Formal Approaches to Teamwork

John Grant, Sarit Kraus and Don Perlis

1 Introduction

Cooperation matters. Many everyday tasks cannot be done at all by a single agent, and many others are done more effectively by multiple agents. Moving a very heavy object is an example of the first sort, and moving a very long (but not heavy) object can be of the second. Many researchers have investigated the automation of cooperative behavior, producing a large body of published work with philosophical, formal, and implementational aspects. In this paper we focus on the second of these; our aim is to provide a comparative overview of a (hopefully somewhat representative) variety of approaches. Toward this end we have chosen a simple example of cooperative behavior (two agents working together to move a heavy object) and attempted to treat this example in some detail in each of the six distinct formal approaches, in order to reveal their differences and similarities, and noting where we needed to make some changes either to the example or to the formalism.

It is useful initially to ask just what it takes for agents to cooperate. Several authors provide answers stressing different aspects of cooperation. Bratman [1992] gives four criteria that a multi-agent activity must meet to be a "shared cooperative activity:" mutual responsiveness, commitment to the joint activity, commitment to mutual support, and formation of subplans that mesh with one another. Two of the papers we will consider later in detail deal with some key concepts about teams and cooperation. [Sonenberg *et al.*, 1992] states that a team of agents must have mutual beliefs, joint goals, joint plans, and joint intentions. [Wooldridge and Jennings, 1999] give four stages in cooperative behavior: recognition of the possibility for cooperation, team formation, plan formation, and execution.

There is now a substantial literature on teamwork. Some papers present non-formal philosophical discussion on what constitute teams (e.g., [Tuomela and Miller, 1988; Tuomela, 1991; Searle, 1990; Bratman, 1992]). Some of these non-formal models served as the basis for the formal ones. Others take a practical approach and provide techniques to implement teams on top of formal approaches (e.g., [Jennings, 1995; Tambe, 1997]). There are many other proposed interesting formalizations including [Singh, 1991; Castelfranchi, 1995; Singh, 1998]. For space reasons we have chosen to deal in detail with just six approaches (in some cases we combined several papers of the authors). We leave for future work an extended survey of formal teamwork models.

The formal models of teamwork build on formal models of mental states of individuals. There is substantial work on models of knowledge and belief such as [Megiddo, 1989; Konolige, 1986; Levesque, 1984; Shoham, 1989; Moore, 1985; Kraus and Lehmann, 1988; Gmytrasiewicz *et al.*, 1992; Haass, 1983; Dubois and Prade, 1994; Fagin *et al.*, 1995; Grant *et al.*, 2000]. Others discussed models of goals and intentions including [Singh and Asher, 1993; Rao and Georgeff, 1991b; Cohen and Levesque, 1990; Perrault, 1990; Cohen *et al.*, 1990; Moore, 1985; Konolige and Pollack, 1993; Rao and Georgeff, 1991a; Schut *et al.*, 2001; Kraus *et al.*, 1998; Kumar *et al.*, 2000; Alonso, 1997; Wooldridge, 2000]. To survey and compare all these works is not in the scope of our paper.

In Section 2 we describe an example task for two agents. We tried to find an example that is neither trivial nor difficult and that requires the cooperation of two agents as a team. The agents must go to a location where they find a large block they have to push to a new location. They must push together and do this twice in order to avoid an obstacle.

Section 3 contains the main part of the paper: we show how to express the example in the six formalisms. In each case we give names to the approaches based on the terminology used by the authors. Except for $SharedPlans$ these are not official names for the formalisms discussed. We order the formalizations based on the order in which the key papers appeared. We call the approach of Cohen and Levesque (sometimes with co-authors) $Joint\ Intentions$ [Levesque *et al.*, 1990]. The approach of the Sonenberg group we call $Team\ Plans$ [Sonenberg *et al.*, 1992]. We use the term $SharedPlans$ for the work of Grosz and Kraus [Grosz and Kraus, 1996]. We refer to Wooldridge and Jennings's model as *Cooperative Problem Solving* [Wooldridge and Jennings, 1999] and to Dunin-Keplicz and Verbrugge's by *Collective Intentions* [Dunin-Keplicz and Verbrugge, 2002]. Finally, we call the Grant, Kraus and Perlis approach *Cooperative Subcontracting* [Grant *et al.*, 2005].

In Section 4 we compare the six approaches based on several criteria, including their focus, formalism, representation of time, and facility for expressing complex plans. The paper ends with a brief conclusion and suggestion for future work in Section 5.

2 Example Description

We illustrate the six approaches we have chosen by expressing the same example in all of them. We chose an example that is neither particularly difficult, nor trivial, and that shows some important aspects of the formalizations of teamwork.

Agents A_1 and A_2 have the joint goal of pushing block Bl from location $locO$ (origin) to $locD$ (destination), starting at time t. There is an obstacle in the way so they have to push the block twice, once in the north direction, and then to the west. First the agents need to (plan to) be at $locO$, to be on the south side of Bl, and to do a push together; then they need to be on the west side of Bl and do another push. Bl will be at location $locT$ (temporary location) after the first push. The two agents cannot be in the same place simultaneously.

We assume that all the actions needed for performing the task are specified in a recipe. Hence, there is no need to plan in order to satisfy preconditions. In some systems the subactions are not specified in recipes and we will need to adjust the notation. Not surprisingly, as different authors stress different aspects of teamwork, the particular details are not equally well suited to all the approaches.

We assume that before the agents start the task, A_1 is at $loc1$ and A_2 is at $loc2$. Each agent is associated with a specified speed at which it moves. We do not deal with the speed separately. Thus the time it takes to move between two places is the distance between these places divided by the speed. Both the distance between $loc1$ and $locO$ and the distance between $loc2$ and $locO$ are known.

Then A_1 goes to the west corner of the south side of $locO$, $WC\_SS(locO)$ and A_2 goes to the east corner of the south side of $locO$, $EC\_SS(locO)$. For simplicity we assume that the "Go" action is a basic-level action. We assume basic-level actions are executable at will if appropriate situational conditions hold, and do not define this further (see Pollack's argument that this is a reasonable assumption in a computational setting [Pollack, 1986]).

The "Push" action is a multi-agent action (and thus a complex action). It requires two agents, one doing LPush, push from the left, and one doing RPush, push from the right. The push action takes one time unit.

Then agent A_1 goes to the north corner of the west side of $locT$ and A_2 goes to the south corner of the west side of $locT$ and they push again.

Notations:

- The action to be performed is $\alpha = Move(\{A_1, A_2\}, Bl, locO, locD)$, i.e., A_1 and A_2 will move block Bl from $locO$ to $locD$.
- We denote the group of the agents $GR = \{A_1, A_2\}$

- There are six subactions in the recipe for α.
  $\beta_1 = Go(A_1, loc1, WC\_SS(locO))$
  $\beta_2 = Go(A_2, loc2, EC\_SS(locO))$
  $\beta_3 = Push(\{A_1, A_2\}, Bl, locO, locT)$
  $\beta_4 = Go(A_1, WC\_SS(locT), NC\_WS(locT))$
  $\beta_5 = Go(A_2, EC\_SS(locT), SC\_WS(locT))$
  $\beta_6 = Push(\{A_1, A_2\}, Bl, locT, locD)$

- The recipe for α consists of the six subactions and a set of order constraints. We use the predicate $Before(\beta, \beta')$ to denote that β should be performed before β'. In the example, β_1 and β_2 should be performed before β_3; β_3 should be performed before β_4 and β_5 and the last two should be performed before β_6. The instantiated recipe for $\alpha = Move(\{A_1, A_2\}, Bl, locO, locD)$ is: $R_\alpha = \{\{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6\},$ $\{Before(\beta_1, \beta_3), Before(\beta_2, \beta_3), Before(\beta_3, \beta_4), Before(\beta_3, \beta_5),$ $Before(\beta_4, \beta_6), Before(\beta_5, \beta_6)\}\}$

The recipe for the action $\beta_i = Push(GR, Bl, locX, locY)$ (with arbitrary locations $locX$ and $locY$) consists of two actions: $\gamma_i^1 = LPush(A_1, Bl, locX, locY)$ and $\gamma_i^2 = RPush(A_2, Bl, locX, locY)$. Denote by $Sim(\beta, \beta')$ the constraint that β and β' should be performed simultaneously. The recipe for such an action is: $\{\{\gamma_i^1, \gamma_i^2\}, \{Sim(\gamma_i^1, \gamma_i^2)\}\}$

We denote by t_p the time when the plans, intentions, and beliefs are formed and by t the time when the agents start performing α.

For each action β, T_β denotes the time interval during which β is performed. We write k_1 and k_2 for the amount of time it takes for agent A_1 and A_2 respectively to go from their initial location to $locO$, and $k = max(k_1, k_2)$.

1. $T_{\beta_1} = [t, t + k_1]$.
2. $T_{\beta_2} = [t, t + k_2]$.
3. β_3 will start after β_1 and β_2. Thus
   $T_{\beta_3} = [t + k + 1, t + k + 2]$.
4. $T_{\beta_4} = [t + k + 3, t + k + 4]$.
5. $T_{\beta_5} = [t + k + 3, t + k + 4]$.
6. β_6 will start after β_4 and β_5. Thus $T_{\beta_6} = [t + k + 5, t + k + 6]$.
7. Thus $T_\alpha = [t, t + k + 6]$.

3 Formal Approaches

This is the main part of the paper. Here we express the example presented in the previous section in six formalizations, also making some comments about these approaches along the way. We present mainly those aspects that are directly relevant to the example and omit many important details discussed by the authors. In the presentations we use the same symbol α as a generic symbol for actions in definitions as well as for the specific action to move the block *Bl* from *locO* to *locD*. The meaning should always be clear from the context. We also make some changes to notation for the sake of uniformity: in particular we change Lisp notation, (Pxy) to P(x,y). In the following section we will also compare the six approaches in various ways.

3.1 Joint Intentions

Cohen and Levesque with several of their colleagues presented a series of papers on models for teamwork ([Levesque *et al.*, 1990; Cohen and Levesque, 1991b; Cohen and Levesque, 1991a; Cohen *et al.*, 1997]. Here we consider mainly the first paper of the series by Levesque, Cohen and Nunes [Levesque *et al.*, 1990] which is the most formal one. The other papers build on it. That paper focuses on the persistence of joint intentions and when agents will drop joint intentions.

They use constructs of dynamic logic to describe sequences of actions and modal operators to express time associated propositions. In their logic, $\alpha_1;\alpha_2$ is action composition and $\alpha_1||\alpha_2$ is the concurrent occurrence of α_1 and α_2. $p?$ is a test action: if p is true, the action succeeds, but if p is false it fails.

They do not deal with groups of agents performing complex subactions; in fact, they do not deal with expressing actions at different levels. Therefore we consider only basic-level actions when expressing our example in their system.

We denote $\alpha = [(\beta_1||\beta_2);(\gamma_3^1||\gamma_3^2);(\beta_4||\beta_5);(\gamma_6^1||\gamma_6^2)]$ that is:

$$\begin{aligned}
&\alpha = \\
&\quad [(Go(A_1, loc1, WC_SS(locO))||Go(A_2, loc2, EC_SS(locO)));\\
&\quad\ (LPush(A_1, Bl, locO, locT)||RPush(A_2, Bl, locO, locT));\\
&\quad\ (Go(A_1, WC_SS(locT), NC_WS(locT))||\\
&\quad\ \ Go(A_2, EC_SS(locT), SC_WS(locT)));\\
&\quad\ (LPush(A_1, Bl, locT, locD)||RPush(A_2, Bl, locT, locD))]
\end{aligned}$$

Levesque, Cohen and Nunes [Levesque *et al.*, 1990] do not show how to express exact times. In particular, they do not describe how to specify joint intentions for specific times. However, in another paper [Cohen and Levesque, 1990] that considers single agent intentions, they do present a way to specifying exact time. They introduce the concept of a time proposition such as $2:30PM/3/6/85$. Using a time proposition we can express tests of the form $2:30PM/3/6/85?$. Writing $2:30PM/7/20/05?; \beta_1; 2:35PM/7/20/05?$ will express the sequence of events where β_1 will be performed July 20 between $2:30PM$ and $2:35PM$. In particular, suppose that in our example the agents should move the block on July 20, 2005. So if we say that t is 2PM and have specific values for k_1 and k_2, say 5 minutes and 10 minutes respectively, and we assume that the time unit is 1 minute, we can write the following:

$\alpha =$ [2:00PM/7/20/05?;
(β_1 2:05PM/7/20/05?||β_2 2:10PM/7/20/05?);
2:11PM/7/20/05?($\gamma_3^1||\gamma_3^2$);
2:13PM/7/20/05?; ($\beta_4||\beta_5$);
2:15PM/7/20/05?; ($\gamma_6^1||\gamma_6^2$);
2:16PM/7/20/05?]

To avoid the complexity of dealing with time, we consider the α without the time expression in the rest of this section.

In addition to dynamic logic constructs, [Levesque *et al.*, 1990] use temporal modal logic operators:

- UNTIL(p, q) specifies that until p is true q will remain true.
- $\Box p$ indicates that p is true from now on.
- $\Diamond p$ indicates that p is true at some point in the future.
- DONE(x, y, α) indicates that α has just happened and x and y are its agents.
- DOING(x, y, α) indicates that x and y are doing α.

They also apply modal operators for mental states:

- GOAL(x, p) says that x has p as a goal.
- BEL(x, p) says that x has p as a belief.
- MB(x, y, p) says that p is mutually believed by x and y.

Using these operators they define the concept of a persistent goal:

- $PGOAL(x, p, q)$ means that x has a persistent goal that p will be true because of q.

$$\begin{aligned}&\text{PGOAL}(x, p, q) =_{def}\\&\quad \text{BEL}(x, \neg p) \wedge \text{GOAL}(x, \diamond p) \wedge\\&\quad UNTIL([\text{BEL}(x, p) \vee \text{BEL}(x, \Box\neg p) \vee \text{BEL}(x, \neg q)]\\&\qquad \text{GOAL}(x, \diamond p))\end{aligned}$$

  Intuitively this means that x has a persistent goal that p become true if it believes that p is false but wants it to become true until x comes to believe that p is true or p will never be true or that q is false.

They also define two more complex notions with respect to two agents:

- $\text{MG}(x, y, p) =_{def} \text{MB}(x, y, \text{GOAL}(x, \diamond p) \wedge \text{GOAL}(y, \diamond p))$. Intuitively, MG(mutual goal) means that both agents would like p to become true eventually and this is mutually believed.

- $$\begin{aligned}&\text{WG}(x, y, p) =_{def}\\&[\neg\text{BEL}(x, p) \wedge \text{GOAL}(x, \diamond p)] \vee\\&[\text{BEL}(x, p) \wedge \text{GOAL}(x, \diamond \text{MB}(x, y, p))] \vee\\&[\text{BEL}(x, \Box\neg p) \wedge \text{GOAL}(x, \diamond \text{MB}(x, y, \Box\neg p))]\end{aligned}$$

  This form of ”weak goal” (WG) involves three mutually exclusive cases: either x has the goal that p will be true sometime in the future, or believes that p is true and wants to make it mutually believed or believes that p will never be true from now on and wants to make that mutually believed.

- They also define WMG expressing that both agents have a ”weak goal” that is mutually believed:

  $\text{WMG}(x, y, p) =_{def} \text{MB}(x, y, \text{WG}(x, y, p) \wedge \text{WG}(y, x, p))$.

Suppose q is a reason why the agents want to perform α. JPG, joint persistent goal, in our example with respect to q would be expressed as follows:

$$\begin{aligned}&\text{JPG}(A_1, A_2, \text{DONE}(A_1, A_2, \alpha), q) =^{def}\\&\ \text{MB}(A_1, A_2, \neg\text{DONE}(A_1, A_2, \alpha)) \wedge \text{MG}(A_1, A_2, \text{DONE}(A_1, A_2, \alpha)) \wedge\\&\ \text{UNTIL}([\text{MB}(A_1, A_2, \text{DONE}(A_1, A_2, \alpha)) \vee \text{MB}(A_1, A_2, \Box\neg\text{DONE}(A_1, A_2, \alpha))\\&\qquad\qquad \vee \text{MB}(A_1, A_2, \neg q)],\\&\qquad \text{WMG}(A_1, A_2, \text{DONE}(A_1, A_2, \alpha)))\end{aligned}$$

The two clauses indicate that A_1 and A_2 mutually believe that α has not been done yet and that they both have the mutual goal for α to be done eventually. The clause that starts with UNTIL specifies when the agents will stop having the persistent goal and what goals and beliefs A_1 and A_2 will have until then. In particular, the agents will stop having the JPG when either they mutually believe that α has been done, or they will mutually believe that α will never be done or they will mutually believe that the reason for doing α, namely q, is not true any more. Until then, they will have a weak mutual goal, which means that they mutually believe that both of them have a weak goal with respect to the other agent. A_1 having a WG with respect to A_2 that α will be done means that one of the following is true: (i) A_1 does not believe that the α has been done and has a goal that it will be eventually done. (ii) A_1 believes that α has been done and has a goal that this belief will become a mutual belief with A_2. (iii) A_1 believes that α will never be done and has a goal that this belief will become jointly believed with A_2.

Finally, we are ready to present their notion of joint intentions:

$$
\begin{array}{l}
\mathrm{JI}(A_1, A_2, \alpha, q) =^{def} \\
\quad \mathrm{JPG}(A_1, A_2, \\
\quad\quad \mathrm{DONE}(A_1, A_2, \\
\quad\quad\quad \mathrm{UNTIL}(DONE(A_1, A_2, \alpha), \\
\quad\quad\quad \mathrm{MB}(A_1, A_2, \mathrm{DOING}(A_1, A_2, \alpha)))?; \alpha), \\
\quad\quad q)
\end{array}
$$

A_1 and A_2 having joint intentions to do α means that they have a joint persistent goal to do α while mutually believing (throughout the execution of α) that they are doing it.

The main question is what are the individual intentions and beliefs that A_1 and A_2 have when they have joint intentions. Levesque, Cohen and Nunes prove that each agent has a persistent goal that α will be done, i.e., $PGOAL(A_1, \alpha, q)$ and $PGOAL(A_2, \alpha, q)$, but it is not clear what will be the intentions and beliefs toward the subactions, β_i. They say that if an agent is committed to doing $\beta_1; \beta_2$, it does not follow that the agent is committed to each action separately; for example, it does not make sense to be committed to push the block if going to $locO$ failed. More non formal discussion on these issues are presented in [Cohen and Levesque, 1991b; Cohen *et al.*, 1997]. In their paper [Cohen and Levesque, 1991a] they prove that if the agents have a joint persistent goal that $\beta_1; \beta_2$ will be done, if they mutually believe that β_2 has not been done, then they will keep having a joint persistent goal that β_2 will be done at least as long they have the persistent goal that $\beta_1; \beta_2$ will be done. In our example, if A_1 and A_2

mutually believe that the push has not been done they will keep on having a persistent joint goal to do it at least while they have a persistent goal to do α.

3.2 Team Plans

Sonenberg et al in a series of papers [Sonenberg *et al.*, 1992; Kinny *et al.*, 1994; Rao *et al.*, 1992; Tidhar *et al.*, 1996] focus on groups performing complex plans to achieve a goal. They discuss how to find a suitable team for a task; how to synchronize the establishment of joint goals and the adoption of intentions; how to assign roles; and how to maintain proper temporal relations while executing different parts of the plan by different members of the team.

Here we consider the most extended version [Sonenberg *et al.*, 1992]. They apply three modal operators:

- $\mathrm{BEL}(A, \phi)$ means that an agent A believes that ϕ is true.
- $\mathrm{GOAL}(A, \phi)$ means that A has the goal ϕ.
- $\mathrm{INTEND}(A, \pi)$ means that A intends to do π.

Teams play an important role in their model. An individual agent is a team and a set of agents (either ordered or not) is a team. They also refer to team roles that we denote using $a, b, ...$ We use τ to denote actual teams. Using the individual agent operators they define team operators:

- $\mathrm{MBEL}(\tau, \phi)$ means that the team τ jointly believe ϕ.
- $\mathrm{JGOAL}(\tau, \phi)$ means that the team τ has the joint goal ϕ.
- $\mathrm{JINTEND}(\tau, \pi)$ means that the team τ jointly intend to do π.

They reduce all the multi-agent attitudes to single agent attitudes and provide mechanisms for expressing which part of a joint plan is the responsibility of each member of the team.

As in the Joint Intentions approach, they also apply dynamic logic for simple plan expressions. However, their notation is slightly different. If α is a primitive action, then $*(\tau, \alpha)$ denotes the performance of α by the team τ; $!(\tau, \phi)$ the achievement of ϕ by τ; $(?\tau, \phi)$ the testing of ϕ by τ; and $\hat{}(\tau, \phi)$ that τ is waiting for ϕ to become true. They use ; for sequencing, & for and-parallelism and | for non deterministic choice; $ in front of a string indicates that the string is a variable. (This notation is slightly different from the one in [Levesque *et al.*, 1990] where parallelism is denoted by ||.) They do not discuss the representation of explicit times.

A plan is a tuple $(p, \phi_{purpose}, \phi_{precondition}, \rho_{body})$ where p is a unique plan name, $\phi_{purpose}$ is a goal formula that motivates the plan, $\phi_{precondition}$ is the preconditions that must be true in order to perform the plan and ρ_{body} is a plan graph or a plan expression. The plan includes team variables that represent the roles of the plan. There may be a library of plans that are available with non-assigned roles. Once a team is formed, the roles are assigned to specific individual agents or specific teams. For readability we simplified the notation of teams and roles. The syntax of [Sonenberg *et al.*, 1992] is much richer and allows dealing with very complex role structures and teams.

In our Move example, $\phi_{purpose}$ is
$\mathrm{JGOAL}((a, b, (a, b)), \mathrm{done}(!((a, b, (a, b))Move(Bl, locO, locD))))$.

In the team there are the individual agent roles a and b as well as a team (a, b) associated with the multi-agent actions.

The preconditions are $at(loc1, a, t)$ and $at(loc2, b, t)$.

The body is as follows:

$$
\begin{array}{l}
(*(a, Go(a, loc1, WC\_SS(locO)))\& * (b, Go(b, loc2, EC\_SS(locO))));\\
\;\; * ((a, b), Push((a, b), Bl, locO, locT));\\
(*(a, Go(a, WC\_SS(locT), NC\_WS(locT)))\&\\
\;\; * (b, Go(b, EC\_SS(locT), SC\_WS(locT))));\\
\;\; * ((a, b), Push((a, b), Bl, locT, locD))
\end{array}
$$

When the actual team is formed, the role a is assigned to A_1 and b to A_2. Any agent may form a team. Assume that A_1 forms the team of our example and serves as the team leader. It needs to send A_2 the joint goal to move the block, as well as the joint plan and role assignments. A_1 is assigned the role a and A_2 the role b. The authors describe two protocols A_1 and A_2 can use to form the relevant beliefs and intentions (commit-and-cancel and agree-and-execute). If the agents agree on the plan p and the role assignments, they form a joint intention of the form

$$
\begin{array}{l}
\mathrm{JINTEND}((A_1, A_2, (A_1, A_2)), p) \equiv\\
\;\; \mathrm{INTEND}(A_1, p) \wedge \mathrm{BEL}(A_1, \mathrm{JINTEND}((A_1, A_2, (A_1, A_2)), p))\wedge\\
\;\; \mathrm{INTEND}(A_2, p) \wedge \mathrm{BEL}(A_2, \mathrm{JINTEND}((A_1, A_2, (A_1, A_2)), p))\wedge\\
\;\; \mathrm{JINTEND}((A_1, A_2), p) \wedge \mathrm{MBEL}((A_1, A_2), \mathrm{JINTEND}((A_1, A_2, (A_1, A_2)), p))
\end{array}
$$

The first two lines specify that each agent intends to perform the complex plan p and believes that A_1, A_2 and the team (A_1, A_2) have a joint intention for doing p. The third line indicates that the team (A_1, A_2) jointly intend p and mutually believe that they all have a joint intention to do the plan p.

For the execution of the plan, each agent transforms the plan to the subplan that it needs to perform based on the agent's role. The function *try* attempts to execute a plan expression and returns the value true if successful, and false otherwise. For example, for the action *Go* that A_1 needs to perform, it will have the following in its executable plan:

$$
\begin{aligned}
&*(A_1, \$r = try(Go(A_1, loc1, WC\_SS(locO))));\\
&(?(A_1, \$r); !(A_1, broadcast(succeeded(Go(A_1, loc1, WC\_SS(locO)))))|\\
&(?(A_1, \neg\$r); !(A_1, broadcast(failed(Go(A_1, loc1, WC\_SS(locO)))));\\
&\quad * (A_1, fail))).
\end{aligned}
$$

That is, A_1 will try to perform the *Go* action. The result of this attempt will be stored in $\$r$. If A_1 succeeds, it will broadcast this success; if A_1 fails it will broadcast its failure.

For the part of the plan where A_1 is not involved, such as the *Go* action for A_2, it will include the following:

$$
\begin{aligned}
&\hat{}(A_1, failed(Go(A_2, loc2, EC\_SS(locO))\vee\\
&\quad succeeded(Go(A_2, loc2, EC\_SS(locO)))));\\
&?(A_1, succeeded(Go(A_2, loc2, EC\_SS(locO))))
\end{aligned}
$$

That is, A_1 will wait to know whether A_2 succeeded or failed and only then will it continues.

The transformed plan mimics the structure of the original plan, but it includes communications guaranteeing that each role-player knows when a given subplan has succeeded or failed. This common knowledge ensures that the temporal order of the actions and the behavior in response to failure in the original plan are preserved in the execution of the transformed plans. For this a new conjunction has been added to the JINTEND definition. In particular in our example the following will be added to the specification of JINTEND$((A_1, A_2, (A_1, A_2)), p)$:

$$
\begin{aligned}
&\text{BEL}(A_1,\\
&\quad done(!(A_1, transform(a, p))) \wedge done(!(A_2, transform(b, p)))\wedge\\
&\quad done(!((A_1, A_2), transform((a, b), p))) \to done(!((A_1, A_2, (A_1, A_2)), p))) \wedge\\
&\text{BEL}(A_2,\\
&\quad done(!(A_1, transform(a, p))) \wedge done(!(A_2, transform(b, p)))\wedge\\
&\quad done(!((A_1, A_2), transform((a, b), p))) \to done(!((A_1, A_2, (A_1, A_2)), p))) \wedge\\
&\text{MBEL}((A_1, A_2),\\
&\quad done(!(A_1, transform(a, p))) \wedge done(!(A_2, transform(b, p)))\wedge\\
&\quad done(!((A_1, A_2), transform((a, b), p))) \to done(!((A_1, A_2, (A_1, A_2)), p)))
\end{aligned}
$$

Intuitively this means that each agent believes that if each agent and subteam do their part of the plan as specified by the "transform" function (including the broadcast and the waiting parts) then the team will perform the plan. Similarly, the subteam (A_1, A_2) has such a mutual belief.

3.3 SharedPlans

The SharedPlan formalization of collaboration [Grosz and Kraus, 1996; Grosz and Kraus, 1999; Grosz and Kraus, 1993] is based on a mental-state view of plans [Bratman, 1987; Pollack, 1990]: agents are said to have plans when they have a particular set of intentions and beliefs. Both *shared plans* and *individual plans* are considered. Shared plans are constructed by groups of collaborating agents and include subsidiary shared plans [Lochbaum, 1994] formed by subgroups as well as subsidiary individual plans formed by individual participants in the group activity. The formalization distinguishes between complete plans—those in which all the requisite beliefs and intentions have been established—and partial plans.

The plan definitions apply intention operators.

- $\text{Int.To}(G, \alpha, T_i, T_\alpha, \text{con}(\alpha), IC_\alpha)$ represents agent G's intention at time T_i to do action α at time T_α under the constraints $\text{con}(\alpha)$ in the context IC_α.
- $\text{Int.Th}(G, prop, T_i, T_{prop}, \text{IC}_{prop})$ represents an agent's (G) intention at time T_i that a certain proposition *prop* hold in the intentional context IC_{prop} at time T_{prop}.

The significant distinction between intentions-to and intentions-that is not in the types of objects each relates, but in their connection to means-ends reasoning and in their different presumptions about an agent's ability to act in service of the intention. An Int.To commits an agent to means-ends reasoning [Bratman, 1987] and, at some point, to acting. In contrast, an Int.Th does not directly engender such behavior. Int.Ths form the basis for meshing subplans, helping one's collaborators, and coordinating status updates, all of which play an important role in collaborative plans; any of these activities may lead to the adoption of an Int.To and thus indirectly to means-ends reasoning. In addition, an agent can only adopt an Int.To toward an action for which it is the agent and that it believes it can do.

SharedPlan applies also the belief operators:

- $\text{Bel}(G, prop, T_{bel})$ means that G believes at time T_{bel} that *prop* is true.
- $\text{MB}(\text{GR}, prop, T_{bel})$ means that the group GR mutually believe at time T_{bel} that *prop* is true.

In addition, two meta-predicates are defined to represent the beliefs agents have about their own and their collaborators' abilities:

- $\text{CBA}(G, \alpha, R_\alpha, T_\alpha, \Theta)$ ("can bring about") represents an agent's (G) ability to do the action α using the recipe R_α at time T_α and under constraints Θ.

- CBAG(GR, $\alpha, R_\alpha, T_\alpha, \Theta$) (Can bring about group) represents the ability of the group GR to do the action α using the recipe R_α at time T_α and under constraints Θ.

The two main meta-predicates that are defined in the *SharedPlans* formalization are:

- FSP(P, GR, $\alpha, T_p, T_\alpha, R_\alpha, \Theta_\alpha, IC_\alpha$) means that the group GR has at time T_p a full plan P to do the multi-agent action α at time T_α using the recipe R_α under the constraint Θ_α in the context of IC_α.

- PSP(P, GR, $\alpha, T_p, T_\alpha, \Theta_\alpha, IC_\alpha$) means that the group GR has at time T_p a partial plan P to do the multi-agent action α at time T_α under the constraint Θ_α in the context of IC_α.

The meta-predicate FSP is used to represent the situation in which a group of agents has completely determined the recipe by which they are going to do some group activity, and members of the group have adopted Int.Tos toward all the basic-level actions in the recipe as well as Int.Ths toward the actions of the group and its other members. Most typically a group of agents will not have such a complete plan until after they have done some of the actions in the recipe. Groups, like the individual agents they comprise, typically have only partial plans for which the meta-predicate PSP is used. We will also use the meta-predicate FIP, similarly defined, for individual agents.

Now we describe the full SharedPlan of our example. Here α is $Move(\{A_1, A_2\}, Bl, locO, locD)$, and for the rest of this subsection we use α for this specific action. *Recipes*(α) are all the possible recipes for α and R_{Move} is the recipe presented in Section 2 that consists of $\beta_1 - \beta_6$ with the order constraints described there.

Next we describe the constituents of FSP($P, \{A_1, A_2\}, \alpha, t_p, [T, T^e_\alpha], R_{Move}, \theta_\alpha, IC_\alpha$), i.e., the full shared plan P of A_1,A_2 at time t_p to move the block Bl.

0. The group $GR = \{A_1, A_2\}$ has mutual belief that both of them are committed to the success of the group's doing α; in particular each agent has Int.Th at time t_p that they do α in the context of the FSP.

$$\begin{aligned}&\text{MB}(\{A_1, A_2\}, (\forall G \in \{A_1, A_2\})\\ &\quad \text{Int.Th}(G, \text{Do}(\{A_1, A_2\}, \alpha, [t, t + k + 6]),\\ &\qquad\qquad t_p, [t, t + k + 6], FSP(\alpha)), t_p)\end{aligned}$$

1. The group $\{A_1, A_2\}$ has mutual belief that the recipe R_α that they agreed to use is actually a recipe for α.

$\text{MB}(\{A_1, A_2\}, R_\alpha \in \mathit{Recipes}(\alpha), t_p)$

For each of the actions of the recipe R_α, the agents have agreed who will do it and they adopt intentions and form beliefs. There are two cases: single-agent actions and multi-agent actions. In the recipe there are four single-agent actions, $\beta_1 = \mathit{Go}(A_1, \mathit{loc}1, \mathit{WC\_SS}(\mathit{locO}))$, $\beta_2 = \mathit{Go}(A_2, \mathit{loc}2, \mathit{EC\_SS}(\mathit{locO}))$, $\beta_4 = \mathit{Go}(A, \mathit{WC\_SS}(\mathit{locT}), \mathit{NC\_WS}(\mathit{locT}))$, $\beta_5 = \mathit{Go}(A_2, \mathit{EC\_SS}(\mathit{locT}), \mathit{SC\_WS}(\mathit{locT}))$ and two multi-agent actions $\beta_3 = \mathit{Push}(\mathit{GR}, \mathit{Bl}, \mathit{locO}, \mathit{locT})$ and $\beta_6 = \mathit{Push}(\mathit{GR}, \mathit{Bl}, \mathit{locT}, \mathit{locD})$.

We first consider one of the single-agent actions, β_1 only; the others are similar.

- A_1 has the intention to do β_1 at time $T_{\beta_1} = [t, t{+}k_1]$ under the relevant constraints from the recipe, i.e., $\{\mathit{Before}(\beta_1, \beta_3)\}$. Its context for doing it is the FSP of α. Formally,
  $\text{Int.To}(A_1, \beta_1, t_p, T_{\beta_1}, \{\mathit{Before}(\beta_1, \beta_3)\}, \mathit{FSP}(\alpha))$.
- Both agents have the mutual belief that A_1 intends to do β_1
  $\text{MB}(\{A_1, A_2\}, \text{Int.To}(A_1, \beta_1, t_p, T_{\beta_1}, \{\mathit{Before}(\beta_1, \beta_3)\}, \mathit{FSP}(\alpha)), t_p)$.
- Both agents mutually believe that A_1 can do β_1. As β_1 is a basic-level action, R_{β_1} is empty.
  $\text{MB}(\{A_1, A_2\}, \text{CBA}(A_1, \beta_1, R_{\mathit{Empty}}, T_{\beta_1}, \{\mathit{Before}(\beta_1, \beta_3)\}), t_p)$.
- A_2 is committed to the success of A_1 to perform β_1 by having an Int.Th that it will be able to perform it. This intention is mutually believed by the two agents.
  $\text{MB}(\{A_1, A_2\},$
  $\quad \text{Int.Th}(A_2, (\exists R_{\beta_1})\text{CBA}(A_1, \beta_1, R_{\beta_1}, T_{\beta_1},$
  $\qquad \{\mathit{Before}(\beta_1, \beta_3)\}), t_p, T_{\beta_1}, \beta_1), t_p)$
  If β_1 is a basic level action, R_{β_1} is empty.
- If β_1 were not a basic level action, A_1 should have a full individual plan for doing it and both agents should have mutual belief about it.
  - $\neg\text{basic.level}(\beta_1)$
  - There should be a recipe for the *Go* action, R_{β_1}, and full individual plan P_{β_1} of A_1 to do β_1 at T_{β_1}. The context is the full shared plan of α and the action should be done under the constraint $\mathit{Before}(\beta_1, \beta_3)$. Formally,
    $(\exists P_{\beta_1}, R_{\beta_1})$
    $\text{FIP}(P_{\beta_1}, A_1, \beta_1, t_p, T_{\beta_1}, R_{\beta_1}, \{\mathit{Before}(\beta_1, \beta_3)\}, \mathit{FSP}(\alpha))$.

– A_1 and A_2 mutually believe that such a plan and a recipe exists and that A_1 can bring about the performance of α according to the recipe. Note that A_2 does not need to know the recipe nor the details of the plan. It just needs to believe that A_1 has them.

$$\text{MB}(\{A_1, A_2\}, (\exists P_{\beta_1}, R_{\beta_1}) \\ [\text{CBA}(A_1, \beta_1, R_{\beta_1}, T_{\beta_1}, \{Before(\beta_1, \beta_3)\}) \wedge \\ \text{FIP}(P_{\beta_1}, A, \beta_1, t_p, T_{\beta_1}, R_{\beta_1}, \{Before(\beta_1, \beta_3)\}, FSP(\alpha))], t_p).$$

The clauses associated with the other single-agent actions of the recipe R_α, namely β_2 β_4 β_5, are similar to those of β_1. The intentions and beliefs associated with the multi-agent action $\beta_3 = Push(\{A_1, A_2\}, Bl, locO, locT)$ is different from that of the single-agent actions. They include the following:

- There should be a full SharedPlan, P_{β_3}, for the agents to do the Push using a recipe R_{β_3}. This recipe consists of $\gamma_3^1 = LPush(A_1, Bl, locX, locY)$ and $\gamma_3^2 = RPush(A_2, Bl, locX, locY)$ and the constraint $\{Sim(\gamma_i^1, \gamma_i^2)\}$. The relevant constraints for the Push are $\{Before(\beta_3, \beta_4), Before(\beta_3, \beta_5), Before(\beta_1, \beta_3), Before(\beta_2, \beta_3)\}$ and the context is the full SharedPlans for α:
$$(\exists P_{\beta_3}, R_{\beta_3})\ \text{FSP}(P_{\beta_3}, \{A_1, A_2\}, \beta_3, t_p, T_{\beta_3}, R_{\beta_3}, \{Before(\beta_3, \beta_4), \\ Before(\beta_3, \beta_5), Before(\beta_1, \beta_3), Before(\beta_2, \beta_3)\}, \text{FSP}(\alpha)).$$

- A_1 and A_2 mutually believe that they can perform β_3 and that they have a full SharedPlan to do it. Note that if the group for the Move included more agents, the entire group would have had this mutual belief. In our case since the subgroup and the entire group are the same, this clause can be inferred from the previous one.
$$\text{MB}(\{A_1, A_2\}, (\exists P_{\beta_3}, R_{\beta_3}) \\ [\text{CBAG}(\{A_1, A_2\}, \beta_3, R_{\beta_3}, T_{\beta_3}, \{Before(\beta_3, \beta_4), Before(\beta_3, \beta_5), \\ Before(\beta_1, \beta_3), Before(\beta_2, \beta_3)\}) \wedge \\ \text{FSP}(P_{\beta_3}, \{A_1, A_2\}, \beta_3, t_p, T_{\beta_3}, R_{\beta_3}, \{Before(\beta_3, \beta_4), Before(\beta_3, \beta_5), \\ Before(\beta_1, \beta_3), Before(\beta_2, \beta_3)\}, FSP(\alpha))], t_p)$$

- If there had been additional members in the group, they should have had Int.Th that the subgroup will succeed in performing β_3.

An important aspect of the *SharedPlans* model is that it motivates helpful behavior of the group members. If something goes wrong and one of the agents is not able to perform its action, other members in the group will consider helping. For example, if A_1 is running out of fuel and is not able to perform a *Go* action, A_2 will try to help it. This is motivated by the Int.Th of A_2 that A_1 will succeed in performing the *Go*. In addition, this Int.Th (via axioms included in the *SharedPlans* model) also motivates the following: if A_2 is able to reduce the costs of A_1 doing the Go action, without bearing a big loss, it will help A_1 too. For example, if A_2 finds out that there is an obstacle in A_1's way, it will let A_1 know.

3.4 Cooperative Problem Solving

The paper [Wooldridge and Jennings, 1999] provides a model of cooperative problem solving starting with the recognition of the potential for cooperation to the team action itself. The notation is based on a branching-time tree formed of states connected by arcs representing primitive actions. Paths through the tree can be characterized by sequences of actions. Terms may denote agents, groups of agents, sequences of actions, and other objects in the environment. Action expressions are formed by using ";" for sequencing, "—" for non-deterministic choice, "*" for iteration, and "?" for test actions.

The important modal operators are as follows:

- $Bel(a, \phi)$ means that agent a has belief ϕ.
- $Goal(a, \phi)$ means that agent a has goal ϕ.
- $Agts(\alpha, g)$ means that g is an agent group required to do α.
- $Agt(\alpha, a)$ means that a is the only agent for α.
- $Happens(\alpha)$ means that action α is the first thing that happens on a given path.
- $A\ \phi$ means that ϕ holds on all paths (inevitably ϕ).

The following are important examples of some of the operators that can be defined using these concepts.

- Mutual belief, *M-Bel* is defined using the everyone believes, *E-Bel*, operator as follows:
  $E\text{-}Bel(g, \phi, 0) =_{def} \phi$
  $E\text{-}Bel(g, \phi, \ell + 1) =_{def} \forall a(a \in g \rightarrow Bel(a, E\text{-}Bel(g, \phi, \ell)))$, where ℓ represents the level of nesting for *E-Bel*
  $M\text{-}Bel(a, \phi) =_{def} E\text{-}Bel(g, \phi, \ell)$ for all levels ℓ.
- $Does(\alpha) =_{def} A\ Happens(\alpha)$ meaning that α happens on all paths.
- $Achieves(\alpha, \phi) =_{def}$
  $E(Happens(\alpha)) \wedge A((Happens(\alpha) \rightarrow Happens(\alpha; \phi?))$
  meaning that α achieves the goal ϕ if α happens on some path and on all paths if α happens then ϕ is true afterwards. (Note: E is the dual of A.)

- $Able_1(\alpha, \phi) =_{def} \exists\alpha(Bel(a, (Agt(\alpha, a) \wedge Achieves(\alpha, \phi)) \wedge$
  $Agt(\alpha, a) \wedge Achieves(\alpha, \phi))$
  An agent has type 1 ability to achieve ϕ if it believes that some action can achieve ϕ and it can do the action.
  $Able(a, \phi) =_{def} Able_1(a, \phi) \vee Able_1(a, Able_1(a, \phi))$
  An agent has the ability to achieve ϕ if either it has type 1 ability for it or it has type 1 ability to bring about a situation where it has type 1 ability to achieve it.

- $J\text{-}Able_1(g, \phi) =_{def} \exists\alpha$
  $(M\text{-}Bel(g, Agts(\alpha, g) \wedge Achieves(\alpha, \phi)) \wedge Agts(\alpha, g) \wedge Achieves(\alpha, \phi))$
  Type 1 joint ability of a group of agents g means that there is an action that achieves ϕ that the group is required to do and this is mutually believed.
  $J\text{-}Able(g, \phi) =_{def} J\text{-}Able_1(g, \phi) \vee J\text{-}Able_1(g, J\text{-}Able_1(g, \phi))$
  Joint ability is obtained from type 1 joint ability analogously as ability is obtained from type 1 ability.

- $J\text{-}Commit(g, \phi, \psi, \chi, c)$ has a fairly complex definition. It defines joint commitment with the parameters g for the group, ϕ for the goal, ψ for the motivation, χ for the pre-condition, and c for the convention. We do not give the details here.

- $J\text{-}Intend(g, \alpha, \psi) =_{def} M\text{-}Bel(g, Agts(\alpha, g)) \wedge$
  $J\text{-}Commit(g, A \diamond Happens(M{-}Bel(g, A\, Happens(\alpha)); \alpha), \psi, \chi_{soc}, c_{soc})$
  J-Intend is the joint intention of the group g with respect to action α and motivation ψ. Here $\diamond\, \phi$ means that ϕ is eventually satisfied; χ_{soc} and c_{soc} involve social pre-conditions and conventions. We can read joint intention to mean that the group g has a joint commitment that eventually g will believe that α will happen next, and then α happens next.

- $Team(g, \phi, , a) =_{def}$
  $\exists\alpha(M\text{-}Bel(g, Achieves(\alpha, \phi)) \wedge J\text{-}Intend(g, \alpha, Goal(a, \phi))$
  The agents g form a team to accomplish agent a's goal ϕ if the agents in g mutually believe that some action sequence α achieves ϕ and they jointly intend to do it.

We did not include here the details of various other steps including individual commitments, intentions, and the potential for cooperation.

Consider now how our example can be formalized using these concepts. Initially an agent, say A_1, is presented with the goal of having the block Bl at location $locD$. We can write this as $Goal(A_1, At(Bl, locD))$. As

[Wooldridge and Jennings, 1999] does not explicitly deal with parallel actions, we add parallelism by using the extra symbol &. There is also no discussion of complex subactions, so we consider only basic-level actions and write $\alpha = (\beta_1 \& \beta_2); (\gamma_3^1 \& \gamma_3^2); (\beta_4 \& \beta_5); (\gamma_6^1 \& \gamma_6^2)$. This formalism also does not have explicit time representation. In this case the group $g = (A_1, A_2)$. So we would get to $M\text{-}Bel((A_1, A_2), Achieves(\alpha, At(Bl, locD)))$. The preconditions here include $At(A_1, loc1)$, $At(A_2, loc2)$, $At(Bl, locO)$. If everything goes well, we get to $Team((A_1, A_2), At(Bl, locD), \alpha)$ at which point A_1 and A_2 form a team to apply the sequence of actions α to move the block from $locO$ to $locD$. There will also be individual intentions to do the various subactions: for example A_1 to do β_1, A_2 to do β_2, and so on.

3.5 Collective Intentions

In a series of three papers, [Dunin-Keplicz and Verbrugge, 2002], [Dunin-Keplicz and Verbrugge, 2003], and [Dunin-Keplicz and Verbrugge, 2004] Dunin-Keplicz and Verbrugge formally characterize the concept of a team having a collective intention and collective commitment towards a common goal. Many aspects of teamwork get a thorough formal treatment. They give a modal logic axiomatization for the predicates and prove the completeness of an important portion of their formalization, collective intentions, by the use of Kripke models.

The language contains propositional symbols, symbols for agents and atomic actions, and a large number of predicate symbols. Individual actions are built up from atomic actions by using sequential composition: $(\alpha_1; \alpha_2)$, non-deterministic choice: $(\alpha_1 \cup \alpha_2)$, iteration: (α^*), as well as several other predicate symbols. Using individual actions, social plan expressions are formed by writing $< \alpha, a >$ (agent a associated with action α), $stit(G, \phi)$ (group G sees to it that ϕ), $confirm(\phi)$, as well as by combining social plan expressions Γ and Δ as $< \Gamma; \Delta >$ (sequential composition) and $< \Gamma || \Delta >$ (parallelism).

Various modalities are also defined. A partial list includes:

epistemic - versions of belief

motivational - versions of goals, intentions and commitments

temporal - versions of $done$, $succeeded$, $failed$

abilities and opportunities - $able(a, \alpha)$, $opp(a, \alpha)$

Now we list some of the important epistemic and motivational predicate symbols with their meanings; however we change the subscripts to parameters:

- $BEL(a, \phi)$ means that agent a believes ϕ.
- *E-BEL*(G, ϕ) means that every agent in group G believes ϕ.
- *C-BEL*(G, ϕ) means that group G collectively believes ϕ.
- $GOAL(a, \phi)$ means that agent a has goal ϕ.
- $INT(a, \phi)$ means that agent a has the intention to make ϕ true.
- *E-INT*(G, ϕ) means that every agent in G has an individual intention to make ϕ true.
- *M-INT*(G, ϕ) means that G has a mutual intention to make ϕ true.
- *C-INT*(G, ϕ) means that G has a collective intention to make ϕ true.
- $COMM(a, b, \phi)$ means that agent a commits to agent b to make ϕ true.
- *R-COMM*(G, P, ϕ) means that group G has a robust collective commitment to achieve ϕ by plan P.
- *S-COMM*(G, P, ϕ) means that group G has a strong collective commitment to achieve ϕ by plan P.
- *W-COMM*(G, P, ϕ) means that group G has a weak collective commitment to achieve ϕ by plan P.
- *T-COMM*(G, P, ϕ) means that group G has a team commitment to achieve ϕ by plan P.
- *D-COMM*(G, P, ϕ) means that group G has a distributed commitment to achieve ϕ by plan P.

E-BEL refers to the individual belief of each agent in the group. Collective belief, *C-BEL*, is a recursive extension of *E-BEL*: every agent believes that every agent believes $\ldots\ \phi$. For intentions, again *E-INT* refers to the individual intention of each agent in a group. Mutual intention, *M-INT*, is a recursive extension of *E-INT*: every agent intends that every agent intends $\ldots\ \phi$. Collective intention, *C-INT*, extends mutual intention by requiring also the collective belief of the group in the mutual intention. Commitment, $COMM$, involves an intention, a goal, and a collective belief in the intention and the goal. The various types of group commitments differ in some subtle details involving differences in the awareness of the responsibilities of the various agents in the group. The group commitments listed above go from the strongest to the weakest versions. For example, in the strongest version, robust commitment, essentially every agent is aware of all the intentions and commitments of all the other agents; while in the weakest version, distributed commitment, there is not even a collective intention and agents are aware only of their subactions.

To indicate the approach of the papers, which rely heavily on axiomatization of the concepts, we include several representative axioms. For introspection axioms we give only the positive versions; there are also negative introspection axioms.

A4 $BEL(a, \phi) \rightarrow BEL(a, BEL(a, \phi))$ (positive introspection)

R2 From ϕ infer $BEL(a, \phi)$ (all agents believe the true propositions)

C1 $E\text{-}BEL(G, \phi) \leftrightarrow \bigwedge_{a \in G} BEL(a, \phi)$ (the meaning of *E-BEL*)

C2 $C\text{-}BEL(G, \phi) \leftrightarrow E\text{-}BEL(G, \phi \wedge C\text{-}BEL(G, \phi))$ (collective belief in terms of *E-BEL*)

A2$_G$ $GOAL(a, \phi) \wedge GOAL(a, \phi \rightarrow \psi) \rightarrow GOAL(a, \psi)$ (goal distribution)

A7$_{GB}$ $GOAL(a, \phi) \rightarrow BEL(a, GOAL(a, \phi))$ (positive introspection for goals)

A2$_I$ $INT(a, \phi) \wedge INT(a, \phi \rightarrow \psi) \rightarrow INT(a, \psi)$ (intention distribution)

A7$_{IB}$ $INT(a, \phi) \rightarrow BEL(a, INT(a, \phi))$ (positive introspection for intentions)

A9$_{IG}$ $INT(a, \phi) \rightarrow GOAL(a, \phi)$ (intention implies goal)

M1 $E\text{-}INT(G, \phi) \leftrightarrow \bigwedge_{a \in G} INT(a, \phi)$ (the meaning of *E-INT*)

M2 $M\text{-}INT(G, \phi) \leftrightarrow E\text{-}INT(G, \phi \wedge M\text{-}INT(G, \phi))$ (mutual intention in terms of *E-INT*)

M3 $C\text{-}INT(G, \phi) \leftrightarrow M\text{-}INT(G, \phi) \wedge C\text{-}BEL(G, M\text{-}INT(G, \phi))$ (collective intention in terms of collective belief and mutual intention)

SC1 $COMM(a, b, \phi) \leftrightarrow INT(a, \phi) \wedge GOAL(b, stit(a, \phi))$ $\wedge C\text{-}BEL((a, b), INT(a, \phi) \wedge GOAL(b, stit(a, \phi)))$ (commitment for a proposition in terms of intentions, goals, and collective belief)

SC2 $COMM(a, b, \alpha) \leftrightarrow INT(a, \alpha) \wedge GOAL(b, done(a, \alpha)) \wedge$ $C\text{-}BEL((a, b), INT(a, \phi) \wedge GOAL(b, done(a, \alpha)))$ (commitment for an action in terms of intentions, goals, and collective belief)

Consider now how our example can be formalized in this framework. A social plan for it can be expressed as

$P =<<<<< \beta_1, A_1 > || < \beta_2, A_2 >>; << \gamma_3^1, A_1 > || < \gamma_3^2, A_2 >>>;$
$<< \beta_4, A_1 > || < \beta_5, A_2 >>>; << \gamma_6^1, A_1 > || < \gamma_6^2, A_2 >>>$

This will lead to the various individual commitments, such as $COMM(A_1, A_2, \beta_1)$, $COMM(A_2, A_1, \beta_2)$, and so on and at some point to the group commitment, probably *S-COMM*$((A_1, A_2), P, \alpha)$ in this case. The execution of α will follow the commitments; there is no explicit time representation. The last two papers go into substantial detail about what happens if some action fails during execution.

3.6 Cooperative Subcontracting

Grant, Kraus, and Perlis [Grant *et al.*, 2005; Grant *et al.*, 2002] provide a formal theory for the handling of agent intentions in a framework that allows multiple agents doing various tasks together. We briefly describe those aspects of the language that are needed for the formalization of our example; we omit many important details, such as how agents communicate. There are, in fact, several languages. The part of the language described here is really a metalanguage that has names for formulas in the language of the agents. This (meta)language is a sorted first-order language with different sorts for time, agents, actions, recipes, and formulas. To conform to the example given for this paper in the different formalisms we slightly change the symbols used in the original paper. The following symbols are used: for time: t, i, j, k; for agents: A_i; for actions: α, etc.; for recipes: r; for formulas: f.

Actions are either basic-level or complex; for a complex action there is always a recipe.

- $BL(\alpha, d)$ means that α is basic level and takes d units of time.
- $Rec(\alpha, r)$ means that α has recipe r. We only deal with the single recipe case here.
- $Mem(r, \beta, i, j, k)$ means that in recipe r the subaction β is one of the ith subactions (there may be more than one) starting at time j and ending at time k. These times are offsets relative to the beginning of the action.

Agents have beliefs, such as that they can do an action.

- $Bel(t, A_i, f)$ means that at time t agent A_i believes in the statement expressed by the formula f.
- $CanDo(t, A_i, \alpha)$ means that at time t agent A_i can do (at least start to do) action α.

We are primarily interested in the concept of intention. There are three main predicates for this.

- $ATD(t, A_i, A_j, \beta, \alpha, t')$ means that at time t agent A_i asks agent A_j to do action β as a subaction of action α at time t'.

- $PotInt(t, A_i, A_j, \beta, \alpha, t')$ means that at time t agent A_i directly assisting agent A_j has the potential intention to do action β as a subaction of action α at time t'.

- Int is defined exactly as $PotInt$, but it stands for the actual intention.

The axioms of this system show what happens in a time step given various initial conditions; for example, how an agent gains potential intentions and intentions.

The general process is as follows. The agent is asked to do an action at a particular time. This triggers a potential intention for the agent. This potential intention may lead to an intention assuming certain conditions are satisfied. In particular, the agent must believe that it can do the action at the right time, possibly with help from other agents. The parts of the action the agent cannot do, it can subcontract to other agents. A subcontracting means that an agent is asked to do an action leading again to a potential intention. The time for each action is explicitly represented by a time value. The recipes for the actions are known by the agents, as needed.

In our example, the action α is the action of moving the block from *locO* to *locD*. There are two agents, A_1 and A_2. We must use subcontracting for the joint action; hence initially A_1 is asked to do this move and A_1 must subcontract some subactions to A_2. Some of the actions are done in parallel. Since our system requires that the first subaction of every action must be a single action, we add an essential dummy first subaction, as needed. In the case of the agents going to the block, this first subaction may be "prepare to go".

Later we will only sketch the attainment of intentions and the actual actions, but for clarity we write the recipe out in detail.

- $Rec(\alpha, r)$: r is the recipe for α: "move the block from *locO* to *locD*".

- $Mem(r, \beta_1, 1, 1, 1)$: The "prepare to go" subaction starts and ends at time 1.

- $Mem(r, \beta_{21}, 2, 2, k_1 + 2)$: The "go to *locO*" subaction for agent A_1 takes k_1 time steps.

- $Mem(r, \beta_{22}, 2, 2, k_2 + 2)$: The "go to *locO*" subaction for agent A_2 takes k_2 time steps. Note that both β_{21} and β_{22} are considered the second subaction. Next recall that $k = max(k_1, k_2)$.
- $Mem(r, \beta_3, 3, k+3, k+6)$: The "push" subaction is a complex action.
- $Mem(r, \beta_{41}, 4, k+7, k+8)$: The first agent goes to the new position.
- $Mem(r, \beta_{42}, 4, k+7, k+8)$: The second agent goes to the new position.
- $Mem(r, \beta_3, 5, k+9, k+12)$: Another "push" is done.
- $Rec(\beta_3, r')$: r' is the recipe for "push".
- $Mem(r', \gamma_1, 1, 1, 1)$: The "prepare to push" subaction is analogous to the "prepare to go" subaction.
- $Mem(r', \gamma_{21}, 2, 2, 3)$: The "push from left" subaction takes one time step.
- $Mem(r', \gamma_{22}, 2, 2, 3)$: The "push from right" subaction takes one time step.

Next we briefly sketch some of the steps involved in the getting and keeping of intentions by the two agents. In our system we allow the retraction of intentions for various reasons but assume it doesn't happen here. In order to indicate clearly the time value, we write it first separately and leave it out of the formula itself. We intersperse the timed actions with various explanations.

time 0 $ATD(\_, A_1, \alpha, \_, t)$
A_1 is asked to do action α at time t.

time 1 $PotInt(A_1, \_, \alpha, \_, t)$
$Bel(A_1, "CanDo(t, A_1, \alpha)")$
The ATD triggers a $PotInt$ for A_1 to do action α at time t. We also assume that A_1 believes it can do α, possibly with help from other agents. This belief triggers potential intentions for the subactions of α while the potential intention for α is inherited.

time 2 $PotInt(A_1, \_, \alpha, \_, t)$
$PotInt(A_1, A_1, \beta_1, \alpha, t+1)$
$Bel(A_1, "CanDo(t+1, A_1, \beta_1)")$
The first formula shows the inheritance of $PotInt$. The second formula is the $PotInt$ for β_1. There is a similar $PotInt$ for all the βs. The third formula shows A_1's belief about β_1.

time 3 $Int(A_1, A_1, \beta_1, \alpha, t+1)$
$ATD(A_1, A_2, \beta_{22}, \alpha, t+2)$
The first formula gives the first intention for A_1. A similar intention is obtained for all subactions that A_1 believes it can do. The second formula shows A_1 asking A_2 to do β_{22}. This is a subaction A_1 cannot do. Similarly A_1 asks A_2 to do β_{42}.

time 4 $PotInt(A_2, A_1, \beta_{22}, \alpha, t+2)$
$Int(A_1, A_1, \gamma_1, \beta_3, t+k+4)$
$ATD(A_1, A_2, \gamma_{22}, \beta_3, t+k+5)$
The first formula shows that A_2 obtains a potential intention for β_{22}. The second formula shows the intention of A_1 to do γ_1. A_1 must ask A_2 to do γ_{22}. Continuing this process we just show the main formula at time 8.

time 8 $Int(A_1, \_, \alpha, \_, t)$
This is the point where A_1 gets an intention for α as all subactions have been assigned and for each subaction of α either A_1 or A_2 has an intention to do it.

Finally, we sketch the initiation of the first subactions:

time t $Ini(A_1, \_, \alpha)$
Given the continuing intention of A_1 to do α, then at time t A_1 actually initiates the action.

time t+1 $Ini(A_1, A_1, \beta_1)$
$Bel(A_1, "Done(t+1, \beta_1)")$
A_1 initiates β_1 which is just the preparation subaction and gets done immediately.

time t+2 $Ini(A_1, A_1, \beta_{21})$
$Ini(A_2, A_1, \beta_{22})$
A_1 initiates its subaction β_{21} and A_2 initiates its subaction β_{22}.

Assuming the agents can do all the subactions at the right time, they all get done. We do not show the details.

4 Comparison

There are many ways in which the formal approaches presented in the previous section can be compared. We have chosen several criteria for comparison and present it here. We do not claim that this is an exhaustive list.

We start by considering the focus of these approaches. We have found that this difference in focus leads in a natural way to differences in the way that different authors handle many issues. The focus of Joint Intentions is the persistence of joint intentions. The focus of Team Plans is on the complexity of actions and the roles of agents. SharedPlans emphasizes individual intentions and actions needed for teamwork. Cooperative Problem Solving emphasizes joint commitment to a goal and cooperation. Collective Intentions emphasizes the formalization of collective intentions in as multi-modal logical framework that allows for a proof of the completeness of the logic. Cooperative Subcontracting focuses on subcontracting between cooperative agents.

Consider now the formalisms themselves. Along this dimension we can distinguish approaches that use a modal logic and ones that have a syntactic approach. The modal logic approach uses various modalities in the syntax and Kripke models for the semantics. Typically, in dealing with plans and actions, some aspects of dynamic logic are included. Among the six systems, Joint Intentions, Team Plans, Cooperative Problem Solving, and Collective Intentions use the modal logic approach. SharedPlans and Cooperative Subcontracting both use the syntactic approach; the latter also deals with semantics via minimal models.

Another issue separating the various formalizations is whether in the system teamwork is based on the intentions and actions of single agents (working together) or primarily on mutual beliefs and joint intentions. This is basically a matter of degree because in all cases at some point agents must have their own beliefs and intentions. Here we find that Joint Intentions, Cooperative Problem Solving, and Collective Intentions emphasize collective/joint intentions, while Team Plans, SharedPlans, and Cooperative Subcontracting have more emphasis on the intentions of single agents.

We can also consider the representation of time. Joint Intentions, SharedPlans, and Cooperative Subcontracting allow for the explicit representation of time in the language. Team Plans has time in the semantics in the form of a time tree. Cooperative Problem Solving also does not have explicit time values in the syntax and deal with it through path formulas, expressing properties of a single path through a branching time structure. Collective Intentions does not deal with time.

Systems differ also in their facility for expressing complex plans. Team Plans places great emphasis on this issue. SharedPlans and Cooperative Subcontracting also possess specific constructs to express complex plans. Joint Intentions, Cooperative Problem Solving, and Collective Intentions do not deal with this issue.

Now we recall the four stages of cooperative behavior as given in the Introduction from the Cooperative Problem Solving Approach: recognition for the possibility of cooperation, team formation, plan formation, and execution. We modify some of the explanation there to provide a more comprehensive formulation of these four steps. Recognition is about agent abilities and beliefs, as well as the potential for cooperation. Team formation involves mutual beliefs, intentions, and commitments. Plan formation consists of possibly negotiation between agents as well as formulating the actions and their sequence that need to be done. Finally, team action should be the actual activity of the agents in accomplishing their goal. Using these definitions, we find that except for Joint Intentions, which emphasizes mainly team formation, the other approaches appear to contain at least some aspects of each of the four stages.

We also recall that Bratman [1992] gives four criteria that a multi-agent activity must meet to be a "shared cooperative activity": mutual responsiveness, commitment to the joint activity, commitment to mutual support, and formation of subplans that mesh with one another. All works consider the issue of the commitment to the joint activity. The requirements from an agent to be considered being committed to the team activity varied from one model to the other. With respect to the other requirements, SharedPlans considers all the other three of Bratman's requirement explicitly (Section 7 of [Grosz and Kraus, 1996]). Team Plans studies carefully the issue of coordination and synchronization of team activities and thus handles the issue of meshing subplans. It also handles mutual responsiveness in action, namely the responsiveness required when plans must be modified to cope with problems in execution. This may also lead to mutual support, but the details are left to future work (Section 5.3 of [Sonenberg *et al.*, 1992]). Collective Intentions provides a formal model for the dynamic change in collective commitments and intentions and thus handles mutual responsiveness and, in some ways, the meshing of subplans and mutual support. The original Joint Intentions paper [Levesque *et al.*, 1990] considers only the issue of the commitments to the joint activity and defines how agents disband a team. In a later paper, [Cohen *et al.*, 1997], issues such as commitment to mutual support and mutual responsiveness are discussed. The Cooperative Problem Solving considers implicitly the issue of meshing subplans and mutual support in their plan formation process. Cooperative Subcontracting deals with mutual responsiveness and with the meshing of subplans both in the planning phase and in the execution phase.

5 Conclusions and Future Work

We compared six approaches to teamwork from a substantial literature on the subject. Our primary technique was to see how these approaches handle a specific task of two agents pushing a block. Teamwork is a highly complex issue and we used the example to illustrate some important aspects of it.

We did not deal with some other issues associated with teamwork such as the details of preconditions, communication between agents, and intention reconsideration. Our examination of the six formalisms with respect to our example of two cooperating agents indicates that there are (at least) two sub-areas of emphasis in the literature: formation of the preconditions for cooperative action, and enactment based upon the cooperative preconditions. Some approaches tend to cover both of these, but none is entirely adequate for an end-to-end treatment that can describe the entire agent processing. Clearly, there is room for more work, especially integrative work; in particular, two interrelated components in need of more attention are explicit representations of time and primitive actions.

Acknowledgement

We wish to thank Artur Garcez for many helpful comments. This research was supported by NSF under grant IIS-0222914 and by AFOSR and ONR.

BIBLIOGRAPHY

[Alonso, 1997] E. Alonso. A formal framework for the representation of negotiation protocols, 1997.

[Bratman, 1987] M. E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, 1987.

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

[Castelfranchi, 1995] C. Castelfranchi. Commitments: From individual intentions to groups and organizations. In *ICMAS 95*, 1995.

[Cohen and Levesque, 1990] P. R. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:263–310, 1990.

[Cohen and Levesque, 1991a] P. R. Cohen and H. J. Levesque. Confirmation and joint action. In *Proce. of IJCAI 1991*, 1991.

[Cohen and Levesque, 1991b] P. R. Cohen and H. J. Levesque. Teamwork. *Noûs*, pages 487–512, 1991.

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

[Cohen *et al.*, 1997] P. R. Cohen, H. J. Levesque, and I. Smith. On team formation. In J. Hintikka and R. Tuomela, editors, *Contemporary Action Theory Synthese.* 1997.

[Dubois and Prade, 1994] D. Dubois and H. Prade. A survey of belief revision and updating rules in various uncertainty models. *International Journal of Intelligent Systems*, 9:61–100, 1994.

[Dunin-Keplicz and Verbrugge, 2002] B. Dunin-Keplicz and R. Verbrugge. Collective intentions. *Fundamenta Informatica*, 49:271–295, 2002.

[Dunin-Keplicz and Verbrugge, 2003] B. Dunin-Keplicz and R. Verbrugge. Evolution of collective commitments during teamwork. *Fundamenta Informaticae*, 56(4):329–371, 2003.

[Dunin-Keplicz and Verbrugge, 2004] B. Dunin-Keplicz and R. Verbrugge. A tuning machine for cooperative problem solving. *Fundam. Inform.*, 63(2-3):283–307, 2004.

[Fagin *et al.*, 1995] R. Fagin, J. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge.* MIT Press, 1995.

[Gmytrasiewicz *et al.*, 1992] P. Gmytrasiewicz, E. H. Durfee, and D. Wehe. A logic of knowledge and belief for recursive modeling: Preliminary report. In *Proc. of AAAI-92*, pages 628–634, California, 1992.

[Grant *et al.*, 2000] J. Grant, S. Kraus, and D. Perlis. A logic for characterizing multiple bounded agents. *Autonomous Agents and Multi-Agent Systems Journal*, 3(4):351–387, 2000.

[Grant *et al.*, 2002] J. Grant, S. Kraus, and D. Perlis. A logic-based model of intentions for multi-agent subcontracting. In *Proc. of AAAI-2002*, pages 320–325, 2002.

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

[Grosz and Kraus, 1993] B. J. Grosz and Sarit Kraus. Collaborative plans for group activities. In *Proceedings of IJCAI-93*, pages 367–373, Chambery, France, August 1993.

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

[Grosz and Kraus, 1999] B. J. Grosz and S. Kraus. The evolution of sharedplans. In A. Rao and M. Wooldridge, editors, *Foundations and Theories of Rational Agency*, pages 227–262. Kluwer Academic Publishers, 1999.

[Haass, 1983] A. Haass. The syntactic theory of belief and knowledge. *Artificial Intelligence*, 28(3):245–293, 1983.

[Jennings, 1995] N. R. Jennings. Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence Journal*, 75(2):1–46, 1995.

[Kinny *et al.*, 1994] 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 (LNAI-830)*, Amsterdam, The Netherlands, 1994. Springer Verlag.

[Konolige and Pollack, 1993] K. Konolige and M. E. Pollack. A representationalist theory of intention. In *Proc. of IJCAI-93*, pages 390–395, Chambery, France, August 1993.

[Konolige, 1986] K. Konolige. *A Deduction Model of Belief.* Pitman, London, 1986.

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

[Kraus *et al.*, 1998] S. Kraus, K. Sycara, and A. Evenchik. Reaching agreements through argumentation: a logical model and implementation. *Artificial Intelligence*, 104(1-2):1–69, 1998.

[Kumar *et al.*, 2000] S. Kumar, M. J. Huber, D.R. McGee, P.R. Cohen, and H.J. Levesque. Semantics of agent communication languages for group interaction. In *AAAI 2000*, pages 42–47, 2000.

[Levesque *et al.*, 1990] H. J. Levesque, P. R. Cohen, and J. Nunes. On acting together. In *Proceedings of AAAI-90*, pages 94–99, Boston, MA, July 1990.

[Levesque, 1984] H. J. Levesque. A logic of implicit and explicit belief. In *Proc. of AAAI-84*, pages 198–202, Austin, TX, 1984.

[Lochbaum, 1994] K. Lochbaum. *Using Collaborative Plans to Model the Intentional Structure of Discourse*. PhD thesis, Harvard University, 1994. Available as Tech Report TR-25-94.

[Megiddo, 1989] N. Megiddo. On computable beliefs of rational machines. *Games and Economic Behavior*, 1:144–169, 1989.

[Moore, 1985] R. Moore. A formal theory of knowledge and action. In J. Hobbs and R. Moore, editors, *Formal Theories of the Commonsesnse World*. ABLEX publishing, Norwood, N.J., 1985.

[Perrault, 1990] R. Perrault. An application of default logic to speech act theory. In P.R. Cohen, J.L. Morgan, and M.E. Pollack, editors, *Intentions in Communication*, pages 161–185. Bradford Books at MIT Press, 1990.

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

[Pollack, 1990] M. E. Pollack. Plans as complex mental attitudes. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 77–103. MIT Press, 1990.

[Rao and Georgeff, 1991a] A. Rao and M. Georgeff. Deliberation and its role in the formation of intention. In *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*, San Mateo, California, 1991. Morgan Kaufmann Publishers, Inc.

[Rao and Georgeff, 1991b] A. Rao and M. P Georgeff. Asymmetry thesis and side-effect problems in linear-time and branching-time intention logics. In *Proc. of IJCAI-91*, pages 498–504, Sydney, Australia, August 1991.

[Rao *et al.*, 1992] A. Rao, M. P. Georgeff, and E. A Sonenberg. Social plans: A preliminary report. In *Decentralized Artificial Intelligence, Volume 3*, pages 57–76. Elsevier Science Publishers, 1992.

[Schut *et al.*, 2001] M. Schut, M. Wooldridge, and S. Parsons. Reasoning about intentions in uncertain domains. In *Proceedings of the Sixth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU-2001)*, Toulouse, France, September 2001.

[Searle, 1990] J. R. Searle. Collective intentions and actions. In *Intentions in Communication*, chapter 19. The MIT Press, 1990.

[Shoham, 1989] Y. Shoham. Belief as defeasible knowledge. In *Proc. Eleventh Int'l Joint Conf. on Artificial Intelligence*, Detroit, MI, 1989. Int'l Joint Conferences on Artificial Intelligence, Inc.

[Singh and Asher, 1993] M. P. Singh and N. M. Asher. A logic of intentions and beliefs. *Journal of Philosophical Logic*, 22:513–544, 1993.

[Singh, 1991] M. P. Singh. Group ability and structure. In Y. Demazeau and J.-P. Müller, editors, *Decentralized AI 2 — Proceedings of the Second European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW-90)*, pages 127–146, Saint-Quentin en Yvelines, France, 1991. Elsevier Science B.V.: Amsterdam, Netherland.

[Singh, 1998] M. P. Singh. The intentions of teams: Team structure, endodeixis, and exodeixis. In *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI)*, pages 303–307. Wiley, 1998.

[Sonenberg *et al.*, 1992] E. Sonenberg, G. Tidhar, E. Werner, D. Kinny, M. Ljungberg, and A. Rao. Planned team activity. Technical Report 26, Australian Artificial Intelligence Institute, Australia, 1992.

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

[Tidhar *et al.*, 1996] G. Tidhar, A. Rao, and E. Sonenberg. Guided team selection. In *Proceedings of Second International Conference on Multi-Agent Systems*, pages 369–376, 1996.

[Tuomela and Miller, 1988] R. Tuomela and K. Miller. We-intentions. *Philosophical Studies*, 53:367–389, 1988.

[Tuomela, 1991] R. Tuomela. We will do it: An analysis of group-intentions. *Philosophy and Phenomenological Research*, 51:249–277, 1991.

[Wooldridge and Jennings, 1999] M. Wooldridge and N. R. Jennings. The cooperative problem-solving process. *Journal of Logic and Computation*, 9(4):563–592, 1999.

[Wooldridge, 2000] M. Wooldridge. *Reasoning about Rational Agents.* The MIT Press, 2000.