The Role of Dialogue in Cooperative Problem Solving

Barbara Dunin-Kęplicz
Institute of Informatics
Warsaw University
Banacha 2, 02-097 Warsaw
Poland
e-mail: `keplicz@mimuw.edu.pl`

Rineke Verbrugge
Cognitive Science and Engineering
University of Groningen
Grote Kruisstraat 2/1, 9712 TS Groningen
The Netherlands
e-mail: `rineke@ai.rug.nl`

Abstract

Though communication is a vital ingredient of Computer Supported Cooperative Work (CSCW) as well as Cooperative Problem Solving (CPS) in multiagent systems (MAS), the in-depth analysis of different types of communication has been missing in the MAS literature. This paper presents an investigation into the role the specific types of dialogue play during consecutive stages of CPS in BDI systems, i.e., *potential recognition*, *team formation*, *plan formation*, and *team action*. Our approach is based on the typology of dialogues given by Krabbe and Walton ([16]). The relevant dialogue types are: *persuasion*, *negotiation*, *inquiry*, *deliberation*, and *information seeking*. In the present paper informal characteristics of these types are presented.

1 Introduction

In Computer Supported Cooperative Work (CSCW) one of the main issues is the study of how groups work, and how the technology enhancing group interaction can be implemented. On the other hand, in multiagent systems (MAS), especially in BDI (Belief, Desire, Intention) Systems the crucial notion is that of Cooperative Problem Solving or teamwork. In any form of CPS the vital ingredients are *cooperation*, *coordination*, and *negotiation*. Designers of MAS try to achieve a good balance between communication and reasoning, but *communication* is always necessary. In recent systems models of communication range from rather inflexible communication protocols to more sophisticated constructions based on speech acts. What is missing is a more in-depth analysis of different types of communication. The main purpose of this paper is to fill this gap. This sort of discussion may be profitable whenever communication technologies are considered, e.g. in MAS or in CSCW.

Our approach to communication is based on the typology of dialogues given by Krabbe and Walton ([16]), working in the strong tradition of argumentation theory and informal logic. In their framework the dialogue types are: *persuasion*, *negotiation*, *inquiry*, *deliberation*, *information seeking*, and *eristics*. As we consider only cooperative teams, eristics, i.e. fighting between agents, has been left out. In the paper we first give a short characterization of the above dialogue types. Next, we discuss their role during teamwork.

In our BDI framework, CPS is modelled explicitly. We base our analysis on the [17] model, containing the consecutive stages of *potential recognition*, *team formation*, *plan formation*, and *team action*. These stages are viewed as levels of abstraction and constitute together an abstract architecture. When defining the levels we abstract from particular methods and algorithms meant to realize level-associated goals, but instead formulate their final results and associate them with appropriate individual, social, and collective motivational attitudes. Now, on the basis of the four-level structure we show which dialogue types are involved in CPS on different stages. The first, informal, part of the discussion is presented in this paper. The second part will include a formal normative model of CPS related dialogue types.

The theory of collective motivational attitudes, with collective commitment as a central concept, focuses on static aspects of CPS. In [4, 5] we have provided a formalization of social and collective motivational attitudes of agents within strictly cooperative groups. Special attention was paid to the strongest attitude: *collective commitment*. We find that commitments are stronger than intentions, in the sense that it is commitments, not intentions, that trigger action execution. Our formal characterization of collective commitments is viewed from the perspective of teamwork, which is reflected in a plan towards achieving a given goal. The collective commitment hinges on pairwise social commitments between the team members to realize the actions resulting from the generated decomposition of the overall goal.

The proper treatment of collective commitments in a dynamically changing environment entails the maintenance of all individual, social and collective motivational attitudes involved throughout the whole process. In other words, a dynamic environment poses the problem that team members may fail to bring their tasks to a good end or new opportunities may appear. This leads to the so-called *reconfiguration problem* which may be treated as the core of plan execution. It has been the main subject of [6]: in terms of generic levels and their interplay an efficient and flexible *reconfiguration algorithm* has been specified. The algorithm is meant to be generic: a pattern of behavior is described in terms of abstract level-associated procedures, viewed as a sort of black box with specified input and output parameters. The next step should be a transformation of *black* boxes into *glass* ones. The investigation into the role that specific types of dialogue play at different stages of CPS, presented in this paper, is an integral part of this process.

The paper is structured in the following manner. Section 2 presents characteristics of dialogue types that appear in CPS. Section 3 includes a brief logical background. The CPS process can be divided into four stages. On the basis of this structure, in section 4 different dialogue types that are involved during consecutive stages are described. Finally, section 5 presents conclusions and further research.

2 Dialogue types in CPS

Dialogue can be defined as the process of communication among two or more agents. The agents speak in turn, for example asking questions and giving replies, and take into account, at each turn, what has occurred previously. The score of the dialogue is kept by each agent as an *attitude store*, to which propositions may be added or retracted during the dialogue in an orderly way. These propositions represent informational attitudes like individual beliefs ($\text{BEL}(a, \varphi)$) or collective beliefs within a group ($\text{C-BEL}_G(\varphi)$). In our MAS context they may also represent motivational attitudes like individual goals ($\text{GOAL}(a, \varphi)$), individual intentions ($\text{INT}(a, \varphi)$), collective intentions($\text{C-INT}_G(\varphi)$), social commitments ($\text{COMM}(a, b, \varphi)$), and collective commitments ($\text{SC-COMM}_{G,P}(\varphi)$ and $\text{WC-COMM}_{G,P}(\varphi)$).

Krabbe and Walton [16] provide a typology of dialogue types between two agents, with an emphasis on the persuasion dialogue. For each type of dialogue, they formulate *an initial situation, a primary goal*, and *a set of rules*. These constitute a *normative model*, representing the ideal way reasonable, cooperative agents participate in the type of dialogue in question, and not a record of real dialogues. Below, their typology is shortly explained and adapted to the MAS and CSCW perspective. In the course of communication, there often occurs a shift from one type of dialogue to another. A special kind of shift, called *embedding*, occurs when the second dialogue is functionally related to the first one and improves its quality. For example, a persuasion dialogue about the truth of a proposition may need an information-seeking phase.

2.1 Persuasion

A persuasion dialogue arises from a conflict of opinions. It may be that one agent believes φ while some others either believe a contrary proposition ψ_i (where $\varphi \wedge \psi_i$ is inconsistent) or just have doubt about φ. The goal of a persuasion dialogue is to resolve the conflict by verbal means, in such a way that a stable agreement results. In the multi-agent context, this would mean that the end result is a collective belief $\text{C-BEL}_G(\chi)$, where χ may be the φ of the first agent or one of the ψ_i's that the other agents started out with, or yet another conclusion that replaces the original belief. Clearly, belief revision takes place here.

Initially, all agents have attitude stores consisting of *theses* and *concessions*. Here, the theses are assertions they are prepared to defend (like φ for the first agent above), while concessions are propositions that are taken for granted for the sake of argument. Krabbe and Walton provide many rules governing the ways that the attitude stores may change during persuasion, especially the situations in which concessions and theses may be retracted. In the MAS context we need to define a new kind of persuasion not found in Krabbe and Walton, namely *persuasion with respect to motivational attitudes*. This type of dialogue arises from a conflict of intentions, where one agent intends to achieve φ, while some other agents have a conflicting intention to achieve ψ_i (where φ and ψ_i are inconsistent) or simply lack any positive motivational attitude with respect to φ. The main goal of this type of persuasion is to resolve this conflict in such a way that a stable collective intention results.

2.2 Negotiation

The initial situation of negotiation is a conflict of interests, together with a need for cooperation. The main goal is to make a deal. Thus, the selling and buying of goods and services, that is often described in the MAS-literature, is only one of the many contexts where negotiation plays a role in MAS. Negotiation and persuasion are often not distinguished adequately. One has to keep in mind that the main aim of negotiation is not to convince the others that one's view is right, as it is in persuasion, but to make a deal about a course of action. Also, in general it is not allowed to embed negotiation into persuasion, because bargaining and substituting offers for arguments does not lead to a stable agreement with respect to a collective belief. On the other hand, persuasion may be embedded in negotiation, e.g. when setting up the agenda. The rules governing negotiation include severe restrictions on retracting concessions, which are represented mostly as courses of action. For example, if an agent has conceded that it is prepared to execute some action (e.g. to sell a product for a certain price) it may not generally retract this concession.

2.3 Inquiry

The initial situation of inquiry is one where all agents are ignorant about the solution to some question or open problem. The main goal is the growth of knowledge, leading to agreement about the conclusive answer of the question. This goal is attained by a incremental process of argument which builds on established facts in order to prove conclusions beyond a reasonable doubt. Both information retrieval and reasoning are used in this process. In terms of collective informational attitudes, the end result is at least as strong as $\text{C-BEL}_G(\varphi)$. In some contexts it may even be $\text{C-KNOW}_G(\varphi)$. If one agent reaches an intermediate or final conclusion before the others, it needs to convince the others by persuasion. Thus, a persuasion dialogue may be embedded in inquiry. On the other hand, in persuasion sometimes an open problem appears which needs to be settled before the persuasion can continue; in this case, inquiry is embedded in the persuasion dialogue.

2.4 Deliberation

Deliberation as a type of dialogue is similar to inquiry, and different from both persuasion and negotiation as it starts from an open problem, rather than from a conflict of opinion. The initial situation for deliberation is a need for action performance and is concerned with the future. The main goal is to reach a decision on how to act, for the time being. The kind of reasoning that is central to deliberation, and in general to BDI systems, is *practical reasoning*: goal-directed, knowledge-based reasoning where an agent considers different means of achieving a goal. A typical example of practical inference is a means-ends analysis linking a particular goal or intention an agent is committed to with a, possibly complex, action.

2.5 Information seeking

The initial situation of information seeking occurs when one agent is ignorant about the truth of a certain proposition and seeks information from other agents in order to gain more knowledge. In terms of informational attitudes, the end result is a new individual belief $\text{BEL}(a, \varphi)$ of the information seeking

$\mathrm{BEL}(a, \varphi)$	agent a has the belief that φ
$\text{C-BEL}_G(\varphi)$	group G has the collective belief that φ
$\mathrm{GOAL}(a, \varphi)$	agent a has as a goal that φ be true
$\mathrm{GOAL}(a, \alpha)$	agent a has as a goal to do α
$\mathit{stit}(a, \varphi)$	agent a sees to it that φ holds
$\mathit{done}(a, \alpha)$	agent a has done α at the previous moment
$\mathrm{INT}(a, \varphi)$	agent a has the intention to make φ true
$\mathrm{INT}(a, \alpha)$	agent a has the intention to do α
$\text{E-INT}_G(\varphi)$	every agent in group G has the individual intention to make φ true
$\text{M-INT}_G(\varphi)$	group G has the mutual intention to make φ true
$\text{C-INT}_G(\varphi)$	group G has the collective intention to make φ true
$\mathrm{COMM}(a, b, \alpha)$	agent a commits to agent b to do α
$\text{WC-COMM}_{G,P}(\varphi)$	group G has a weak collective commitment to make φ true by plan P
$\text{SC-COMM}_{G,P}(\varphi)$	group G has a strong collective commitment to make φ true by plan P

Table 1: Formulas and their intended meaning

agent a. In contrast to the case of inquiry, the need for proof is not strong in information seeking, and it is not necessary to establish a collective belief. An example of this dialogue type occurs in expert consultation, which is useful if the questioner has no direct access to information.

3 Logical background

We propose the use of multi-modal logics to formalize agents' informational and motivational attitudes. In our framework the axioms relating motivational attitudes of agents appear with respect to *propositions*, reflecting particular states of affairs that an agent aims for, and/or with respect to *actions*. Table 1 gives basic formulas appearing in this paper, together with their intended meanings. The symbol φ denotes a proposition and α an action. The individual motivational attitudes are primitive but are governed by axiom systems and corresponding semantics, while the social and collective motivational attitudes are defined by axioms in terms of the individual ones.

3.1 Kripke models

Each Kripke model for a set of agents $\mathcal{A} = \{1, \ldots, n\}$, and a set of formulas $\mathcal{F}or$ on the basis of propositional atoms $\mathcal{P}$ consists of a set of worlds, a set of accessibility relations between worlds, and a valuation of the propositional atoms, as well as predefined truth functions for abilities, opportunities and willingness, as follows.

Definition 3.1 (Kripke model)
A Kripke model is a tuple $\mathcal{M} = (W, T, \{B_i : i \in \mathcal{A}\}, \{G_i : i \in \mathcal{A}\}, \{I_i : i \in \mathcal{A}\}, \mathit{Val}, \mathit{abl}, \mathit{op}, \mathit{will})$, such that

1. W is a set of possible worlds, or states;
2. T is a discrete tree-like partial ordering, representing branching time on W;
3. For all $i \in \mathcal{A}$, it holds that $B_i, G_i, I_i \subseteq W \times W$. They stand for the accessibility relations for each agent w.r.t. beliefs, goals, and intentions, respectively. For example, $(w_1, w_2) \in B_i$ means that w_2 is an epistemic alternative for agent i in state w_1.
4. $\mathit{Val} : \mathcal{P} \times W \to \{0, 1\}$ is the function that assigns the truth values to propositional formulas in states.
5. $\mathit{abl} : \mathcal{A} \times \mathcal{F}or \to \{0, 1\}$ is the ability function such that $\mathit{abl}(i, \varphi) = 1$ indicates that agent i is able to achieve φ.

6. $op : \mathcal{A}\times\mathcal{F}or \to (W \to \{0,1\})$ is the opportunity function such that $op(i,\varphi)(w) = 1$ indicates that in world w, agent i has the opportunity to achieve φ.
7. $will : \mathcal{A}\times\mathcal{F}or \to (W \to \{0,1\})$ is the willingness function such that $wil(i,\varphi)(w) = 1$ indicates that in world w, agent i is willing to participate in team formation with respect to overall goal φ.

The aspect of ability (cf. the *abl*-function) considers whether the agents can perform the right type of tasks. It does not depend on the situation, but may be viewed as an inherent property of the agent itself. The aspect of opportunity (cf. the *op*-function) takes into account the possibilities of task performance in the present situation, involving resources and possibly other properties. The aspect of willingness (cf. the *will*-function) takes into account a very weak motivational attitude of an agent with respect to a given overall goal φ: it does not necessarily intend φ in the current situation, but it is prepared to participate in team formation (possibly resulting in the agent adopting an individual, and even its part in a mutual, intention towards φ). Both abilities and opportunities, as well as willingness, are modeled in the above definition in a rather static way. It is possible to make more refined definitions, using a language that includes dynamic and/or temporal operators (see e.g. [14] and [2]). We have chosen not to do so here, because these concepts are not the main focus of this paper. More in general, the dynamic and temporal aspects of the logic will not be explicitly treated in this paper. At this stage, it is possible to define the "non-modal" truth conditions.

Definition 3.2 (Semantics for the non-modal operators)
Truth at worlds in the Kripke model is defined by induction as follows:

- $\mathcal{M}, v \models p \Leftrightarrow Val(p, w) = 1$;
- $\mathcal{M}, v \models \neg\varphi \Leftrightarrow \mathcal{M}, v \not\models \varphi$;
- $\mathcal{M}, v \models \varphi \wedge \psi \Leftrightarrow \mathcal{M}, v \models \varphi$ and $\mathcal{M}, v \models \psi$;
- $\mathcal{M}, v \models able(i,\varphi) \Leftrightarrow abl(i,\varphi) = 1$;
- $\mathcal{M}, v \models opp(i,\varphi) \Leftrightarrow op(i,\varphi)(v) = 1$;
- $\mathcal{M}, v \models willing(i,\varphi) \Leftrightarrow will(i,\varphi)(v) = 1$.

We view individual beliefs, goals, and intentions as primitive notions and use well-known modal logics to formalize them.

3.2 Beliefs

For the epistemic operator BEL, referring to an agent's belief, the standard modal system $\mathrm{KD45}_n$ for n agents is used [8], with the standard semantics where the relationsB_i are serial, transitive, and euclidean. One can then define modal operators for group beliefs. The formula $\text{E-BEL}_G(\varphi)$ is meant to stand for "every agent in group G believes φ". It is defined semantically as $w \models \text{E-BEL}_G(\varphi)$ iff for all $i \in G$, $w \models \mathrm{BEL}(i,\varphi)$), which corresponds to the following axiom:

E-BEL $\text{E-BEL}_G(\varphi) \leftrightarrow \bigwedge_{i\in G} \mathrm{BEL}(i,\varphi)$

A traditional way of lifting single-agent concepts to multiagent ones is through the use of *collective belief* $\text{C-BEL}_G(\varphi)$. This rather strong operator is similar to the more usual one of common knowledge. $\text{C-BEL}_G(\varphi)$ is meant to be true if everyone in G believes φ, everyone in G believes that everyone in G believes φ, etc. Define t to be *G-reachable* from s if there is a path in the Kripke model from s to t along belief-accessibility arrows B_i that are associated with members i of G. Then the following property holds (see [8]):

$$s \models \text{C-BEL}_G(\varphi) \text{ iff } t \models \varphi \text{ for all } t \text{ that are } G\text{-reachable from } s.$$

The following axiom and rule can be added to the union of $KD45_n$ and E-BEL_G, resulting in the system $KD45_n^C$, which is sound and complete with respect to Kripke models where all n accessibility relations are transitive, serial and euclidean [8]:

C-BEL $\text{C-BEL}_G(\varphi) \leftrightarrow \text{E-BEL}_G(\varphi \wedge \text{C-BEL}_G(\varphi))$

RC1 From $\varphi \to \text{E-BEL}_G(\psi \wedge \varphi)$ infer $\varphi \to \text{C-BEL}_G(\psi)$ (Induction Rule)

3.3 Individual motivational attitudes

Our framework to describe motivational attitudes and related aspects is minimal in the sense that we aim to deal with concise necessary and sufficient conditions. Additional aspects appearing on the stage in specific cases may be addressed by refining the system and adding new axioms. This subsection focuses on individual goals and intentions, and gives a short overview of our choice of axioms from [11], and the corresponding semantic conditions .

As to Kripke semantics, the binary accessibility relations G_i and $I - i$ correspond to each agent's goals (or desires) and intentions, all of which lead from a world to a world. For example, wI_iv may be interpreted as "in world w, from agent i's perspective, world v is consistent with its intentions". Evaluation of formulas at worlds is defined in the obvious manner inspired by epistemic logic. For $i = 1, \ldots, n$ we have:

$M, w \models \text{GOAL}(i, \varphi)$ iff $\forall v$ with wG_iv, $M, v \models \varphi$
$M, w \models \text{INT}(i, \varphi)$ iff $\forall v$ with wI_iv, $M, v \models \varphi$

For the motivational operators GOAL and INT the axioms and rules include the system K_n. For $i = 1, \ldots, n$ the following axioms and rules are included:

A1 All instantiations of tautologies of the propositional calculus

R1 From φ and $\varphi \to \psi$ infer ψ (Modus Ponens)

A2$_G$ $\text{GOAL}(i, \varphi) \wedge \text{GOAL}(i, \varphi \to \psi) \to \text{GOAL}(i, \psi)$ (Goal Distribution)

A2$_I$ $\text{INT}(i, \varphi) \wedge \text{INT}(i, \varphi \to \psi) \to \text{INT}(i, \psi)$ (Intention Distribution)

R2$_G$ From φ infer $\text{GOAL}(i, \varphi)$ (Goal Generalization)

R2$_I$ From φ infer $\text{INT}(i, \varphi)$ (Intention Generalization)

In a BDI system, an agent's activity starts from goals. As the agent may have many different objectives, its goals need not be consistent with each other. Then, the agent chooses a limited number of its goals to be intentions. It is not the main focus of this paper to discuss how intentions are formed from a set of goals (but see [2]). However, goals and their relation with intentions form an important part of BDI-theory, so goals are first-class citizens in our system. In any case, we assume that intentions are chosen in such a way that consistency is preserved. Thus for intentions we assume, as Rao and Georgeff do, that they should be consistent. This can be formulated as follows:

A6$_I$ $\neg\text{INT}(i, \bot)$ for $i = 1, \ldots, n$ (Intention Consistency Axiom)

Rao and Georgeff also add an analogous axiom for the consistency of goals. However, it was argued above that an agent's goals are not necessarily consistent with each other. Thus, we adopt the basic system K_n for goals. Nevertheless, in the presented approach other choices may be adopted without consequences for the rest of the definitions in this paper.

It is not hard to prove soundness and completeness of the basic axiom systems for goals and intentions with respect to suitable classes of models by a tableau method, and also give decidability results using a small model theorem.

Interdependencies between belief and individual motivational attitudes are expressed by the following axioms:

A7$_{GB}$ $\text{GOAL}(a, \varphi) \to \text{BEL}(a, \text{GOAL}(a, \varphi))$ (Positive Introspection for Goals).

A7$_{IB}$ $\text{INT}(a, \varphi) \to \text{BEL}(a, \text{INT}(a, \varphi))$ (Positive Introspection for Intentions).

A8$_{GB}$ $\neg\text{GOAL}(a, \varphi) \to \text{BEL}(a, \neg\text{GOAL}(a, \varphi))$ (Negative Introspection for Goals).

A8$_{IB}$ $\text{INT}(a, \varphi) \to \text{BEL}(a, \text{INT}(a, \varphi))$ (Negative Introspection for Intentions).

The semantic property corresponding to $\mathbf{A7}_{IB}$ is $\forall x, y, z((xB_a y \wedge yI_a z) \to xI_a z)$, analogously for $\mathbf{A7}_{GB}$. The semantic property that corresponds to $\mathbf{A8}_{IB}$ is $\forall x, y, z((xI_a y \wedge xB_a z) \to zI_a y)$, analogously for $\mathbf{A8}_{GB}$. We also assume that every intention corresponds to a goal:

$\mathbf{A9}_{ID}$ $\mathrm{INT}(a, \varphi) \to \mathrm{GOAL}(a, \varphi)$ (Intention implies goal)

The corresponding semantic property is that $G_a \subseteq I_a$.

To model teamwork, individual attitudes naturally do not suffice. In other work we extensively discussed the pair-wise notion of social commitments, as well as collective notions like collective belief, collective intention, and collective commitment [5, 7]. This paper will present only brief reminders of the definitions.

4 The four levels of CPS

We base our analysis of CPS in MAS and CSCW on the four-stage model of [17], containing the consecutive stages of *potential recognition*, *team formation*, *plan formation* and *team action*. These stages will be viewed as levels of abstraction and constitute together an abstract architecture. Taking into account the unpredictable environment, all four stages have a dynamic character and require methods reflecting this. When defining the levels we abstract from particular methods and algorithms meant to realize level-associated goals, but instead formulate their final results and associate them with appropriate individual, social, and collective motivational attitudes (see [6] for extended discussion). Now, we will concentrate on different types of dialogue involved in the realization of level-oriented goals.

Level 1: the potential recognition level

The starting point is the overall goal that should be achieved by the Cooperative Problem Solving process. In this paper, as in [17], we restrict ourselves to the situation with one fixed overall goal and one agent that takes the initiative to achieve that goal. Analogous to [17], we consider CPS to begin when some agent in a multi-agent environment recognizes the potential for cooperative action to reach the overall goal φ of the system.

The first task of the *initiator* is to form a partial (abstract) plan for the achievement of the overall goal. On the basis of the (type of) subgoals that it recognizes it will determine which agents might be most suited to form the team. In order to determine this match the initiator tries to find out the properties of the agents, being interested in three aspects, namely their *abilities*, *opportunities*, and their *willingness* to participate in team formation.

The aspect of ability considers whether the agents can perform the right type of tasks. Ability does not depend on the situation, but may be viewed as some kind of inherent property of the agent itself. The aspect of opportunities takes into account the possibilities of task performance in the present situation. It involves resources and possibly other such properties. The aspect of willingness considers their mental attitude towards participating in the team. Very capable agents that do not want to do the job are of no use. The components of agent's suitability are represented as follows:

1. the individual ability of agent b to achieve a goal ψ is denoted by $able(b, \psi)$.
2. the resources available to agent b are reflected by the opportunity that agent b has to achieve ψ, denoted by $opp(b, \psi)$.
3. the willingness of agent b to participate in team formation is denoted by $willing(b, \varphi)$.

A simple static semantics for these notions has been given in subsection 3.1.

Towards a potential of cooperation

In the previous subsection we have described what type of information the initiating agent tries to gather in order to start team formation. In this subsection we will describe how the information is collected, and how it leads to the formal outcome.

The output at this stage is the "potential for cooperation" POTCOOP(a, φ) that the initiator a sees with respect to φ. This means that φ is a goal of a (GOAL(a, φ)), and there is a group G, which is a subset of the given set of relevant agents T, such that a believes that G can collectively achieve φ (C-CAN$_G(\varphi)$) and are willing to participate in team formation; and either a cannot or doesn't desire to achieve φ in isolation. This is expressed by the following definitions (see [6] for more discussion).

$$\begin{aligned}
&\mathrm{POTCOOP}(a,\varphi) \leftrightarrow \mathrm{GOAL}(a,\varphi) \wedge \\
&\quad \exists G \subseteq T \mathrm{BEL}(a, \text{C-CAN}_G(\varphi) \wedge \forall i \in G \mathit{willing}(i,\varphi)) \wedge \\
&\quad (\neg \mathrm{CAN}(a,\varphi) \vee \neg \mathrm{GOAL}(a, \mathit{stit}(\varphi)))
\end{aligned}$$

$$\mathrm{CAN}(a,\varphi) \leftrightarrow \mathit{able}(a,\varphi) \wedge \mathit{opp}(a,\varphi).$$

POTCOOP(a, φ) is derived from the information collected from the individual agents. To derive C-CAN$_G(\varphi)$ the initiator compares the information obtained about the other agents against a partial abstract plan for the overall goal φ. For this purpose the overall goal φ is split into a number of subgoals $\varphi_1, \ldots, \varphi_n$. These subgoals can be viewed as instrumental to the overall goal. Together they *realize* φ and are compared with the individual abilities and opportunities that the agents are believed to have. Formally:

$$\begin{aligned}
&\text{C-CAN}_G(\varphi) \leftrightarrow \\
&\quad \exists \varphi_1, \ldots, \exists \varphi_n (\mathit{realize}(< \varphi_1, \ldots, \varphi_n >, \varphi) \wedge \\
&\quad \forall i \leq n \exists j \in G(\mathit{able}(j, \varphi_i) \wedge \mathit{opp}(j, \varphi_i)))
\end{aligned}$$

Here, $(\mathit{realize}(< \varphi_1, \ldots, \varphi_n >, \varphi)$ intuitively means: "if $\varphi_1, \ldots, \varphi_n$ are achieved, then φ holds". Technically, one may need some extra-logical, context-dependent, reasoning to show the implication.

As output of a successful outcome of this stage, track is kept of the sequence $(G_1, \ldots, G_n)$ of all relevant groups $G \subseteq T$ for which a sees potential for cooperation with respect to the main goal. It will be of use when the need for reconfiguration appears.

Information seeking during potential recognition

Ultimately, the initiator has to form beliefs about the abilities, opportunities, and willingness of the individual agents in order to derive POTCOOP(a, φ). This can be done by a asking every agent about its properties and the agent responding with the requested information. Formally this can be expressed as follows, where ψ may stand for any of $\mathit{willing}(i, \varphi)$, $\mathit{able}(i, \varphi_i)$, $\mathit{opp}(i, \varphi_i)$, etc.

The formulas below are based on the formal theory on *speech acts* developed in [15] and [3]. The illocutionary acts $ASS_{i,j}$ and $REQ_{i,j}$ stand for an assertion respectively a request made by agent i to agent j. For the full characterizations, see [15]. Here it suffices to say that asserting a proposition entails that the speaker believes it; requesting another agent to perform an action α entails that the speaker wants the hearer to do α.

$$\begin{aligned}
&[REQ_{a,i}(\textbf{ if } \psi \textbf{ then } ASS_{i,a}(\psi) \textbf{ else } ASS_{i,a}(\neg\psi))] \\
&[ASS_{i,a}(\psi)]\mathrm{BEL}(a,\psi)
\end{aligned}$$

This is based on a dynamic logic formula of the form $[\alpha_1][\alpha_2]\psi$, meaning that if α_1 is performed then always a situation arises such that if α_2 is performed then in the resulting state ψ will always hold.

In the above case $alpha_1$ is the complex action

$$REQ_{a,i}(\textbf{ if } \psi \textbf{ then } ASS_{i,a}(\psi) \textbf{ else } ASS_{i,a}(\neg\psi)),$$

where $REQ_{a,i}(\alpha)$ stands for agent a requesting agent i to perform the action α.

In an analogous way, an information seeking dialogue about all ingredients of POTCOOP(a, φ) takes place. In principle, the schema of all necessary questions may be rather complex. Finally, the formal result of the potential recognition stage is that agent a has the belief that it is possible to form a team to achieve the overall goal:

$$\mathrm{BEL}(a, \text{C-CAN}_G(\varphi) \wedge \forall i \in G \mathit{willing}(i,\varphi))$$

with respect to the relevant groups $G \subseteq T$, or that it is not. It would be useful for the other agents to know this fact as well, in order to revise their motivational and informational attitudes. This information might be divulged to all agents by using a broadcast message of a stating the formal outcome of this stage.

For more in-depth analysis of the speech acts that occur in these information seeking dialogues, see [1, 2].

Level 2: the team formation level

Suppose that agent a sees the potential for cooperation in order to achieve φ. Somewhat different from [17], we find that during the team formation stage agent a attempts to bring it about in some group G that the group has the *collective intention* to make the overall goal φ true. The input of this stage is agent a, a formula φ and sequence of potential groups $(G_1, \ldots, G_n)$ as output by stage 1. The output of a successful outcome of this stage is one group G from $(G_1, \ldots, G_n)$ together with a collective intention $\text{C-INT}_G(\varphi)$ among G to achieve φ.

Collective intentions

As we focus on strictly cooperative groups, the definition of collective intention is rather strong. In such groups, a necessary condition for a collective intention is that all members of the group have the associated individual intention $(\text{INT}(i, \varphi))$. Moreover, to exclude the case of competition, all agents should *intend* all members to have the associated individual intention, as well as the intention that all members have the individual intention, and so on; we call such a mutual intention $\text{M-INT}_G(\varphi)$. Furthermore, all members of the group are aware of this mutual intention, that is, they have a collective belief about this $(\text{C-BEL}_G(\text{M-INT}_G(\varphi)))$.

In order to formalize the above two conditions, $\text{E-INT}_G(\varphi)$ (standing for "everyone intends") is defined by the following axiom, corresponding to the semantic condition that $w \models \text{E-INT}_G(\varphi)$ iff for all $i \in G$, $w \models \text{INT}(i, \varphi)$:

(E-INT) $\text{E-INT}_G(\varphi) \leftrightarrow \bigwedge_{i \in G} \text{INT}(i, \varphi)$.

The mutual intention $\text{M-INT}_G(\varphi)$ is meant to be true if everyone in G intends φ, everyone in G intends that everyone in G intends φ, etc. As we do not have infinite formulas to express this, let $\text{E-INT}^1_G(\varphi)$ be an abbreviation for $\text{E-INT}_G(\varphi)$, and let $\text{E-INT}^{k+1}_G(\varphi)$ for $k \geq 1$ be an abbreviation of $\text{E-INT}_G(\text{E-INT}^k_G(\varphi))$. Thus we have $w \models \text{M-INT}_G(\varphi)$ iff $w \models \text{E-INT}^k_G(\varphi)$ for all $k \geq 1$. Define world t to be *G-reachable* from world s if there is a path from s to t along accessibility arrows I_i that are associated with members i of G. Then the following property holds (see [8] for an analogous property for common knowledge):

$$s \models \text{M-INT}_G(\varphi) \text{ iff}$$
$$t \models \varphi \text{ for all } t \text{ that are } G\text{-reachable from } s.$$

Using this property, it can be shown that the following axiom and rule can be soundly added to the union of KD_n and (E-INT):

(M-INT) $\text{M-INT}_G(\varphi) \leftrightarrow \text{E-INT}_G(\varphi \wedge \text{M-INT}_G(\varphi))$

(R-INT) From $\varphi \to \text{E-INT}_G(\psi \wedge \varphi)$ infer $\varphi \to \text{M-INT}_G(\psi)$ (Induction Rule)

The resulting system is called $KD_n^{\text{M-INT}_G}$, and it is sound and complete with respect to Kripke models where all n accessibility relations are serial (by a proof analogous to the one for common knowledge in [8]). Finally, the collective intention is defined formally:

(C-INT) $\text{C-INT}_G(\varphi) \leftrightarrow \text{M-INT}_G(\varphi) \wedge \text{C-BEL}_G(\text{M-INT}_G(\varphi))$

Note that this definition is different from the one given in [4, 5]; The new definition is extensively discussed and compared with others in [7]. Let us remark that, even though $\text{C-INT}_G(\varphi)$ seems to be an infinite concept, collective intentions may be established in practice in a finite number of steps.

Formula (M-INT) makes evident that a crucial step for the initiator is to persuade all members of a potential team to take the overall goal as individual intention. To establish the second part of the conjunction, the initiator also persuades each member to take on the intention that there be a mutual intention among the potential team to contribute to the overall goal, in order to foster cooperation from the start.

Persuasion during team formation

The main type of dialogue at this stage is *persuasion with respect to motivational attitudes.* This arises from a conflict of intentions, where $\mathrm{INT}(a, \varphi)$ while some other agents in the potential team have a conflicting intention $\mathrm{INT}(i, \psi_i)$ or simply lack any positive motivational attitude with respect to φ. In the case of team formation the persuasion is one-sided so that in the end all agents adopt a's intention to achieve φ. In contrast to persuasion with respect to informational attitudes, bargaining may be appropriate within a persuasion dialogue with respect to motivational attitudes. For example, during team formation potential team members may be reasonably persuaded using an embedded negotiation about return favors from agent a. In [1, 2], a detailed analysis of the persuasion dialogue during team formation is given, including formal rules in terms of speech acts.

When all the individual motivational attitudes are established within the team, the initiator broadcasts the fact $\text{E-INT}_G(\varphi)) \wedge \text{E-INT}_G(\text{E-INT}_G(\varphi))$, by which the necessary collective beliefs are established and the collective intention is in place.

Goal of the persuasion dialogue

The goal of a persuasion dialogue is to establish a collective intention within a group. The above formula **M3** makes evident that a crucial step for the initiator is to persuade all members of a potential team to take the overall goal as an individual intention. To establish the higher levels of the mutual intention, the initiator also persuades each member to take on the intention that all members of the potential team have the mutual intention, in order to strengthen cooperation from the start. It suffices if the initiator persuades all members of a potential team to take on an individual intention towards φ and the intention that there be a mutual intention among that team. Formally, for all $i \in G$, agent a seeks to establish $\mathrm{INT}(i, \varphi \wedge \text{M-INT}_G(\varphi))$. For this implies by axiom **M1** that $\text{E-INT}_G(\varphi \wedge \text{M-INT}_G(\varphi))$, which in turn implies by axiom **M2** that $\text{M-INT}_G(\varphi)$. When all the individual motivational attitudes are established within the team, the initiator broadcasts the fact $\text{M-INT}_G(\varphi)$, by which the necessary collective belief $\text{C-BEL}_G(\text{M-INT}_G(\varphi)))$ is established and the collective intention is in place.

This will be achieved during a persuasion dialogue, which according to [16] consists of three main stages: information exchange, rigorous persuasion and completion. In our case the information exchange already started in the potential recognition stage. Let us remind the reader that we extended the concept of persuasion to include also intentions, and not only beliefs. The final result of the team formation stage is reached when for one potential team all the persuasion dialogues have been concluded successfully.

Information exchange during persuasion

During the information exchange the agents make clear their initial stand with respect to the overall goal and to form part of a certain team to achieve it. These issues are expressed partly in the form of intentions and beliefs. Other beliefs supporting or related to the above issues might also be exchanged already. Only when a conflict arises about these issues a persuasion dialogue has to take place. In each persuasion there are two parties or roles; the proponent (P) and the opponent (O). In our case the proponent is the initiator and the opponent the other agent.

The stands the other agents take about the above issues are seen as its initial *concessions.* Concessions are beliefs and intentions that an agent takes on for the sake of argument, but need not be prepared to defend. The agents will also have private attitudes that may only become apparent later on during the dialogue. The stand of the initiator is seen as the initial thesis that it is prepared to defend during the dialogue. The initial conflict description consists of a set of O's initial concessions and P's initial thesis.

Rigorous persuasion

During the rigorous persuasion stage the agents exchange arguments to challenge or support a thesis. The following rules can be used to govern these moves adapted from [16].

1. Starting with O the two parties move alternately according to the rules of the game.
2. Each move consists of either a challenge, a question, a statement, a challenge or question accompanied by a statement, or a final remark.
3. The game is highly asymmetrical. All P's statements are assertions, and called *theses*, all O's statements are called *concessions*. P is doing the questioning and O all the challenging.
4. The initial move by O challenges P's initial thesis. It is P's goal to make O concede the thesis. P can do this by questioning O and thus bridge the gap between the initial concessions of O and the thesis, or by making an assertion to clinch the argument if acceptable.
5. Each move for O is to pertain to P's preceding move. If this move was a question, then O has to answer it. If it was an assertion, then O has to challenge it.
6. Each party may give up, using the final remark $ASS_{a,i}(quit)$ for the initiator, or $ASS_{i,a}(\mathrm{INT}(i, \varphi \wedge \text{M-INT}_G(\varphi)))$.

   If O's concessions imply P's thesis, then P can end the dialogue by the final remark: $ASS_{a,i}(won)$. In our system we assume that we have the following rule: $[ASS_{a,i}(won)]OBL(ASS_{i,a}(\mathrm{INT}(i, \varphi) \wedge \text{M-INT}_G(\varphi)))$, which means that agent i is obliged to state that it has been persuaded and accepts its role in the team. This does not mean that i will actually make this assertion! Just that there is an obligation (according to the rules of the persuasion "game"). The modal operator OBL is taken from deontic logic [3].
7. All challenges have to follow logical rules. For example, a thesis $A \wedge B$ can be challenged by challenging one of the two conjuncts. For a complete set of rules for the propositional part of the logic we refer to [16].

In the completion stage the outcome is made explicit, such that the agents either have a collective belief and/or intention or they know that they differ in opinion.

Speech acts during persuasion

In contrast to other approaches, for example [16], we need to monitor agent's informational and motivational attitudes during persuasion, as well as during other stages of CPS. In the course of dialogue we are concerned with assertions and challenges with respect to beliefs, and concessions and requests with respect to both informational and motivational attitudes. A detailed discussion and formal implementation of these speech acts is presented in [1, 2].

Level 3: the plan generation/formation level

The input of this stage is a team G together with its collective intention to achieve a goal φ. The final successful outcome of this stage is the weak (or strong) collective commitment of the group G based on the social plan P, namely $\text{WC-COMM}_{G,P}(\varphi)$ (resp. $\text{SC-COMM}_{G,P}(\varphi)$). Before defining it we need to introduce a notion of *social* i.e bilateral commitment between two agents.

A social commitment between two agents is not as strong as a collective commitment among them, but stronger than an individual intention of one agent. If an agent commits to a second agent to do something, then the first agent should have the *intention* to do that. Moreover, the first agents commits to the second one only if the second one is *interested* in the first one fulfilling its intention. Finally, agents need to be aware about the situation. Such awareness, expressed in terms of collective belief, is generally achieved by communication:

$$\mathrm{COMM}(a, b, \alpha) \leftrightarrow \mathrm{INT}(a, \alpha) \wedge \mathrm{GOAL}(b, done(a, \alpha)) \wedge$$
$$\text{C-BEL}_{a,b}(\mathrm{INT}(a, \alpha) \wedge \mathrm{GOAL}(b, done(a, \alpha)))$$

Collective commitments

A collective commitment among a group G based on a social plan P can only be established or maintained if the group has the associated collective intention C-INT$_G(\varphi)$. Furthermore, the plan should result from the main goal by task division, means-end-analysis, and action allocation, as reflected in *constitute*(φ, P). In addition, for every one of the actions α_i that make up plan P, there should be one agent a in the group who is socially committed to at least one (mostly other) agent b in the group to do the action, reflected in social commitments between pairs of agents (COMM(a, b, α_i)). Moreover, there should be a collective belief in the whole group that the plan will be entirely realised, i.e. that all actions have been adopted by committed members of the group. The defining axiom below reflects all these characteristics.

$$\begin{aligned}\text{WC-COMM}_{G,P}(\varphi) \leftrightarrow\ & \text{C-INT}_G(\varphi) \wedge \mathit{constitute}(\varphi, P)\\ & \wedge \bigwedge_{\alpha_i \in P} \bigvee_{a,b \in G} \text{COMM}(a, b, \alpha_i) \wedge\\ & \text{C-BEL}_G(\bigwedge_{\alpha_i \in P} \bigvee_{a,b \in G} \text{COMM}(a, b, \alpha_i))\end{aligned}$$

The above definition can be viewed as a *weak* form of collective commitment: no collective decision making is assumed. Otherwise, if the group is planning collectively, so that in the end the plan is completely known to all members, one conjunct C-BEL$_G$(*constitute*(φ, P)) reflecting this is added to get the *strong* definition of collective commitment:

$$\begin{aligned}\text{SC-COMM}_{G,P}(\varphi) \leftrightarrow\ & \text{C-INT}_G(\varphi) \wedge \mathit{constitute}(P, \varphi)\\ & \wedge \text{C-BEL}_G(\mathit{constitute}(\varphi, P)) \wedge\\ & \bigwedge_{\alpha_i \in P} \bigvee_{a,b \in G} \text{COMM}(a, b, \alpha_i) \wedge\\ & \text{C-BEL}_G(\bigwedge_{\alpha_i \in P} \bigvee_{a,b \in G} \text{COMM}(a, b, \alpha_i))\end{aligned}$$

These two versions of the definition reflect different social structures in which the group is involved. In the weak case, not all members are aware of the social plan as a whole, though they are aware about their share in it. Note that agents' limited orientation in the task distribution may be done on purpose. It is always assumed, however, that the overall goal is known to everybody.

In the AI and MAS literature many methods of planning have been proposed. From our perspective, the most interesting one is *planning from first principles*. In the MAS context, we may assume that this method includes *task division* and *task allocation*. We will discuss this method now.

During a successful run of the plan generation stage, first, an adequate task division of φ into a sequence of (possibly complex) subtasks $\varphi_1, \dots, \varphi_n$ is constructed. These subgoals are viewed as instrumental to the overall goal and are compared with the individual capabilities and opportunities that the agents are believed to have.

This step is followed by *task allocation*, in which actions are associated to the subtasks constructed in task division, and these are given to team members. This results in pairs $< \alpha_i, b >$ of a possibly complex action α_i that realizes task φ_i and an agent b. For an appropriate allocation of actions to agents, the agents' abilities and resources are taken into account. To make a social plan complete, the temporal structure among the pairs $< \alpha_i, b >$ should be established. This is reflected in the recursive definition of a *social plan expression*, adopted from [12] and inspired by dynamic logic. The result of the whole planning process is a plan P, represented as a social plan expression. Note that there is a close interplay between task division and task allocation.

Deliberation during plan generation

When using planning from first principles, *deliberation* is typically the main kind of dialogue during plan generation. The desired outcome is a collective commitment based on a social plan. It turns out that, in order to achieve this goal, different kinds of dialogue are phases within this complex stage.

The main part of *task division* is discussion of proposals. Krabbe and Walton define this as a subtype of the persuasion dialogue. Thus, *persuasion* is embedded in deliberation. The outcome of this phase should be a sequence of subtasks $\varphi_1, \ldots, \varphi_n$ and a collective belief that execution of these subtasks leads to realization of the main goal, reflected in

$$\exists\varphi_1, \ldots, \exists\varphi_n \text{C-BEL}_G(\mathit{realize}(<\varphi_1, \ldots, \varphi_n>, \varphi)).$$

Inquiry and *information seeking* are the types of dialogue usually needed during task division.

During *task allocation*, a social plan is devised to realize the sequence of subtasks $\varphi_1, \ldots, \varphi_n$, and finally to constitute the overall goal φ (as reflected in $\mathit{constitute}(\varphi, P)$). The first part of task allocation should be a sort of means-end analysis. This may be seen as *inquiry*, resulting in a collective belief that the actions α_i are means that realize the subtasks φ_i, formally:

$$\text{C-BEL}_G(\bigwedge_{i=1}^{n} \mathit{means\text{-}for}(\alpha_i, \varphi_i)).$$

A subphase of this occurs when team members are asked about their abilities, opportunities, etc. In this case *information seeking* is embedded in inquiry.

Even though agents have established a collective intention $\text{C-INT}_G(\varphi)$, being self-interested they usually still have a conflict of interests during task allocation. Thus, the agents may *negotiate* to devise a social plan reflecting the agents' individual interests in a way accepted by all of them. When this takes place, the end product is a number of social commitments $\text{COMM}(a, b, \alpha_i)$ between pairs of agents. In effect, all agents from the group socially commit to carry out their respective actions, and to communicate about these social commitments in order to establish pairwise mutual beliefs about them, as reflected by the defining axiom for social commitments. This information exchange concludes the collective part of plan generation, by which the collective commitment $\text{SC-COMM}_{G,P}(\varphi)$ (or in some situations $\text{WC-COMM}_{G,P}(\varphi)$ is established.

To sum up, a number of functional embeddings is appropriate in the deliberation dialogue. First, during task division, a persuasion dialogue is embedded into it, which may have inquiry and information seeking as subphases in its turn.

Then, during task allocation, first an inquiry (possibly with its own subphases of information seeking) is embedded into deliberation. Then, if necessary, negotiation (possibly with its own subphases of inquiry and information seking) is embedded into the deliberation dialogue. Finally, in order to establish the pairwise social commitments between team members, information seeking may be embedded into deliberation.

Another planning method adopted in MAS is e.g. using a *plan library* of more or less instantiated partial plans. Input and output of this phase is obviously the same as before. However, the dialogue types involved are less complex here, depending on the degree of plan instantiation. Because we deal with a variety of possibilities here, we will not discuss it in detail.

Level 4: the plan execution / team action level

Even though the three stages of potential recognition, team formation and plan formation have been extensively discussed in the MAS and AI literature, the important phase of collective team action has received relatively little attention. The requirements of a constantly changing environment lead to the *reconfiguration problem*: when maintaining a collective intention during plan execution, it is crucial that agents replan properly and efficiently when some members do not fulfill their delegated subtasks or are presented with new opportunities. The solution of this problem leads to the *reconfiguration algorithm* formulated in terms of the abstract levels and their (complex) interplay. See [6] for a presentation of the algorithm.

The plan execution level includes *execution of actions* realizing the agents' own subtasks from the collective commitment SC-COMM$_{G,P}(\varphi)$ or WC-COMM$_{G,P}(\varphi)$. Because many different situations may occur during execution of agent-specific actions, this usually leads to the *reconfiguration procedure.* In terms of motivational attitudes, plan realization amounts to the maintenance of *social commitments* and associated *individual intentions.* During this process communication, including all types of dialogue, cooperation, and coordination take place.

The successful outcome of the stage of plan execution is that all subtasks making up the social plan P have been carried out by the agents who were socially committed to do them, reflected in the team's collective commitment and that by the success of their actions the goal φ of the collective commitment has been achieved. The successful realization of team action finishes the evolution of the team and its motivational attitudes. Before this expected situation takes place, all aspects of evolution are treated in the reconfiguration process.

Dialogues during plan execution

Depending on the domain and the collective commitment in question, all types of dialogue may occur during *execution of actions* by the team members aiming to realize their individual subtasks. For example, in the case of scientific collaboration (see [6] for an extensive discussion of an example of this), the whole action execution stage may be viewed as *inquiry.* When team members ask the others for intermediate and final results of their individual research, *information-seeking* is embedded into inquiry. *Persuasion* also occurs as a phase of the inquiry when one agent has found a result that the others do not know yet. Thus, by information-seeking and persuasion, the agents' distributed beliefs are transformed into collective beliefs.

In the *reconfiguration procedure* an assignment of new social and individual motivational attitudes is made to create a new collective commitment. This may be viewed as a *revision* of motivational, and therefore informational, attitudes. From a pragmatic point of view, the proper treatment of revision is possible, because it is assumed that agents are obliged to communicate about changes. On the other hand, the presence of social commitments and common beliefs about them *only* between partners for whom they are relevant ensures that motivational attitudes revision is as efficient as possible.

During reconfiguration, if an obstacle appears, the problems are solved by moving up in the hierarchy of levels, but as little as possible. Thus, one moves to the nearest point up where a different choice is possible. If such a point does not exist anymore, the reconfiguration algorithm fails. In technical terms, depth-first search is used. This means that exactly those dialogue types play a role that correspond to the stages up in the hierarchy which need to be revisited in order to make a new collective commitment.

Let us trace the evolution of the system during a realization of the reconfiguration algorithm. Let us focus on the points where failure of the four main stages of the process takes place.

The failure of the potential recognition level meaning that agent a does not see any potential for cooperation with respect to the goal φ, leads to the total failure of the system and ends up the evolution of the system.

The failure of the team formation level, meaning that the collective intention C-INT$_G(\varphi)$ cannot be established among any of the teams, requires a return to the potential recognition stage to construct a new sequence of potential teams. As we said before, mostly information seeking is involved here.

The failure of the task division level requires a return to the team formation level in order to establish a collective intention in the chosen new team may be viewed as the evolution of the team together with the evolution of motivational attitudes on the level of collective intention and respective individual attitudes. In this case persuasion is the main type of dialogue involved.

The failure of the task allocation level requires a return to the task division level in order to create a new sequence of tasks. This may be viewed as the evolution of the first part of a social plan P to realize the goal φ. Here, deliberation with all other embedded dialogue types takes place.

When, finally, a collective commitment WC-COMM$_{G,P}(\varphi)$ or SC-COMM$_{G,P}(\varphi)$ is established, after the failure of some actions together constituting the social plan P, the reasons of failure have to be recognized (see [6] for details). Depending on the outcome of this analysis either the system fails, or returns to the task divison or the task allocation level. Then again, different types of dialogue take place. It is an important subject of further research to give formal rules about the relation between the dialogues taking place at each stage during subsequent returns to that stage.

5 Conclusions

In our previous work [6] it was shown how CPS can be divided into four stages, each of them resulting in specific motivational attitudes that are formally described. In the present paper the stages are related to specific types of dialogue by which these motivational attitudes can be formed. It is also shown for each stage which dialogue types may be appropriately embedded into the other. Our approach is based on the typology of dialogues given by Krabbe and Walton ([16]), which has also inspired [13]. In that paper, Reed gives a formal definition of dialogue frames and functional embeddings, together with some simple examples of dialogues formulated in his formal language. However, he does not explain how these dialogues actually play a role in cooperative problem solving, nor does he associate them to the construction of individual, social, and collective informational and motivational attitudes, as we do.

In [1, 2], the analysis of the present paper is taken a step further to the level of speech acts. It is shown for the first two stages of potential recognition and team formation which rules govern the dialogues and how the moves within these dialogues can be expressed by formally represented speech acts. The complete formal description of the dialogues has thus been made possible. In this way, it will be possible to prove that in given circumstances the dialogue results in a certain outcome. This is of prime importance if one wants to construct a MAS for automated cooperative problem solving.

Related work can be found in [10], who also present an agent architecture and a representation for agent communication. In the discussion they note that their own "negotiation" in fact covers a number of Walton and Krabbe's different dialogue types. We find the more fine-grained typology to be very useful when designing agents for teamwork: one can use specific sets of rules governing each type of dialogue as well as the possible embeddings between the different types. Thus desired kinds of communication are allowed and harmful ones prevented, without completely fixing any protocol. Also [10] use multi-context logic whereas we stick to (multi-)modal logic.

The emphasis on pre-planning (here in the stage of potential recognition) and establishing appropriate collective attitudes for teamwork is shared with [9]. Nevertheless, the intentional component in their definition of collective plan is much weaker than our collective intention: Grosz and Kraus' agents involved in a collective plan have individual intentions towards the overall goal and a collective belief about these intentions; intentions with respect to the other agents play a part only at the level of individual sub-actions of the collective plan. We stress, however, that team members' intentions about their colleagues' motivation to achieve the overall goal play an important role in keeping the team on track even if their plan has to be changed radically due to a changing environment (see also [6]).

The first issue for further research is to complete the formal normative model of CPS related dialogue types. Thus, one needs to construct a complete set of formal rules for all the types of dialogue and indicate how these are implemented through formal speech acts. In particular, formal rules should be formulated for negotiation, deliberation, and inquiry. This would make it possible to extend the analysis initiated in [1, 2] to the last two stages of cooperative problem solving, namely plan formation and execution. As part of the latter, a formal analysis of the dialogues taking place during the reconfiguration procedure is of prime importance.

Acknowledgements

We would like to thank E. Krabbe and F. Dignum for interesting discussions about the subject of this paper. This work is supported by the Polish KBN Grant 7T11C 006 20.

References

[1] F. Dignum, B. Dunin-Kęplicz, and R. Verbrugge. Agent theory for team formation by dialogue. In: C. Castelfranchi en Y. Lesperance (eds.) *Proceedings of Agent Theories, Architectures, and Languages (ATAL 2000)*, Boston, 2000, pp. 141-156. (To appear in the Springer LNAI Series, 2001.)

[2] F. Dignum, B. Dunin-Kęplicz, and R. Verbrugge. Creating collective intention through dialogue. In J. Cunningham en D. Gabbay (eds.), *Proceedings International Conference on Formal and Applied Practical Reasoning (FAPR 2000)*,), London, Imperial College, 2000, pp. 145-158. (To appear in the Logic Journal of the IGPL, 2001.)

[3] F. Dignum and H. Weigand. Communication and deontic logic. In: R. Wieringa and R. Feenstra (eds.), *Information Systems, Correctness and Reusability*, Singapore, World Scientific, 1995, pp. 242–260.

[4] B. Dunin-Kęplicz and R. Verbrugge. Collective commitments. In M. Tokoro, editor, *Proceedings Second International Conference on Multi-Agent Systems*, pages 56–63, Menlo Park (CA), AAAI-Press, 1996.

[5] B. Dunin-Kęplicz and R. Verbrugge. Collective motivational attitudes in cooperative problem solving. In V. Gorodetsky, editor, *Proceedings of the First International Workshop of Eastern and Central Europe on Multi-agent Systems (CEEMAS'99)*, pages 22–41, St. Petersburg, 1999.

[6] B. Dunin-Kęplicz and R. Verbrugge. A reconfiguration algorithm for distributed problem solving. *Electronic Modeling*, 22(2): 68–86, 2000.

[7] B. Dunin-Kęplicz and R. Verbrugge. Collective intention constitutes a group. In Y. Demazeau and F. J. Garijo (eds.), *Proceedings of the Tenth European Workshop on Multi-Agent Systems (MAAMAW01)*, Annecy, 2001.

[8] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about Knowledge.* Cambridge MA, MIT Press, 1995.

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

[10] S. Parsons, M. Wooldridge, and N. Jennings. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8:261–292, 1998.

[11] A. Rao and M. Georgeff. Modeling rational agents within a BDI-architecture. In R. Fikes and E. Sandewall, editors, *Proceedings of the Second Conference on Knowledge Representation and Reasoning*, pages 473–484. Morgan Kaufman, 1991.

[12] A. Rao, M. Georgeff, and E. Sonenberg. Social plans: A preliminary report. In E. Werner and Y. Demazeau, editors, *Decentralized A.I.-3*, pages 57–76, Amsterdam, Elsevier, 1992.

[13] C. Reed. Dialogue frames in agent communication. In Y. Demazeau, editor. *Proceedings Third International Conference on Multi-Agent Systems*, Los Alamitos (CA), IEEE Computer Society, 1998.. pages 246–253.

[14] B. van Linder, W. van der Hoek, and J.-J. C. Meyer. Formalising abilities and opportunities of agents. *Fundamenta Informaticae*, 34:53–101, 1998.

[15] J. R. Searle. *Speech Acts: An Essay in the Philosophy of Language.* Cambridge, Cambridge University Press, 1969.

[16] D. Walton and E. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning.* Albany (NY), State University of New York Press, 1995.

[17] M. Wooldridge and N. Jennings. Cooperative problem solving. *Journal of Logic and Computation*, 9:563–592, 1999.