Dialogue in Team Formation

Frank Dignum[1]
Barbara Dunin-Kęplicz[2]
Rineke Verbrugge[3]

[1]Faculty of Mathematics and Computing Science
Technical University Eindhoven
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
e-mail: dignum@win.tue.nl
Tel: +31-40-2473705, Fax: +31-40-2463992
[2]Institute of Informatics
Warsaw University
Banacha 2, 02-097 Warsaw, Poland
E-mail: keplicz@mimuw.edu.pl
Tel: +48 22 6584594, Fax: +48 22 6583164
[3]Cognitive Science and Engineering
University of Groningen
Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands
E-mail: rineke@tcw2.ppsw.rug.nl
Tel: +31-50-3636334, Fax: +31-50-3636784

Abstract. The process of cooperative problem solving can be divided into four stages. First, finding potential team members, then forming a team followed by constructing a plan for that team. Finally, the plan is executed by the team. Traditionally, very simple protocols like the Contract Net protocol are used for performing the first two stages of the process. And often the team is already taken for granted. In an open environment (like in e.g. electronic commerce) however, there can be discussion among the agents in order to form a team that can achieve the collective goal of solving the problem. For these cases fixed protocols like contract net do not suffice. In this paper we present an alternative solution, using structured dialogues that can be shown to lead to the required team formation. The dialogues are described formally (using some modal logics), thus making it possible to actually prove that a certain dialogue has a specific outcome.

1 Introduction

Contract Net [12] is often proposed as a simple but effective and efficient way to distribute tasks over a number of agents in order to achieve a common goal. The reason of the success of this approach is the fact that it is completely fixed and thus easy to implement. In our opinion, however,

[0] This research was supported by ESPRIT under the Grant CRIT-2 No. 20288.

the inflexibility is a disadvantage of this approach. Agents can either accept a task (by offering to perform it at a certain time and against a certain remuneration) or reject it. It is hardly possible to negotiate about the tasks in this model. E.g. an agent might offer to retrieve some information before offering a certain time-limit and price. It offers a good response time-limit and price, because it already happens to know where the information is located. However, if it had known that it is the only agent that was responding (the other agent being too busy) it could have negotiated a much better price. Although some of these limitations have been remedied in [12] the Contract Net protocol remains very inflexible.

As an alternative to a completely fixed protocol like Contract Net, one could consider a situation in which the participants are free to respond in whatever way suits them best at each point during the team formation. It means that the agents have to deliberate constantly about every move. This will be very time consuming and maybe less practical. We would like to have something in between completely fixed protocols and totally free conversations. A good candidate to fill this position is the theory of dialogue [16]. The dialogue theory tries to give rules for appropriate moves within different types of dialogues. The rules direct the dialogue without completely fixing the order of the moves.

The overall goal of this paper is to give a first account of team formation by means of dialogue. In the next paper [3], we give a generic account of team formation. In [7] a general framework is developed for using dialogues in collective problem solving ([6]). In this framework four stages are distinguished inspired by [17].

The first stage is *potential recognition*. In this stage the agent that takes the initiative for cooperative problem solving tries to find out which agents are potential candidates for achieving a given overall goal and how these can be combined in a team. The second stage is *team formation*. The result of this stage is a collective intention among a team to achieve the overall goal to solve the problem. This is the team that will try to actually achieve the goal. The third stage is *plan formation*. Here the team divides the overall goal into subtasks, associates these with actions and allocates these actions to team members (including a temporal structure). In terms of motivational attitudes the end result of this stage is a collective commitment to perform the social plan that realizes the overall goal. The last stage is *plan execution*. In this stage the team members execute the allocated actions and monitor the appropriate colleagues. If necessary a reconfiguration of the plan can be constructed [8], [9]. In [7] it is shown which types of dialogues play a role in each of these stages and in which way they can be embedded into each other. Also the results in terms of informational and motivational attitudes are given. In the present paper we will refine the theory and indicate all the steps necessary for the first two stages of the process. Thus we will show how the collective intention that is the result of team formation is built up during the dialogues.

The paper is structured in the following manner. Section 2 presents a brief typology of dialogue in CPS. Sections 3 and 4 investigate different dialogue types during potential recognition and team formation, respectively. Finally, in section 5 conclusions and further research are discussed.

2 Typology of dialogue in CPS

Dialogue can be defined as the 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 in the dialogue.

Krabbe and Walton [16] provide a typology of dialogue types between two agents, with an emphasis on the persuasion dialogue. They create a so called *a normative model*, representing the ideal way reasonable, cooperative agents participate in the type of dialogue in question. For each type of dialogue, they formulate *an initial situation, a primary goal*, and *a set of rules*. Below, their typology is briefly explained and adapted to the CPS perspective. In the course of communication among agents, there often occurs a shift from one type of dialogue to another, in particular *embedding* occurs when the second dialogue is functionally related to the first one.

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.

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.

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.

The initial situation for *deliberation* is a need for action performance. 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 is practical reasoning.

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.

The last type of dialogue i.e. *eristics* (verbal fighting between agents) has been left out as we focus on cooperative teams.

3 Potential recognition

Potential recognition is about finding the set of agents that may participate in the formation of the team that tries to achieve the overall goal.

These agents are grouped into several potential teams with whom further discussion will ensue during the team formation. Important starting point of the cooperative problem solving process is the overall goal that should be achieved in this process. This goal describes the solution of the problem (in an abstract way). It can arise in different ways. It can be a goal that is given from outside the system boundaries. E.g. a built in goal of the complete system such as "control the flights of all planes in a certain area" for an air traffic control system. The overall goal can also be formed by one agent that is interested in achieving that goal but does not have the abilities or desire to achieve it by itself. In this case this agent has to announce the goal as a possible collective goal for a group of agents. E.g. an agent tries to gather all possible information about some topic, but has only limited knowledge about the (possible) resources. It might at that moment ask for help of other agents to collectively gather all information.

In this paper both ways to form collective goals can be taken as point of origin. However, we restrict ourselves to the situation with one fixed overall goal and one agent that takes the initiative to achieve that goal. In future work we will also look at the process of collective goal formation, but we leave it out here because it would draw to much attention from the main goal of this paper and is an interesting subject in its own right.

The agent that takes the initiative to start the process of potential recognition can either be appointed or be the agent that is most interested to achieve the goal. The first usually is the case when the goal is given from outside. In that case there is usually one agent whose role it is to check this goal and to organize the team. In the air traffic control system this might be the air traffic coordinator.

The most important difference between the two cases is that in the case of an appointed initiator this agent usually does not have a self-interest in achieving the goal (more than any of the other agents). Consequently it does not have a predetermined role in the process for achieving the goal in mind for itself (except as being the initiator). Of course, often, the initiator will also be the responsible agent for achieving the overall goal, which means it will at least check whether the other agents fulfill their role.

In case the initiating agent is the agent that is most interested to achieve the goal, this agent usually also has a certain role in the process of achieving the overall goal. This role will lead to the achievement of one of its private goals that it finds important and that might coincide partly with the overall goal.

3.1 The end result of potential recognition

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. The initiator is interested in three aspects of the agents that can be approached for team formation.

1. their abilities

2. opportunities
3. their willingness to participate in the team formation

The first aspect considers whether the agents are suitable to participate in the team in a practical way (i.e. whether they can perform the right type of tasks, etc.). This deals with the subjective perspective of agent behaviour. It does not depend on the situation, but may be viewed as some kind of inherent property of the agent itself. The second aspect takes into account the possibilities of an action performance in the present situation. It involves resources and commitment strategies and possibly other such properties. This way the objective aspect of agent behaviour is expressed. The third aspect considers their mental attitude towards participating in the team. Very capable agents that do not want to do the job are of no use, neither are blindly committed agents that are already pursuing other goals.

The components of the 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 commitment strategy of an agent is denoted by *blind* (the agent is blindly committed to achieve its goals), *single-minded* (it can drop a goal only if it believes it can no longer achieve the goal) or *open-minded* (it can also drop a goal if it does no longer desire the goal).
4. the willingness to participate in team formation is denoted by $willing(b, \varphi)$.

We take Rao and Georgeff's semantics as a basis for our semantics, and refer the reader to [14, 15] for details, especially completeness and decidability proofs. However, we give a short description here, providing the background needed to follow this paper.

As a reminder, the temporal structure is a discrete tree branching towards the future, as in Computation Tree Logic CTL, [10]. The different branches in such a time tree denote the optional courses of events that can be chosen by an agent. Primitive events are those events that an agent can perform and that determine a next time point in the tree. The branch between a point and the next point is labeled with the primitive event leading to that point. For example, if there are two branches emanating from a single time point, one labeled "go to dentist" and the other "go shopping", then the agent has a choice of executing either of these events and moving to the next point along the associated branch. The temporal operators include $inevitable(\varphi)$ (in all paths through the point of reference φ holds), $optional(\varphi) \equiv \neg inevitable(\neg\varphi)$, $\Diamond\varphi$ (somewhere later on the same path, φ holds) and φ $\mathbf{U}$ ψ (φ *until* ψ, i.e. either φ holds forever on this path, or, as soon as it stops holding, ψ will hold). Formulas are divided into state formulas (which are true in a particular state) and path formulas (which are true along a certain path).

As to Kripke semantics, we consider each possible world to be a temporal tree structure as described above. Evaluation of formulas is with respect to a world w and a state s, using ternary accessibility relations B_i, D_i and I_i corresponding to each agent's beliefs, goals (or desires), and intentions, all of which lead from a pair of a world and a state in it to a world. Evaluation of formulas at world-state pairs is defined in the obvious manner inspired by CTL and epistemic logic.
$able(b, \psi)$ if ψ is in a pre-given set of agent b's abilities.
$opp(b, \psi)$ is true in a world and a state if there is a branch through that state such that at a later point on that branch, agent b has just seen to it that ψ. In [11] more formal definitions are given for ability and opportunity based on dynamic logic. The spirit of these definitions is captured in the ones given above except that we use states instead of actions.
As to commitments strategies, $\mathrm{COMM}(a, b, \varphi)$ stands for: agent a is socially committed towards agent b to fulfil φ (see [6, 9]). Note that these social commitments are weaker than collective commitments, the strongest motivational attitude discussed in [6, 9, 8]. Let us stress that they are also different from the similar concept used in [16]. The strongest commitment is followed by the *blindly committed* agent, who maintains its commitments until it actually believes that they have been achieved. Formally,

$$\mathrm{COMM}(a, b, \varphi) \to \mathit{inevitable}(\mathrm{COMM}(a, b, \varphi)\ \mathbf{U}\ \mathrm{BEL}(a, \varphi))$$

Single-minded agents may drop social commitments when they do not believe anymore that the commitment is realizable. However, as soon as the agent abandons a commitment, some communication and coordination with the other agent is needed. For open-minded agents, the situation is similar as for single-minded ones, except that they can also drop social commitments if they do not aim for the respective goal anymore. As in the case of single-minded agents, communication and coordination will be involved as expressed by the axiom:

$$\begin{aligned}
&\mathrm{COMM}(a, b, \varphi) \to \\
&\quad \mathit{inevitable}[\mathrm{COMM}(a, b, \varphi)\ \mathbf{U} \\
&\quad \{\mathrm{BEL}(a, \varphi) \vee \\
&\quad (\neg \mathrm{BEL}(a, \mathit{optional}\ \Diamond\varphi) \wedge \\
&\quad \mathit{done}(\mathit{communicate}(a, b, \neg \mathrm{BEL}(a, \mathit{optional}\ \Diamond\varphi))) \wedge \\
&\quad \mathit{done}(\mathit{coordinate}(a, b, \varphi))) \\
&\quad \vee(\neg \mathrm{GOAL}(a, \varphi) \wedge \\
&\quad \mathit{done}(\mathit{communicate}(a, b, \neg \mathrm{GOAL}(a, \varphi))) \wedge \\
&\quad \mathit{done}(\mathit{coordinate}(a, b, \varphi)))\}].
\end{aligned}$$

The willingness to participate in the team formation is modeled as the agent's belief that it is optional that it is possible that it will have the individual intention to reach the overall goal. Formally $\mathit{willing}(b, \varphi)$ stands for $\mathrm{BEL}(b, \mathit{optional} \diamond \mathrm{INT}(b, \varphi)))$

This does not conflict with the agent having different intentions $\text{INT}(b, \psi)$, even if ψ is inconsistent with the overall goal φ. Note that the agent's willingness to participate in team formation does not mean that the agent will necessarily be a member of the resulting team.
Agents can be more or less willing to participate in the team formation. If they have the overall goal already as their private intention, they are completely willing to participate in the team formation. On the other hand, if they do not want to participate at all, they believe never to intend to achieve the overall goal. Different intermediate positions are also possible. Metrics could be constructed in order to measure its willingness to participate. The metric could for instance be based on the percentage of branches in which the intention arises at some point and how far in the future these points lie.

3.2 Towards a potential of cooperation

In the previous section we have described what type of information the initiating agent tries to gather in order to start the team formation. In this section we will describe the formal outcome of this potential recognition stage, how the information is collected, and how this leads to this outcome.
The output at this stage is the "potential for cooperation" that the initiator a sees with respect to φ, denoted as $\text{POTCOOP}(a, \varphi)$, meaning that φ is a goal of a ($\text{GOAL}(a, \varphi)$), and there is a group G such that a believes that G can collectively achieve φ ($\text{C-CAN}_G(\varphi)$) and that they have the right distribution of commitment strategies to do so ($\text{STRAT}(\varphi, G)$) and are willing to participate in team formation. This is expressed by the following definition (of which the refined version may be found in [8]):

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

$\text{POTCOOP}(a, \varphi)$ is derived from the information collected from the individual agents. To derive $\text{C-CAN}_G(\varphi)$ 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 *constitute* φ and are compared with the individual abilities and opportunities that the agents are believed to have. Formally:

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

Unfortunately we cannot define the right distribution of commitment strategies ($\text{STRAT}(\varphi, G)$) in the same way in terms of individual commitment strategies. The reason is that this notion is not compositional. In fact we have to take the overall distribution of commitment strategies as primitive and regard it as a kind of constraint on individual commitments. Whenever the set of individual commitment strategies of the members is a suitable distribution for the whole group then the individual commitment strategies are also suitable. See [7] for a formal definition of suitable individual commitment strategies.

3.3 Information seeking dialogue

The initiator can find out the above properties of the agents by asking them or by using past experience or a database containing part of this information. The questions in this stage form part of an *information seeking* dialogue. The initiator has to form beliefs about the abilities, opportunities, commitment strategies and willingness of the individual agents in order to derive $\mathrm{POTCOOP}(a, \varphi)$. This can be done by asking every agent about its properties and the agent responding with the requested information. Formally this can be expressed as follows:

$$[REQ_{a,i}(ASS_{i,a}(able(i,\varphi_i)) \cup ASS_{i,a}(\neg able(i,\varphi_i)))]$$
$$[ASS_{i,a}(able(i,\varphi_i))]\mathrm{BEL}(a,(able(i,\varphi_i)))$$

The above formula is based on the formal theory on *speech acts* developed in [13] and [5]. 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 φ will always hold.

In the above case α_1 is the complex action $REQ_{a,i}(ASS_{i,a}(able(i,\varphi_i)))\cup ASS_{i,a}(\neg able(i,\varphi_i)$ where $REQ_{a,i}(\alpha)$ stands for agent a requesting agent i to perform the action α. The action it requests i to perform is to either assert that it is able to achieve φ_i or to assert that it is not.

After this request i has three options.

1. It can simply ignore a and not answer at all.
2. It can state that it is not willing to divulge this information (formally expressed as $ASS_{i,a}(\neg(ASS_{i,a}(able(i,\varphi_i)) \cup ASS_{i,a}(\neg able(i,\varphi_i)))))$.
3. it can either assert that it is able to achieve φ (as is described in the formula above) or that it is not, in which case a believes that i is not able to achieve the goal: $[ASS_{i,a}(\neg able(i,\varphi_i))]\mathrm{BEL}(a,(\neg able(i,\varphi_i)).$

Only in the last two cases will a have a resulting belief about the ability of i. Of course in the case that i does not want to divulge the information a can already derive that i is not willing to achieve φ as part of a team.

The same sequence as is shown above to seek information about the abilities of agents can be followed for the opportunities, commitment strategies and willingness of agents to participate in team formation.

We conclude this section with three remarks on some practical aspects on the above procedure of information seeking.

The first is a technical point about the exact questions that are asked by a. In the formulas above, a asks for the ability to achieve a specific goal φ_i. In principle, using this scheme, a should repeat this question for all the subgoals it distinguished to every agent. Of course it is more

practical for agent a to ask each agent i about all the abilities it has with respect to achieving this set of goals. This can also be expressed formally in our logic, but is quite complex, and not very readable. Therefore we leave out this sophistication here.
A second point is about the case when agent i simply does not respond to the question. To avoid agent a waiting indefinitely for an answer, we assume that every speech act has an implicit deadline for reaction incorporated. The modeling of these type of deadlines is described in e.g. [4], but is not further pursued here.
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 or not. It would be useful for the other agents to know this fact as well, in order to revise their motivational and informational attitudes. In case no team could be formed or the agent is not considered as part of such a team formation, the agent can pursue other tasks again.
This information might be divulged to all agents by using a broadcast message of a stating the formal outcome of this stage.

4 Team formation

At the start of this stage, the initiator has a sequence of groups in mind that could be suitable to form a team for achieving the overall goal. The sequence may be created arbitrarily or based on the properties of the agents involved. Although we do not go into details here, the organization of the groups into a sequence is important for the case that the most desirable group can not be formed.
All agents in these potential teams have expressed their willingness to participate in team formation, but do not necessarily have the individual intention to contribute towards achieving that goal yet. In this situation, the initiator tries to persuade all agents from a potential team to take on the intention to achieve the overall goal as well as to perform its own role within the team. This role corresponds to one of the temporary subgoals from a sequence $\varphi_1, \ldots, \varphi_n$ found during potential recognition, of which the initiator believes that $constitute(< \varphi_1, \ldots, \varphi_n >, \varphi)$.
The first decision the initiator has to make is with whom to start a *persuasion dialogue*. This decision could be based on a partial preferential ordering among agents, based on the initiator's information about their relevant abilities and degree of willingness to participate in team formation. It seems logical to start with an agent who has many relevant abilities and is highly willing to participate.
In the allocation of subgoals to agents, a process similar to constraint satisfaction takes place. In this case, the variables may take values that fall within the range corresponding to each agent's abilities and willingness. The aim of constraint satisfaction is to assign a value to each variable without depleting the domains of other variables. In our case we try to assign a temporary subgoal to each agent. However, in this case the values that can be assigned per agent are not fixed from the start: they may be the subject of negotiation. Techniques to find a proper order of agents with whom to start the persuasion dialogue and the preliminary subgoal about which to persuade them may be derived from

well-known techniques from constraint satisfaction[1]. For example, the most constrained agent, e.g. the agent who is the only one able to achieve a certain crucial subgoal, might be a good starting point. Also, once the agent has been selected, the first subgoal about which to start persuading it could be the most crucial one.

4.1 The end result of team formation

The end result that the initiator wants to achieve during this stage is a collective intention (C-INT$_G(\varphi)$ among a team G to reach the overall goal φ. A necessary condition for this is that all members of the group have the associated individual intention INT(i, φ), and that all members of the group are aware of this. That is, they have a collective belief about their intentions (C-BEL$_G$(E-INT$_G(\varphi)$)). Moreover, to exclude the case of competition, all agents should *intend* the other members to have the associated individual intention, and it should be a collective belief in the group that this is so. In order to formalize the above two conditions, E-INT$_G(\varphi)$ (standing for "everyone intends") is defined by E-INT$_G(\varphi) \leftrightarrow \bigwedge_{i \in G}$ INT(i, φ). Formally:

$$
\begin{aligned}
&\text{C-INT}_G(\varphi) \leftrightarrow \text{E-INT}_G(\varphi) \wedge \text{C-BEL}_G(\text{E-INT}_G(\varphi)) \wedge \\
&\quad \text{E-INT}_G(\text{E-INT}_G(\varphi)) \wedge \\
&\quad \text{C-BEL}_G(\text{E-INT}_G(\text{E-INT}_G(\varphi)))
\end{aligned}
$$

The above formula 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. In practice this will mean that the initiator also persuades them to be willing to take on an appropriate temporary subgoal (this is much less strong than an intention [14]). To establish the second part of the above formula, the initiator also persuades each member to take on the intention that all other members of the potential team intend to contribute to the overall goal, in order to foster cooperation from the start. When all the individual motivational attitudes are established within the team, the initiator broadcasts the fact E-INT$_G(\varphi)) \wedge$ E-INT$_G$(E-INT$_G(\varphi)$), by which the necessary collective beliefs are established and the collective intention is in place.

4.2 Persuasion dialogue

During the persuasion dialogue the initiator has to convince the other agent to form an intention to reach the overall goal as part of a team of agents that intend to achieve that goal. According to [16] the persuasion dialogue consists of three main stages: information exchange, rigorous persuasion and completion. In our case the information exchange already started in the potential recognition stage during which abilities, opportunities and willingness of the agents are exchanged.

Information exchange During the information exchange the agents make clear their initial stand with respect to the overall goal, the willingness to take on a certain subgoal and to form part of a certain team to achieve the overall goal. 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
   $ASS_{i,a}(\mathrm{INT}(i,\varphi) \wedge willing(i,\varphi_i) \wedge \bigwedge_{j \in G} \mathrm{INT}(i, \mathrm{INT}(j,\varphi))$. 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:

   $$[ASS_{a,i}(won)]O(ASS_{i,a}(\mathrm{INT}(i,\varphi) \wedge \\ willing(i,\varphi_i) \wedge \bigwedge_{j \in G} \mathrm{INT}(i, \mathrm{INT}(j,\varphi)))$$

   which means that agent i is obliged to state that it gives up and will play 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").

7. All challenges have to follow logical rules. E.g. 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.

Goal formation through speech acts In this section we discuss goals and intentions and their relation. Although we do not give a formal semantics and/or definition, we assume that goals do not necessarily have to be consistent, but intentions do have to be consistent (leading to a K_n system for goals and a KD_n system for intentions). Also we assume that for each intention there is a goal that generated it and thus we have: $\mathrm{INT}(x, \varphi) \rightarrow \mathrm{GOAL}(x, \varphi)$.

Intentions are formed on the basis of beliefs and previously formed goals of a higher abstraction level by a number of formal rules (see [2]). E.g. the built-in goal can be to obey the law, or avoid the punishment. The (instrumental) belief is that driving slower than the speed limit is instrumental for obeying the law. Together with the rule the new goal of driving slower than the speed limit is derived.

The general intention generation rule is represented as follows:

$$\mathrm{GOAL}(x, \varphi) \ \wedge \ \mathrm{BEL}(x, INSTR(\psi, \varphi)) \ \rightarrow \ \mathrm{INT}(x, \psi)$$

It states that if an agent x has a goal φ and it believes that ψ is instrumental in achieving φ then it will have the intention to achieve ψ. If ψ is instrumental in achieving φ then it means achieving ψ gets the agent "closer" to φ in some abstract sense. We do not define this relation any further, but leave it as primitive.

We acknowledge that the above rule is an oversimplification of reality. In most cases it will be too strong. But we see it as an example of a possible rule that links goals and intentions and do not explore more realistic, but also more complex, rules in this paper.

This very general rule for generating intentions can be combined with more specific ones, e.g. goal adoption, in which case an agent adopts the goals of another agent. One more specific rule for goal adoption can be given by:

$$\begin{array}{l} \mathrm{GOAL}(x, helped(x, y)) \wedge \\ \mathrm{BEL}(x, [\mathrm{GOAL}(y, \varphi) \rightarrow INSTR(OBT(y, \varphi), helped(x, y))] \end{array}$$

Agent x has as abstract goal to help y. It believes that if y has a goal φ then it is instrumental for x to achieve its abstract goal to help y if it helps y obtain its goal. Where

$$OBT(y, \varphi) \ \equiv \ \varphi \ \wedge \ \mathrm{GOAL}(y, \varphi) \ \mathbf{U} \ \varphi$$

So, agent y obtained φ if φ is true and it had the goal φ until the moment φ became true.

Together with the general intention generation rule we can now directly derive $\text{INT}(x, OBT(y, \varphi))$ whenever $\text{BEL}(x, \text{GOAL}(y, \varphi))$.
The above sketched mechanism can be used in our setting during persuasion. In this case the initiator tries to get the other agent to concede to higher level goals and instrumental beliefs that together with the general intention generation rule imply the intention to achieve the overall goal of the system. Here follows an example move of the initiator.

$$ASS_{a,i}(\forall j(\text{GOAL}(j, \psi)) \to INSTR(\varphi, \psi))$$

After this speech act agent i knows that the initiator believes that if an agent has the goal ψ then the overall goal φ is instrumental to this goal. Formally (see also [5]):

$$\begin{array}{l}[ASS_{a,i}(\forall j(\text{GOAL}(j, \psi)) \to INSTR(\varphi, \psi))] \\ \text{BEL}(i, \text{BEL}(a, (\forall j(\text{GOAL}(j, \psi)) \to INSTR(\varphi, \psi))))\end{array}$$

Let us assume that i has only two rules for answering. If i does not have a belief that is inconsistent with a belief of a then i will concede. If, on the other hand i does have a belief to the contrary it will challenge the assertion. Formally:

$$\neg\text{BEL}(i, \neg\gamma) \wedge \text{BEL}(i, \text{BEL}(a, \gamma)) \to DO(i, CONCEDE_{i,a}\gamma)$$

$$\text{BEL}(i, \neg\gamma) \wedge \text{BEL}(i, \text{BEL}(a, \gamma)) \to DO(i, CHALLENGE_{i,a}\gamma)$$

where the operator $DO(i, \alpha)$ indicates that α is the next action performed by i. We do not give a formal definition of the $CONCEDE$ and $CHALLENGE$ actions, but will only give their intuitive meaning here. The concede action has as effect that the agent is willing to assume the assertion during the dialogue. The challenge is a combination of a denial and a request to prove. We will come back to it shortly.

Suppose that i did not have a contrary belief then i concedes the above belief by the speech act

$$CONCEDE_{i,a}(\forall j(\text{GOAL}(j, \psi)) \to INSTR(\varphi, \psi))$$

The effect of this speech act is the same as for the ASS, except that a can only assume that i believes the formula during the dialogue and might retract it afterwards.

$$\begin{array}{l}[CONCEDE_{i,a}(\forall j(\text{GOAL}(j, \psi)) \to INSTR(\varphi, \psi))] \\ \text{BEL}(a, \text{BEL}(i, (\forall j(\text{GOAL}(j, \psi)) \to INSTR(\varphi, \psi))))\end{array}$$

Now the formula is a belief of both a and i. In this case, the initiator's next aim in the persuasion will be to get i to concede to the belief that ψ is a goal of i by the question

$$REQ_{a,i}(CONCEDE_{i,a}(\text{GOAL}(i, \psi))).$$

The effect of this request is given formally through:

$$\begin{array}{l}[REQ_{a,i}(CONCEDE_{i,a}(\text{GOAL}(i, \psi))] \\ O(CONCEDE_{i,a}(\text{GOAL}(i, \psi) \cup ASS_{i,a}(\neg\text{GOAL}(i, \psi)))\end{array}$$

That is, i is obliged to either concede it has the goal ψ or assert the negation of it. After i's response the initiator knows whether or not i believes it has the goal ψ.
If i does not concede to the initiator's first speech act, it may instead counter with a challenge by first asserting that it doesn't share the initiator's belief and then asking the initiator to argue for the belief. This is done by the complex speech act $CHALLENGE$. This action can be described as follows:

$$\begin{aligned} CHALLENGE_{i,a}\ & (\forall j(\mathrm{GOAL}(j,\psi) \rightarrow INSTR(\varphi,\psi))) \equiv \\ & \neg ASS_{i,a}(\forall j(\mathrm{GOAL}(j,\psi) \rightarrow INSTR(\varphi,\psi))); \\ & CONCEDE_{i,a}(\mathrm{GOAL}(j,\psi)) \\ & REQ_{i,a}(ASS_{a,i}(PROOF(INSTR(\varphi,\psi)))) \end{aligned}$$

The complete effects of this speech act are quite complex to describe formally, so, we leave it out here. The example gives an idea how the intention to achieve the main goal may be formed by persuasion. For intentions with respect to subgoals and with respect to cooperation with other potential team members, the process is analogous. In the latter case, the initiator tries to get the other agent to form beliefs that will trigger the formation of the intention that all other potential members also intend to achieve the overall goal. The final result of the team formation stage is reached when for one potential team all the persuasion dialogues have been concluded successfully.

Negotiation embedded in the persuasion dialogue Krabbe and Walton ([16]) do not allow for embeddings of negotiations in persuasion dialogues. However, we feel that in contrast to persuasion with respect to informational attitudes, bargaining may be appropriate within a persuasion dialogue with respect to motivational attitudes (intentions). Thus, during team formation the initiator may persuade other agents to take on the overall goal by offering them return favors. Also, an agent may offer willingness to take on a subgoal and trade this for a promise of the initiator to take care during the later stage of plan formation that this will be the (only) subgoal for the agent. A prospective member may even negotiate by trading its membership for the inclusion or exclusion of certain other agents. However, negotiation about the overall goal is not possible in our setting.
Although the topic is too complex to give a complete account on how the negotiation is translated in logical terms, we still will show informally how this process works within the logic underlying our theory.

At the start of the negotiation both agents have a number of possible future courses of action which are preferentially ordered. That is, they have a preferred course of action which they will try to realize first. If the agent notices that this course of action is no longer realizable or preferred it might switch to another branch.
The preferential ordering of the courses of action is based on the beliefs of the agent. E.g. we believe that writing this paper will make us famous. The goal of the negotiation is to change the beliefs of the other agent such

that the preferential orders change so that they fit with each other. Not all beliefs of an agent can be changed at will by speech acts of another agent. This depends on the situation and status of the agents. Therefore, it might be that the agents are unsuccessful in their attempts to change the preferential ordering of the other agent sufficiently (cf. "the bottom line").

In our case the initiator knows at the start of the negotiation that the other agent believes that the overall goal will be its intention during some course of action at some future time. The initiator tries to change the beliefs of the agent in such a way that this branch becomes the preferred course of actions for the other agent. Of course there might be several branches on which the agent adopts the goal, in that case the initiator tries to find its most preferred one.

We will not give a precise set of rules for negotiation, but give some examples of the differences with persuasion. The situation in negotiation is more symmetrical than in the case of persuasion. Both agents may propose theses as well as concessions, so the opponent O has more options than a fixed reaction to the previous move of the proponent. An agent is obliged not to retract its concessions (e.g. to take on a subgoal). The successful end result of negotiation is not that one agent wins, but that a deal is constructed to which both agents agree.

Termination of the negotiation dialogue Because the negotiation dialogue is more symmetrical and both agents can propose new theses and concessions the termination of this type of dialogue is much less clear than that of the persuasion dialogue. Also the end of the negotiation does not imply a winner. Both parties might agree that no deal can be formed and end the negotiation. We assume in our situation that the initiative for this move lays with the initiator. As soon as it finds out that the negotiation cannot be concluded successfully it will stop not only the negotiation but also the persuasion dialogue in which it is embedded.

The negotiation dialogue can thus be concluded in three ways.

1. A successful fit of the courses of actions of the two agents has been found. The initiator announces this by asserting the achieved deal (i.e. common intention).
2. A repetition of moves by one of the two parties is seen as a signal that the relevant beliefs of this party are not to be changed anymore. The initiator then announces the assertion $ASS_{a,i}(quit)$, which also concludes the persuasion dialogue.
3. The second agent i announces $ASS_{i,a}(quit)$; in this case the initiator replies with the assertion $ASS_{a,i}(quit)$, which again also concludes the persuasion dialogue.

In the case of successful termination of negotiation, the agents return to the persuasion dialogue where the deal may be used by the initiator as a concession of the other.

5 Conclusions

In previous work ([8] and [9]) it was shown how Cooperative Problem Solving can be divided into four stages, each of them resulting in specific motivational attitudes that can be formally described. Then in [7] the stages were related to specific types of dialogue (possibly embedded into each other) by which these motivational attitudes can be formed. In this paper, finally, we have shown for the first two stages of potential recognition and team formation which rules govern these dialogues and how the moves within the dialogues can be expressed by formally represented speech acts. The complete description of all elements of the dialogues in a formal logic has thus been made possible. However, many elements of the logic still have to be expanded and improved upon, before one can state that the dialogue has been described formally. Once this is done it will be possible to prove that in given circumstances the dialogue results in a certain outcome. This is of prime importance if you want to construct a MAS for automated cooperative problem solving.

It is clear that, even though the dialogues are governed by strict rules, the reasoning needed to find an appropriate move is highly complex. This implies that the agents also have to contain complex reasoning mechanisms in order to execute the dialogues. It means that, although the result is much more flexible and refined than using a protocol like Contract Net, the process is also more time consuming. For practical cases one should carefully consider what carries more weight and choose the method of team formation accordingly.

A final result of this ongoing research should contain a complete set of formal rules for all the types of dialogue and indicate how these are implemented through formal speech acts. This would make it possible to extend the framework to the next stages of cooperative problem solving (plan formation and execution).

It also should expand on several aspects of the internal reasoning of the agents. One example is the concept of giving a proof as defense of an assertion during the rigorous persuasion. Another is the actual formation of intentions based on previous beliefs and goals.

Another issue for further research is to investigate the consequences of not having a fixed overall goal, but an emerging goal. Another point is to check what should be done in case there is not one initiating agent (but possibly many or none).

Finally it should be investigated how actual proofs can be constructed in an efficient way to prove that the end results of a dialogue are formed through the speech acts given the rules of the dialogue.

References

1. R. Dechter. Enhancement Schemes for Constraint Processing: Backjumping, Learning and Cutset Decomposition. *Journal of Artificial Intelligence*, vol 41, 1990, pp. 273–312.
2. F. Dignum and R. Conte. Intentional agents and goal formation. In M. Singh et.al., editor, *Intelligent Agents IV (LNAI 1365)*, Springer Verlag, 1998, pp. 231-244.

3. F. Dignum, B. Dunin-Kęplicz and R. Verbrugge. *Creating Collective Intention through Dialogue.* Technical Report of Cognitive Science and Engineering, TCW-1999-16, University of Groningen, 1999.
4. F. Dignum and R. Kuiper. Combining dynamic deontic logic and temporal logic for the specification of deadlines. In Jr. R. Sprague, editor, *Proceedings of thirtieth HICSS*, Wailea, Hawaii, 1997.
5. F. Dignum and H. Weigand. Communication and deontic logic. In R. Wieringa and R. Feenstra, editors, *Information Systems, Correctness and Reusability*, World Scientific, Singapore, 1995, pp. 242-260.
6. B. Dunin-Kęplicz and R. Verbrugge. Collective Commitments. In *Proc. Second International Conference on Multi-Agent Systems, ICMAS'96*, IEEE Computer Society Press, Kyoto, 1996, pp. 56-63.
7. B. Dunin-Kęplicz and R. Verbrugge. The Role of Dialogue in Collective Problem Solving: Part I. To appear in *Proceedings of the Agent Day*, Krakow, 1999.
8. B. Dunin-Kęplicz and R. Verbrugge. A Reconfiguration algorithm for distributed problem solving. To appear in *Engineering Simulation: An International Journal of Electrical, Electronic and other Physical Systems*, special issue about distributed systems, 1999.
9. B. Dunin-Kęplicz and R. Verbrugge. Collective motivational attitudes in cooperative problem solving. In V. Gorodetsky et al. (eds.), *Proceedings of The First International Workshop of Central and Eastern Europe on Multi-agent Systems (CEEMAS'99)*, St. Petersburg, 1999, pp. 22-41.
10. E.A. Emerson. Temporal and Modal Logic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, North-Holland, Amsterdam, 1989, pp. 995-1072.
11. B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. Formalising abilities and opportunities of agents. *Fundamenta Informaticae* 34 (1998) pp. 53-101.
12. T. Sandholm and V. Lesser. Issues in automated negotiation and electronic commerce: extending the contract net protocol. In *Proceedings First International Conference on Multiagent Systems (ICMAS95)*, San Francisco, AAAI Press and MIT Press, 1995. pp. 328-335.
13. J.R. Searle and D. Vanderveken. *Foundations of Illocutionary Logic*, Cambridge, Cambridge University Press, 1985.
14. A.S. Rao and M.P. Georgeff. Modeling rational agents within a BDI architecture. In R. Fikes and E. Sandewall (eds.) *Proceedings of Knowledge Representation and Reasoning (KR&R-91)*, San Mateo, Morgan Kaufmann, 1991. pp.473-484.
15. A.S. Rao and M.P. Georgeff. Formal Models and Decision Procedures for Multi-agent Systems. Technical Note 61, Carlton (Victoria), Australian Artificial Intelligence Institute, 1995.
16. D. Walton and E. Krabbe. *Commitment in Dialogue*, SUNY Press, Albany, 1995.
17. M. Wooldridge and N.R. Jennings. Cooperative Problem Solving. *Journal of Logic and Computation* 9 (4) (1999), pp. 563-592.