance of α and the β_i, and (through the context parameter) to the β_i as part of doing α.

We have developed several axioms that are needed to support the role of the *Int.Th* operator in coordinating the intentions and actions of different agents in collaborative activities. We can provide only one example here, an axiom of helpful behavior in the SharedPlan context. This axiom applies whenever an intention of the type represented in Clause (2d) is held; such *Int.Th*'s always arise from G_j and G_k collaborating. The axiom states that G_j will adopt a potential intention to do an action γ that he believes will assist G_k in being able to do β_i (which in this context is contributing to the joint action α) if the group saving from this γ is worth more to G_j than the cost of doing γ. The axiom also addresses Bratman's requirement of minimally cooperatively stable intentions. For example, while chopping onions for the lasagne, Sandy may chop an extra one to assist Jan in making the cheese puffs. These axioms combined with the intention-conflict avoidance axiom provide the basis for agents adopting meshing subplans. Jan's individual plan for making the appetizer cannot produce intentions that conflict with his intention that Sandy succeed in making the main course. For example, Jan cannot *Int.To* an action that would make it impossible for Sandy to boil the lasagne noodles. If potential intentions lead to consideration of adopting such conflicting intentions, the reconciliation axioms will cause one intention to be dropped; some portion of the SharedPlan will then become, or remain, partial.

Clause (3) of the FSP is the extension to cover the case when a constituent in the recipe for α is a group action. From (3a) group G's full SharedPlan for α is seen to include as a component the full SharedPlan of a subgroup G_k to do β_i. As in the case of individual action, there is a distinction between members of this subgroup G_k who must know the recipe for β_i and have the FSP (3a2), and other members of the group who only need to believe there is some such recipe known to the subgroup (3a1); again the scoping difference captures this distinction. The context parameter of the FSP includes the information that β_i is being done as part of doing α. The first group can-bring-about clause (3a1) represents the mutual belief of the group in the ability of the subgroup to do β_i at the appropriate time. The second group can-bring-about clause (3a2) is needed to ensure that the subgroup can actually bring about the FSP according to the given recipe. Again, there is an *Int.Th* clause (3b) that represents a commitment on the part of all members of the group not in G_k to insuring that the members of G_k find a recipe for doing β_i that they can succeed in using.

4.2 Partial SharedPlans

Like partial individual plans, partial SharedPlans differ from full ones in allowing agents to have only a partial recipe for action. And, analogously to the individual case, the minimal requirements for a group of agents having a partial SharedPlan are (1) that the agents mutually believe there is a recipe for α that they can obtain, (2) that they have a full SharedPlan to obtain that recipe, and (3) that every member of the group have an individual intention for elaborating the recipe.

Figure 5 shows the portion of the definition of a partial SharedPlan that represents the mental state of the agents when they have a recipe for α, but recipes for only some of the β_i's involved in doing α. The β_i have been divided into two subgroups: the β_r are actions for which the group has done some deliberation and elaboration, including a decision of who will do the action; the β_k are those actions for which it has not.

Clause (2) gives the conditions that must pertain for single agent β_r. According to Clause (2a), the group must agree on the agent of the action and mutually believe that this agent, G_k, intends to do β_r. That G_k does intend to do β_r may be concluded based on the appropriate BEL and *Int.To* axioms and theorems. This *Int.To* operator may fulfill Clause (2b) of the definition in Figure 2; the agent may have only a partial plan for doing β_r, but it must have an intention to complete that plan.

Clause (2b) is a weaker form of the the can-bring-about constraints in FSPs (Figure 4); notice that there is no clause corresponding to (2c1b) and (2c2b) of FSPs, reflecting the partiality of PSPs. Clause (2c) of the PSP corresponds to (2d) in the FSP; it states that the other agents must all have an intention that the agent be able to perform β_r according to the appropriate recipe. This intention is a commitment on the part of the other agents to assisting G_k.

Clause (3) shows that a similar situation pertains for those β_r that require a subgroup of agents. In this case, however, the subgroup must have a SharedPlan to do the action; this SharedPlan may be either a FSP or a PSP. Again (Clause (3c)), the other members of G must have an intention that the subgroup succeed. Finally (4),

---

Pot.Int.Th for any proposition.

[14] Bratman [Bratman, 1992] and others [Cohen and Levesque, 1991] provide arguments that mutual belief of individual intentions is not sufficient.

$$\mathrm{PSP}(P, G, \alpha, T_p, T_\alpha, C_\alpha)$$

$$\begin{array}{ll}
 & (\exists\{\beta_i, \rho_i\})[\\
(0) & \mathrm{MB}(G, (\exists R_\alpha)[\{\beta_i, \rho_i\} \subseteq R_\alpha \wedge R_\alpha \in \mathit{Recipes}(\alpha)], T_p) \wedge \\
 & \mathit{set}(\beta_i) = \mathit{set}[\beta_k] \cup \mathit{set}[\beta_r] \;\; \wedge \\
(1) & \mathrm{MB}(G, (\forall G_j \in G) \\
 & \quad \mathrm{Int.Th}(G_j, \mathit{Done}(G, \alpha, T_\alpha, R_\alpha), T_p, T_\alpha, C_\alpha), T_p) \wedge \\
 & \\
(2) & (\forall \beta_r \in \mathit{set}[\beta_r] \mathit{s.t.singleagent}(\beta_r) \exists G_k \in G, T_{\beta_r}) \\
(2a) & [\mathrm{MB}(G, \mathrm{Int.To}(G_k, \beta_r, T_p, T_{\beta_r}, C_{\beta_r/\alpha}), T_p) \wedge \\
(2b) & \mathrm{MB}(G, \mathrm{WCBA}(G_k, \beta_r, T_{\beta_r}) \wedge \\
(2c) & \quad (\forall G_j \in G, G_j \neq G_k) \mathrm{Int.Th}(G_j, (\exists R_{\beta_r}) \\
 & \quad \mathrm{CBA}(G_k, \beta_r, R_{\beta_r}, T_{\beta_r}), T_p, T_{\beta_r}, C_{cba/\beta_r/\alpha}), T_p)] \bigwedge \\
 & \\
(3) & (\forall \beta_r \in \mathit{set}[\beta_r] \; \mathit{s.t.} \; \mathit{multiagent}(\beta_r) \exists G_k \subset G, T_{\beta_r}) \\
(3a1) & [\mathrm{MB}(G, (\exists P_{\beta_r}) \mathrm{SP}(P_{\beta_r}, G_k, \beta_r, T_p, T_{\beta_r}, C_{\beta_r/\alpha}), T_p) \wedge \\
(3a2) & (\exists P_{\beta_r}) \mathrm{SP}(P_{\beta_r}, G_k, \beta_r, T_p, T_{\beta_r}, C_{\beta_r/\alpha}) \wedge \\
(3b) & \mathrm{MB}(G, \mathrm{WCBAG}(G_k, \beta_r, T_{\beta_r}) \wedge \\
(3c) & \quad (\forall G_j \in \{G \setminus G_k\}) \mathrm{Int.Th}(G_j, (\exists R_{\beta_r}) \; \mathrm{CBAG}(G_k, \\
 & \quad \beta_r, R_{\beta_r}, T_{\beta_r}), T_p, T_{\beta_r}, C_{cbag/\beta_r/\alpha}), T_p)] \; \bigwedge \\
 & \\
(4) & (\forall \beta_k \in \mathit{set}[\beta_k]) \; (\exists T_{elab}) \\
(4a) & [\mathrm{MB}(G, (\forall G_j \in G) \mathrm{Int.To}(G_j, \\
 & \quad \text{Make-SP/Int}(G_j, G, \beta_k, C_\alpha), T_p, T_{elab}, C_{elab/\beta k/\alpha}) \wedge \\
(4b) & \quad \mathrm{WCBAG}(G, \beta_k, T_{\beta_k}), T_p)]]
\end{array}$$

Figure 5: Core of Definition of Partial SharedPlan

there are the actions β_k. For these we require intentions on the part of every group member to find a recipe and decide who (sub-group or single agent) will do the action. The Make-SP/Int operator corresponds to a group decision making process about who will do β_k and how if the whole group is doing it; discussion of this process is beyond the scope of this paper. Once they have elaborated the plan at least partially, the β_r case holds.

5 Conclusions and Future Work

To provide an account of collaborative activity, Searle [Searle, 1990] introduced the notion of 'we-intention.' Grosz and Sidner [Grosz and Sidner, 1990] argued that such a notion should not be necessary and their initial formulation of SharedPlans avoids use of one. However, the definitions provided in that formulation could only accommodate group activity that directly decomposed into actions of individual agents. Subsequent work in AI on formalizing the plans and intentions of multiple agents has, like Searle's proposal, included some notion of joint intention [Cohen and Levesque, 1990; Rao et al., 1992]. In this paper, we have provided a formulation that again avoids the need for a notion of joint intention. In this work, SharedPlans serve two major roles. They summarize the set of beliefs and intentions needed for collaborative activity, and also provide the rationale for the process of revising beliefs and intentions; consequently, they motivate the collaborative correlate of means-ends reasoning in the plans of an individual agent. SharedPlans ground out in the individual intentions of individual agents and the individual plans that they engender. This formulation accommodates the properties of shared cooperative activity proposed by Bratman [Bratman, 1992]. Intentions to do constituent actions form the basis of each individual's actions. Intentions-that directed toward other agents abilities to act and success in acting, as well as toward the success of the joint activity, ensure cooperation in subplans and helpful behavior.

References

[Allen and Perrault, 1980] J F. Allen and C.R. Perrault. Analyzing intention in utterances. Artificial Intelligence, 15(3):143 178, 1980.

[Balkanski, 1990] C. Balkanski. Modelling act-type relations in collaborative activity. Tr-23-90, Harvard University, 1990.

[Bratman et al., 1988] M. Bratman, D. Israel, and M. Pollack. Plans and resource-bounded practical reasoning. Computational Intelligence, 4:349-355, 1988.

[Bratman, 1987] Michael E. Bratman. Intention, Plans, and Practical Reason. Harvard University Press, Cambridge, Massachusetts and London, England, 1987.

[Bratman, 1992] M E . Bratman. Shared cooperative activity. The Philosophical Review, 101:327-341, 1992.

[Cohen and Leveque, 1991] P. Cohen and H. Leveque. Teamwork. Tr 503, SRI International, 1991.

[Cohen and Levesque, 1990] P. R. Cohen and Hector Levesque. Rational interaction as the basis for communication. In Cohen et al. 1990.

[Cohen el al., 1990] P. R. Cohen, J. Morgan, and M. E. Pollack (editors). Intentions in Communication. MIT Press, 1990.

[Grosz and Sidner, 1986] BJ. Grosz and C.L. Sidner. Attention, intentions, and the structure of discourse. Computational Linguistics, 12(3), 1986.

[Grosz and Sidner, 1990] B.J. Grosz and C.L. Sidner. Plans for discourse. In Cohen et al. 1990.

[Kraus and Lehmann, 1988] S. Kraus and D. Lehmann. Knowledge, belief and time. Theoretical Computer Science, 58:155-174, 1988.

[Kraus, 1993] S. Kraus. Agents contracting tasks in non-collaborative environments. In Proc. of AAAI-98, Washington, D C , 1993.

[Lochbaum et al., 1990] K.E. Lochbaum, B.J. Grosz, and C.L. Sidner. Models of plans to support communication: An initial report. In Proceedings of AAAI-90, Boston, MA, 1990.

[Lochbaum, 1991] K. E. Lochbaum. An algorithm for plan recognition in collaborative discourse. In Proc. ACL-91, Berkeley, CA, 1991.

[Pollack, 1990] M. Pollack. Plans as complex mental attitudes. In Cohen et al. 1990.

[Rao et al., 1992] A. S. Rao, M. P. Georgeff, and E. A Sonenberg. Social plans. In Decentralized Artificial Intelligence, Volume 3. Elsevier Science Pub., 1992.

[Searle, 1990] J.R. Searle. Collective intentionality. In Cohen et al. 1990.