A Reconfiguration Algorithm for Distributed 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@tcw2.ppsw.rug.nl

Abstract

One of the main aspects in Distributed Problem Solving (DPS) in Multi-Agent Systems is that, due to the dynamic and possibly unpredictable environment, team members may fail their tasks or be presented with new opportunities. Thus, it is necessary that team members monitor their performance and replan based on the present situation. This leads to the *reconfiguration problem.*

The paper presents essential aspects of a methodology of teamwork based on the proper maintenance of the individual, social and collective motivational attitudes in a group of heterogeneous agents. The key concept is that of *collective commitment*, formally introduced in [5]. The main phases of DPS, namely construction, maintenance, and realization of collective commitments, are assigned to a four-level abstract architecture. In terms of these generic levels and their interplay an efficient and flexible *reconfiguration algorithm* is specified. Finally, an application of the reconfiguration algorithm in a distributive environment is presented.

1 Introduction

Multi-Agent Systems (MAS) are computational systems in which a collection of (loosely-coupled) autonomous agents interact in order to solve a given problem. As this problem is usually beyond the agents' individual capabilities, agents can communicate, cooperate, coordinate, and negotiate with one another, both in advancement of their own goals as well as for the good of the system as a whole.

Some MAS are referred to as *intentional systems* realizing the practical reasoning paradigm ([1]). The best known and most influential are so-called *belief-desire-intention systems.* BDI-agents are characterised by a "mental state" described in terms of *beliefs*, corresponding to the information the agent has about the environment; *desires*, representing options available to the agent; and *intentions* representing the chosen options.

While beliefs are viewed as agent's *informational* attitudes and desires as its *goals*, intentions and, next, commitments refer to its *motivational* attitudes. From the perspective of teamwork, motivational attitudes are considered on three levels: individual, social (i.e. bilateral), and collective. The process of building up social and collective motivational attitudes starting from individual ones, is formally described in [5]. In the present paper some definitions will be recalled when necessary.

In some recent BDI systems teamwork is modelled explicitly. A dynamic environment poses the problem that team members may fail to bring their tasks to a good end or new opportunities may appear. The explicit model helps the team to monitor its performance and especially to replan based on the present situation. This leads to the so-called *reconfiguration problem* which will be the main subject of this paper.

In order to study the question, we isolate basic stages involved in teamwork in a distributed environment . These are construction, maintenance, and realization of collective commitments. We base our analysis on the four-stage model of [15], 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. This architecture will be profitable in clarifying the nature of dependencies between agents involved: some of them will do domain problem solving, others will be responsible for the proper organization of DPS. 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.

Especially with respect to collective intentions and collective commitments, our analysis differs from the one in [15]. In fact, in [5] we studied teamwork starting from group formation based on a collective intention, and leading up to the strongest motivational attitude, collective commitment. Our approach to collective commitment is plan-based: the collective intention is split up into subtasks, according to a social plan. Next, the task allocation is reflected in social commitments between pairs of agents.

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. For example, the reconfiguration problem mentioned above has only recently come to be discussed. The main contribution of this paper is the *reconfiguration algorithm* formulated in terms of the abstract levels and their (complex) interplay. An example application of the algorithm is extensively discussed. This research is a continuation of the work presented in [5, 6].

The paper is structured in the following manner. In Section 2 we recall the semantics of individual and social motivational attitudes and the formal definition of a collective commitment. Section 3 presents the four level description of the abstract system. In terms of these generic abstract levels a general-purpose reconfiguration algorithm is defined in Section 4. Next, an example refinement of this algorithm is discussed in Section 5. Finally, Section 6 presents conclusions and future perspectives.

2 Preliminaries

In Rao and Georgeff's [12], individual beliefs, goals, and intentions are formalized as primitive notions and given a formal semantics. We take their semantics as a basis for our formalization of collective motivational attitudes, and refer the reader to [12] for details. Extensive descriptions of our own definitions of social commitments and collective intentions, as well as of the well-known notion of collective belief, can be found in [5]. In the present paper, short descriptions of the motivational attitudes are repeated when needed. For ease of notation, we use formulas of a multi-modal logic, as in [5]. Relevant formulas and their meanings are recalled in table 1. As a reminder, the temporal structure

$\mathrm{BEL}(a, \varphi)$	agent a beliefs that φ
$\mathrm{GOAL}(a, \varphi)$	agent a has as a goal that φ be true
$\mathrm{INT}(a, \varphi)$	agent a has the intention to make φ true
$\mathrm{COMM}(a, b, \varphi)$	agent a socially commits to agent b to make φ true
$\text{C-BEL}_G(\varphi)$	group G has the collective (mutual) belief that φ
$\text{C-INT}_G(\varphi)$	group G has the collective intention to make φ true
$\text{C-COMM}_{G,P}(\varphi)$	group G has a collective commitment to make φ true by plan P

Table 1: Individual, social and collective attitudes introduced in [5]

is a discrete tree branching towards the future, as in [7]. The temporal operators include among others *inevitable*(φ) (in all paths through the point of reference φ holds). In [15] a similar multi-modal language and the same time structure are used, so we may easily adapt some of their definitions when needed.

Social commitments between two agents were defined in [5], inspired by the discussion in [2]. A social commitment from one agent to the other is not as strong as a collective commitment among them, but stronger than an individual intention, because the other agent is interested in the first one acting on his agreement. Moreover, the agents are aware about the situation, i.e. about their individual attitudes. Such awareness, expressed in terms of collective belief, is generally achieved by communication:

$$\mathrm{COMM}(a, b, \varphi) \leftrightarrow \mathrm{INT}(a, \varphi) \wedge \mathrm{GOAL}(b, \mathit{stit}(a, \varphi)) \wedge \text{C-BEL}_G(\mathrm{INT}(a, \varphi) \wedge \mathrm{GOAL}(b, \mathit{stit}(a, \varphi))),$$

where $\mathit{stit}(a, \varphi)$ means that agent a sees to it (takes care) that φ becomes true.

In [5] our main contribition was a formal definition of collective commitment. Inspired by [2], we treat collective commitment as the strongest motivational attitude in teamwork. A collective intention may be viewed as inspiration for team activity, whereas the collective commitment reflects the concrete manner of achieving the intention. This concrete manner is provided by planning. We see planning, including means-ends analysis, as a two-step process. The first step is *task division* or decomposition, in which a complex task φ is decomposed into (possibly also complex) subtasks $\varphi_1, \ldots, \varphi_n$. 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 $< \alpha_i, a >$ of a (possibly complex) action α_i that realizes task φ_i and an agent a. To make a social plan complete, the temporal structure among these pairs is established. For a recursive definition of well-formed social plan expressions, see [11].

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. In addition, for every one of the subgoals $\varphi_1, \ldots, \varphi_n$ that together constitute the goal φ in question reflected in the plan P, there should be one agent in the group who is socially committed to at least one (mostly other) agent in the group to fulfil the subgoal. Moreover, there should be a common belief in the whole group that the plan will be entirely realised, i.e. that all subgoals have been adopted by committed members of the group. The defining axiom below reflects all these characteristics.

$$\text{C-COMM}_{G,P}(\varphi) \leftrightarrow \text{C-INT}_G(\varphi) \wedge \bigwedge_{\varphi_i \in P} \bigvee_{a,b \in G} \mathrm{COMM}(a, b, \varphi_i) \wedge$$

$$\text{C-BEL}_G\Big(\bigwedge_{\varphi_i \in P} \bigvee_{a,b \in G} \mathrm{COMM}(a, b, \varphi_i)\Big)$$

$\mathrm{KNOW}(a, \varphi)$	agent a knows that φ
$\text{C-KNOW}_G(\varphi)$	group G has the collective (common) knowledge that φ
$\mathrm{AGENTS}(\alpha, G)$	group G is precisely the set of agents required to perform action α;
$\mathrm{STRAT}(\varphi, G)$	group G have a suitable distribution of commitment strategies to achieve φ;
$\mathit{happens}(\alpha)$	action α happens next on the current path;
$\mathrm{ACHIEVES}(\alpha, \varphi)$	$\mathit{inevitable}(\mathit{happens}(\alpha) \to \mathit{happens}(\alpha; \varphi?))$ i.e. on all paths, after action α happens, proposition φ is true;
$\mathrm{CAN}(a, \varphi)$	$\exists\beta(\mathrm{KNOW}(a, \mathrm{AGENTS}(\beta, \{a\}) \wedge \mathrm{ACHIEVES}(\beta, \varphi)))$ i.e. there is a complex action β such that a knows that a itself is the single agent required to perform β, and that β achieves φ;
$\text{C-CAN}_G(\varphi)$	$\exists\beta(\text{C-KNOW}_G(\mathrm{AGENTS}(\beta, G) \wedge \mathrm{ACHIEVES}(\beta, \varphi)))$.

Table 2: Formulas needed for the definition of potential recognition

In this definition the team as a whole is aware that social commitments between team members have been established in the proper way, without explicit knowledge of each bilateral commitment. This kind of knowledge can be incorporated when needed.

3 The four level abstract description of the system

The analysis of DPS within MAS led us to the distinction of four levels of abstraction, corresponding to the four stages defined in [15]. When characterizing them, we will take the perspective of a BDI architecture: the key point is to bind particular levels of abstraction with appropriate social, collective and individual motivational attitudes. We suppose that the main goal of the system is given at the beginning. The problem of goal selection is beyond the scope of this paper and has been extensively treated elsewhere [4].

3.1 The four levels

Level 1: the *potential recognition* level.
Analogous to [15], we consider DPS to begin when some agent in a multi-agent environment recognizes the potential for cooperative action in order to reach his goal. The input of this stage is an agent a, a goal φ plus a finite set T of agents from whom a potential team may be formed. The output at this stage is the "potential for cooperation" that agent a sees with respect to φ, to be defined below by an analog of the definition of [15]. A few preliminary definitions are needed, which can be looked up in table 2. Most of these are also taken over from [15], except that we prefer non-recursive definitions of individual and group ability (CAN and C-CAN_G). Our choice is motivated by the need to make revision explicit in the algorithm presented in the next section. Moreover, we introduced a new predicate $\mathrm{STRAT}(\varphi, G)$ reflecting possible combinations of commitment strategies suitable for achieving φ. The definitions of agents' commitment strategies, addressing the question whether and in which circumstances the agent is allowed to responsibly drop a social commitment, are introduced in [5]. Inspired by those in [10], they additionally obey the social aspects of communication and coordination.

The strongest commitment strategy is followed by the *blindly committed* agent, who maintains its commitments until it actually believes that they have been achieved. 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 previous case, communication and coordination are needed. It was shown in [5] that the agents' commitment characteristics predispose them to do different tasks. We suggest that agent a take these characteristics into account, starting already from potential recognition; this is reflected in the conjunct $\mathrm{STRAT}(\varphi, G)$ below.

Here follows the definition of "agent a sees potential for cooperation to achieve φ":

$$\begin{aligned}
&\mathrm{POTCOOP}(a,\varphi) \leftrightarrow \mathrm{GOAL}(a,\varphi)\wedge \\
&\quad \exists G \subseteq T \mathrm{BEL}(a, \text{C-CAN}_G(\varphi) \wedge \mathrm{STRAT}(\varphi, G)) \wedge \{\neg \mathrm{CAN}(\alpha,\varphi) \vee \\
&\quad \mathrm{BEL}(a, \neg\exists\alpha[\mathrm{AGENTS}(\alpha,\{a\}) \wedge \mathrm{ACHIEVES}(\alpha,\varphi) \wedge \mathrm{GOAL}(a,\alpha)])\}
\end{aligned}$$

More informally: φ is a goal of a, and there is a group G such that a believes that G can jointly achieve φ and that they have the right distribution of commitment strategies to do so; and either a cannot or doesn't desire to achieve φ in isolation. As output of a successful outcome of this stage, track is also kept of the sequence $(G_1, \ldots, G_n)$ of all relevant groups $G \subseteq T$ for which the condition $\mathrm{BEL}(a, \text{C-CAN}_G(\varphi) \wedge \mathrm{STRAT}(\varphi, G))$ was found to be true in the establishment of $\mathrm{POTCOOP}(a, \varphi)$.

Level 2: the *team formation* level.

Suppose that agent a sees the potential for cooperation to achieve φ. Somewhat different from [15], we find that during the team formation stage agent a attempts to bring it about that some group G has the *collective intention* to make 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 among G to achieve φ, which includes corresponding individual intentions of all group members. This is done by subsequently attempting to establish the collective intention among G_1, G_2 etc., until this succeeds for some G_i. The sequence of still untried potential groups $(G_{i+1}, \ldots, G_n)$ is stored for revision purposes.

Let us recall from [5] the definition of a collective intention. A necessary condition is that all members of the group have the associated individual intention, and that the group is aware of this. Moreover, all members in the group should *intend* the others to have the associated individual intention, and, again, are aware of this. To formalize these conditions, $\text{E-INT}_G(\varphi)$ (standing for "everyone intends") is defined by $\text{E-INT}_G(\varphi) \leftrightarrow \bigwedge_{i \in G} \mathrm{INT}(i, \varphi)$. Now we repeat the axiom defining collective intentions:

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

Level 3: the *plan generation/formation* level.

This level includes *task division* and *task allocation*. The input of this stage is a team G together with its collective intention to achieve φ. During a successful run of this stage, first, an adequate task division of φ into a sequence of subtasks $\varphi_1, \ldots, \varphi_n$ is constructed, ensuring the certain realization of the collective goal. Subsequently, an appropriate subtask allocation to the team members is established with a temporal structure among the tasks. By 'appropriate allocation' we mean that not only agents' abilities and resources, but also their commitment strategies are taken into account. The result of this two-step process is a social plan P (see section 2).

In effect, all agents from a group socially commit to carry out their respective subtasks, and to communicate about these social commitments in order to establish pair-wise mutual beliefs about them. This information exchange concludes the *collective* part of plan generation. The final successful outcome of this stage is collective commitment C-COMM$_{G,P}(\varphi)$ of the group G based on the social plan P. Because of our strong notion of collective commitment, this definition of the successful outcome of the stage of plan generation is somewhat stronger than the one in [15].

Level 4: the *plan execution/team action* level.
This level includes *execution of actions* and the *reconfiguration procedure*: all team members aim at realizing their own subtasks from the collective commitment C-COMM$_{G,P}(\varphi)$. Thus they start executing adequate agent-specific actions. Many different situations may occur, some of which imply reconfiguration among the group, an aspect which is not treated explicitly in [15]. We will discuss reconfiguration in the sequel. In terms of motivational attitudes, plan realization amounts to the maintenance of *social commitments* and associated *individual intentions*. The successful outcome of this stage is that all subtasks have been carried out by the agents who were socially committed to do them and that by the success of their actions the goal φ has been achieved.

Now we can define success and failure of the system. The *system fails* if the overall goal φ is not realizable by any team from the relevant set T; the *system succeeds* otherwise.

4 The reconfiguration method

The dynamics of the system will be reflected in the *cycle of the system*, obeying the consecutive steps of: potential recognition, team formation, plan generation, and plan execution. In the perfect case a group achieves the goal in the way it was planned. In the more common non-perfect case disturbances appear in the first cycle of the system, and some kind of *reconfiguration* is necessary. For this reason, the cycle of the system will be referred to as the *reconfiguration algorithm*. The main purpose of the algorithm is *the proper maintenance of collective commitments*. In order to deal with a variety of situations the reasons of disturbances have to be recognized. Therefore, a few definitions will be introduced. An action execution fails for an *objective reason* if the action is not realizable by anybody in the present team in the current state of the world. An action execution fails for a *subjective reason* if the agent to whom it is delegated does not believe that he can achieve it.

The problem of choosing adequate properties of a reconfiguration method is viewed as an open question, possibly related to the type of application considered. When formulating the algorithm, we chose some intuitive properties corresponding to classical strategies adopted in backtracking. We postulate that the system behaviour should preserve *continuity*: when an obstacle appears, one moves to the nearest point up in the hierarchy of levels where a different choice is possible. If such a point does not exist anymore, the reconfiguration algorithm fails. In other words, depth-first search is used.

The context-independent continuity criterion determines the overall structure of the algorithm. However, as the basis of local decisions a context-sensitive criterion would be valuable. When new rounds at particular levels are necessary, the question arises which results should be preferred based on the failed ones. The answer is context-dependent and requires formulation of a specific notion of the distance between teams, goals, plans, etc. It seems that for a wide class of application domains, the system should behave in a rather *conservative* way. In our case, conservativity (or *inertia*) entails that the collective commitment that is being carried out should change as little as possible.

4.1 Reconfiguration algorithm

The (informal) description of the reconfiguration algorithm below is meant to be generic: a pattern of behavior is described in terms of abstract level-associated procedures (i.e. *potential-recognition, team-formation, task-division, task-allocation, plan-execution*), without fixing any particular method or strategy. Input and output parameters, as well as some other conditions, are commented in the algorithm itself. Each of these complex procedures may succeed or fail - in this sense the structure of each abstract level is analogical. To make the algorithm's structure more transparent, we label each level-associated procedure and use the *goto* statement.

In the algorithm, new social and individual motivational attitudes are assigned to create a new collective commitment. This may be viewed as *revision* of motivational attitudes. The phases of belief revision and motivational attitude revison are distinguished, without splitting collective motivational attitudes into individual and social ones. They are realized by abstract procedures *belief-revision* and *motivational-attitudes-revision*, after *motivational-attitudes-assignment.* The proper treatment of revision is possible, because it is assumed that agents are obliged to communicate about changes.

System failure and success are realized by the complex procedures *system-failure* and *system-success*, respectively. It is also assumed that all the required information is available at the development-time.

Reconfiguration algorithm:

```
begin
{input of the system:
Q - a goal of agent S;
S - the agent that is input to the system and that will recognize potential;
T - a set of agents from whom potential teams are selected}
A: potential-recognition (Q, S, H, T);
    {input: Q, S, T}
    if not (potential-recognition-succeeded) then
        {S does not see any potential for cooperation with respect to goal Q}
        system-failure;
        STOP
    else
    {potential recognition succeeded - output: H - a sequence of teams H =$(G_1, \ldots, G_n)$
     with the potential to realize Q}
    motivational-attitudes-assignment;
    {the attitude POTCOOP$(S, Q)$ is established}
B: team-formation (Q, S, H, G);
    {input: Q, S, H}
    if not (team-formation-succeeded) then
    {the collective intention w.r.t. Q cannot be established among any of the teams from H;
     return to the potential recognition level for S to construct a new sequence of potential
     teams for which it sees cooperation potential w.r.t. Q}
        goto A
    else
```

```
    {team formation succeeded - output G: a set of agents aiming to realize Q; Suppose G is the
      first unused $G_i$ from H for which the collective intention towards Q can be established}
    motivational-attitudes-assignment;
    {the collective intention C-INT$_G(Q)$ of G towards Q is established here}
C: task-division (Q, G, R);
    {input: Q, G}
    if not (task-division-succeeded) then
    {task division failed; return to the team-formation level in order to attempt the first unused
     $G_{i+1}$ from H to establish collective intention towards Q}
    goto B
    fi;
    {task division succeeded - output: R - a sequence of subtasks together realizing Q;
     i.e. the first part of a social plan P}
D: task-allocation (Q, G, R, P);
    {input: Q, G, R}
    if not (task-allocation-succeeded) then
    {task allocation failed; return to the task division level}
      goto C
    fi;
    {task allocation succeeded - output: a social plan P corresponding to subtask sequence R}
        motivational-attitudes-assignment;
        {the collective commitment C-COMM$_{G,P}(Q)$ among G to goal Q with respect to social
        plan P is established (including corresponding social commitments to subtasks from R)}
        {plan execution part starts here}
    E: plan-execution (Q, G, R, P);
        {input Q, G, R, P}
        if plan-execution-succeeded then
        {all actions that constitute the plan P are successfully executed; agents' beliefs and
         motivational attitudes need to be revised to reflect that the overall goal Q is achieved
         as well as agents' subgoals from R}
            belief-revision;
            motivational-attitudes-revision;
            system-success;
            STOP
        elseif
            {some action execution failed: differentiation of reasons for failure}
            F1:if subjective-reason-for-failure(Q, P, G, R, R1, R2) and
            not (objective-reason-for-failure(Q, P, G, R, R1 ,R2)) then
            {R1: the sequence of tasks from R that have been successfully achieved thus far;
             R2: the sequence of pairs (A, X) of tasks that failed plus reasons for failure,
             where X=ob or X=sub}
            {for every action that failed, it failed for a subjective reason, i.e.
            $\forall A \in R \, \forall X((A, X) \in R2 \to X = sub)$}
             belief-revision;
             if task-reallocation-possible then
             {for every action that failed for a subjective reason, there is another team
               member believing it can achieve it, i.e. $\forall A \in R \, \forall M_1, M_2 \in G((A, sub) \in R2$
               $\wedge \mathrm{COMM}(M_1, M_2, A) \to \exists M_3 \in G(M_1 \neq M_3 \wedge \mathrm{BEL}(M_3, \mathrm{CAN}(M_3, A))))$;
               before an attempt at task reallocation based on P, R1 and R2 is made,
               a revision of motivational attitudes is needed}
                   motivational-attitudes-revision;
```

```
          goto D
        elseif
        {no task reallocation is possible: a new task division is needed}
         goto C
      fi
    fi
    {there are also objective reasons for failure, i.e. $\exists A \in R((A, ob) \in R2)$}
    F2:  belief-revision;
    {now it should be investigated whether the objective reason for failure blocks achieving
     the overall goal Q}
    if blocked(Q) then
    {the objective reason for failure blocks achieving the goal Q}
      system failure;
      STOP
    else
    {the objective reason for failure does not block achieving the goal Q;
     a new task division is needed, return to the task division level}
      goto C
    fi
  fi
end
```

Let us trace the cycle of the system during a realization of the reconfiguration algorithm. When consecutive steps of reconfiguration go well, changes of individual, social and collective motivational attitudes are commented in the description of the algorithm. We focus on the points where failure of the four main stages of the process takes place.

The failure of the potential recognition level (see label *A*) meaning that agent S does not see any potential for cooperation with respect to the goal Q, leads to failure of the system. The failure of the team formation level (see label *B*), meaning that the collective intention with respect to Q cannot be established among any of the teams from H, requires a return to the potential recognition stage to construct a new sequence of potential teams. The failure of the task division level (see label *C*) requires a return to the team formation level in order to establish a collective intention in the chosen new team from H and may be viewed as the reconfiguration of the team together with a change of collective intention and respective individual attitudes. The failure of the task allocation level (see label *D*) requires a return to the task division level in order to create a new sequence of tasks. This may be viewed as the reconfiguration of the first part of a social plan P to realize the goal Q. When, finally, a collective commitment is established, the failure of some actions together constituting the social plan P requires the reconfiguration of motivational attitudes on the level of collective commitment.

4.2 Complexity of the reconfiguration algorithm

As the reconfiguration algorithm is abstract, it needs to be instantiated for each particular application. When defining the levels we abstracted from particular methods and algorithms meant to realize level-associated goals, but instead formulated their final results and associated them with appropriate motivational attitudes. Let us stress that the procedures realizing level-associated goals are rather complex.

Especially the plan generation level, including task division and task allocation, is complex from both the AI and the MAS perspective. As for the first perspective, planning is involved. If done from first principles, planning is in general undecidable [3], while for limited domains it may still be tractable. If, on the other hand, a plan library is used, searching the library may be complex as well. As for the MAS perspective, the aspects of communication, negotiation and coordination come to the fore on the level of plan generation. It is quite clear that the procedures being proposed in recent work on negotiation are usually rather complex [13]. The same holds for coordination.

On the level of plan execution, belief revision, which is known to be NP-hard, is repeatedly performed. Also, if there are objectively failed tasks, it needs to be checked whether their failure blocks the overall goal. This kind of consistency check is co-NP complete. Let us stress that all the complex aspects mentioned here have to be treated in any rigorous approach to the reconfiguration problem.

As for the global structure of the reconfiguration algorithm, it is based on backtracking search. In the generic case, when not using any domain-dependent information, context-dependent improvements as obtained in informed search methods cannot be made. However we did use forward checking in the algorithm in the check whether an objectively failed task blocks the overall goal. Next, in order to be certain that the search method is complete and optimal, iterative deepening may be used. It doesn't increase the time or space complexity when compared with depth-first search. Due to its exponential space complexity, breadth-first search is no option for reconfiguration.

Let us stress that our definition of collective commitment ensures efficiency of reconfiguration in two ways. Firstly, all the motivational attitudes occurring in the definition of collective commitment are defined in a non-recursive way. This allows straightforward revision of all relevant motivational attitudes. Secondly, only pair-wise social commitments to subtasks appear in the definition of collective commitment, so that replanning and motivational attitudes revision are made less complex. Thus, in contrast to [15], using our definition of collective commitment it is often sufficient to revise some of the pair-wise commitments, instead of involving the entire team in the replanning process.

5 Example of distributed problem solving

In this section, an example of distributed problem solving is shown in which the reconfiguration algorithm is implemented profitably. The overall goal of the system is to prove a new mathematical theorem. In this application domain, we decided to use the conservativity assumption. The description of the example will not be too precise, e.g. we will not give formal rules governing mathematical provability. We will follow the cycle of the system as given in the reconfiguration algorithm.

5.1 Potential recognition

We assume that all agents work in a distributed environment, communicating, coordinating and negotiating when necessary. Let the system start with an open-minded agent t who looks for cooperation potential for the overall goal $prove(theorem(T))$, from among the set of agents $\{t, l, l_2, c\}$ where both l and l_2 are single-minded agents of whom t believes that they are able to prove the necessary lemmas, and c is a blindly committed agent of whom t believes that it can check all the proofs that will be constructed. Moreover, t cannot prove T on his own, but there is a sequence of potential teams of which t believes that each team can jointly achieve the goal, namely $H =< \{t,l,c\},\{t,l_2,c\},\{t,l,l_2,c\} >$. Thus POTCOOP($t$, *prove*(*theorem*($T$)) is established.

When recognizing the potential for cooperation, we postulate to take agents' commitment characteristic into account. As discussed in [5], blindly committed agents are not at all adaptive. Because in research, plan changes are necessary, it is crucial that the percentage of blindly committed agents in the team is not too high, and that enough open-minded agents are present to pitch in if necessary. Thus, in H above, all three teams contain only one blindly committed member and at least one open-minded one.

5.2 Team formation

During team formation, a team should be formed to adopt the collective intention of proving theorem T. After the potential recognition phase, agent t first tries to get the collective intention established in the first team from his potential team sequence H, namely $G = \{t,l,c\}$. In this case, let's suppose that t succeeds and C-INT$_G$(*prove*(*theorem*(T)) is established. (Otherwise t has to successively try to establish the collective intention in the other two teams.)

5.3 Plan generation

During *task division* a sequence of lemmas is determined from which the theorem T should be proved. More formally, given the theorem T to be proved, there are different divisions $L_0, L_1, \ldots$ of the theorem into sequences of lemmas, where $L_i = (l^i_1, \ldots, l^i_{d(i)})$.

The next step is *task allocation*: specific subtasks are to be allocated to adequate agents having resources and skills sufficient to realize the allocated goal, as well as suitable commitment strategies. In this case, the open-minded agent t (theorem prover) is allocated to prove the theorem T from sequences of lemmas, the single-minded (lemma prover) l is allocated to prove lemmas, and the blindly committed c (proof checker) is allocated to check the others' proofs. Below follows a sample task division and allocation, corresponding to the first division $L_0 = (l^0_1, \ldots, l^0_n)$ of the theorem into lemmas.

- *prove*(*theorem*(T)): to prove the theorem (by the whole group);
- *prove*(*lemma*(l^0_i)): to prove the specific lemmas $l^0_1, \ldots, l^0_n$ (by lemma prover l);
- *check*(*proof*(l^0_i)): to check all specific proofs of the corresponding lemma l^0_i (by proof checker c);
- *prove*(*theorem-from*(L_0)): to prove the theorem from the sequence of lemmas $L_0 =< l^0_1, \ldots, l^0_n >$ (by theorem prover t);
- *check*(*proof-from*(L_0)): to check all proofs of the theorem from the sequence of lemmas L_0 (by proof checker c).

Next, we are ready to give the succesfull condition for a proof of the full theorem:

> There is a division of the theorem T into lemmas such that for all lemmas in the sequence there exists a proof, constructed by the lemma prover, that has been found correct by the proof checker. Also there is a proof of the theorem from the sequence of lemmas, constructed by the theorem prover, which has been found correct by the proof checker.

This condition forms the basis for a simple sequential social plan to prove T:

$$P =<< \mathit{prove}(\mathit{lemma}(l_1^0)), l >; < \mathit{check}(\mathit{proof}(l_1^0)), c >; \dots;$$
$$< \mathit{prove}(\mathit{lemma}(l_n^0)), l >, < \mathit{check}(\mathit{proof}(l_n^0)), c >;$$
$$< \mathit{prove}(\mathit{theorem\text{-}from}(L_0)), t >; < \mathit{check}(\mathit{proof\text{-}from}(L_0)), c >>.$$

In practice, a different temporal structure may be more efficient, e.g. subtasks could be carried out in parallel. Based on the plan P, the group $G = \{t, l, c\}$ is prepared to establish a *collective commitment*:

$$\begin{aligned}
&\text{C-COMM}_{G,P}(\mathit{prove}(\mathit{theorem}(T))) \leftrightarrow \text{C-INT}_G(\mathit{prove}(\mathit{theorem}(T)))\wedge\\
&\quad\bigwedge_{i=1}^{n}(\text{COMM}(l, t, \mathit{prove}(\mathit{lemma}(l_i^0))) \wedge \text{COMM}(c, l, \mathit{check}(\mathit{proof}(l_i^0)))) \wedge\\
&\quad\text{COMM}(t, l, \mathit{prove}(\mathit{theorem\text{-}from}(L_0))) \wedge \text{COMM}(c, t, \mathit{check}(\mathit{proof\text{-}from}(L_0))) \wedge\\
&\quad\text{C-BEL}_G(\bigwedge_{\alpha\in P}\ \bigvee_{a,b\in G} \text{COMM}(a, b, \alpha))
\end{aligned}$$

In the sense of goals to be reached, no more social commitments are needed, although communication (e.g. scientific discussion) about these goals is always possible. The common beliefs are established by bilateral communication between the agents involved.

5.4 Plan execution

The actions to be carried out by the agents in G correspond directly to subtasks in the plan P and reflected in $\text{C-COMM}_{G,P}(\mathit{prove}(\mathit{theorem}(T)))$. During plan execution, the degree of responsibility and thus complexity of the three agents is quite different. The blind agent can work on his delegated tasks without tracking what happens with the other agents. However, the single-minded and especially the open-minded agent have to react when the agent(s) working for them fail to bring their tasks to a good end.

Below we sketch the reconfiguration procedure for the example, corresponding to section 4. As in the general case, when a particular action fails it is important to recognize that each reason gives rise to a different adequate reaction satisfying continuity. We also choose to make local decisions obeying conservativity. Usually the agent for whom the action is executed reacts. Suppose that in the current round, the team is proving the theorem T from lemma sequence $L_j =< l_1^j, \dots, l_{d(j)}^j >$.

The reaction to failure can be subdivided into two cases, first according to different *subjective reasons* for failure of actions: part F1 of the reconfiguration algorithm. Actions that fail for subjective reasons happen when:

sub 1 The proof checker c doesn't finish $\mathit{check}(\mathit{proof}(l_i^j))$ before a pre-set time limit. In this case, the most efficient and conservative way of task reallocation is to let the single-minded lemma prover pitch in for c by checking his own proof, and keep the rest of the plan unchanged. Thus, belief revision in group G is needed, then a new individual intention of l to $\mathit{check}(\mathit{proof}(l_i^j))$ is added, as well as a social commitment toward c to do this, which is reflected in a new collective commitment.

sub 2 The lemma prover l does not believe he can carry out $\mathit{prove}(\mathit{lemma}(l_i^j)))$, but without knowing a counter-example. The most conservative way of task reallocation is to let t pitch in for l if t believes that he can prove the lemma. After the belief revision in group G, agent l drops his social commitment towards t and his individual intention to make the proof, which are both taken over by t towards him and reflected in a new collective commitment.

sub 3 The theorem prover t does not believe that he can do $\mathit{prove}(\mathit{theorem\text{-}from}(L_j)))$, but without knowing a counter-example. The most conservative way of task reallocation is to let l pitch in for t if l believes that he can prove the theorem. After the belief revision in group G, agent t drops his social commitment towards l and his individual intention to make the proof, which are both taken over by l towards him and reflected in a new collective commitment.

In part F2 of the reconfiguration algorithm, objective reasons of failure are considered. First, following the reconfiguration algorithm, it needs to be checked whether the overall goal, namely $\mathit{prove}(\mathit{theorem}(T))$, is blocked. In the example, this happens only in case a counterexample to theorem T is found by c, in which case the system fails.

There are also two cases which do not block the overall goal; in both cases a new task division has to be made. Here follows an elaboration of both cases.

ob 1 The proof checker finds a mistake in a proof; then the action of making the proof has failed, but may be repeated by constructing a new proof of the same lemma or of the theorem. In this case, to obey conservativity, almost the same division of lemmas and the same task division and allocation are proposed. The only change is that the maker of the faulty proof gets as new task to make a new proof, which is then reflected in the social plan and in a social commitment such as $\mathrm{COMM}(l, t, \mathit{prove\text{-}lemma\text{-}diff}(l_i^j, p))$ (l commits to t to make a new proof of lemma l_i^j, different from the faulty proof p), and finally in a new collective commitment.

ob 2 The proof checker finds a counterexample to a lemma; then the action of making the proof has failed, and it has no sense to try again. At the new round of task division, the following happens. In case a counterexample to lemma l_i^j in the sequence of lemmas L_j was found, a new division of the theorem into a lemma sequence L_{j+1} has to be made. To obey conservativity, the new sequence should resemble L_j as closely as possible. Especially important is to conserve as many as possible lemmas from the initial sequence $l_1^j, \ldots, l_{i-1}^j$ which have already been proved and checked. Thus a new sequence of tasks is made corresponding to L_{j+1}.

During task allocation, the task of proving the lemmas is again allocated to l, and the task of proving the theorem from lemma sequence L_{j+1} is allocated to t. Agent c is again allocated to check all proofs. The resulting social plan is:

$$Q = << \mathit{prove}(\mathit{lemma}(l_1^{j+1})), l >; < \mathit{check}(\mathit{proof}(l_1^{j+1})), c >; \ldots;$$
$$< \mathit{prove}(\mathit{lemma}(l_{d(j+1)}^{j+1})), l >, < \mathit{check}(\mathit{proof}(l_{d(j+1)}^{j+1})), c >;$$
$$< \mathit{prove}(\mathit{theorem\text{-}from}(L_{j+1})), t >; < \mathit{check}(\mathit{proof\text{-}from}(L_{j+1})), c >>.$$

This plan is then reflected in a new collective commitment constructed from the previous one (see Subsection 5.3). Namely, both l and t revise their individual intentions and social commitments towards the lemmas that are not in the new sequence L_{j+1}, whereas c adds individual intentions and social commitments to l and t with respect to their new proofs.

After each reconfiguration step, the agents continue executing their actions until a new block appears or the goal of proving the theorem is reached.

6 Discussion and conclusions

In this paper we have provided an abstract multi-level architecture which makes it possible to properly deal with DPS, based on the BDI paradigm. In particular, we stressed the need of the proper treatment of all motivational attitudes — individual, social, and collective, through all stages of DPS. To achieve this we introduced, as in [15], levels of abstraction reflecting the main aspects of DPS, namely potential recognition, team formation, plan generation, and plan execution. The main contribution is the reconfiguration algorithm respecting the property of continuity, which entails that replanning should be more efficient than when other (e.g. blind) methods are used. Also our definition of collective commitment is meant to ensure efficient replanning because of the pair-wise social commitments involved. In addition, it seems that in many domains, conservativity is the appropriate criterion to guide local choices during reconfiguration.

[9] consider related issues, but do not provide a generic reconfiguration method and work with pre-given plans instead of plans that can be adapted on the fly. Also, their definition of collective intention is weaker than ours. In [8] an interesting definition of partial shared plans is given, but reconfiguration is not treated as a separate subject. The need for reconfiguration was recognized by other authors, and some of the above-mentioned papers treat certain aspects of reconfiguration or apply it to specific cases. Our paper, in contrast, provides a rather methodological approach maintaining all motivational attitudes.

When defining a general-purpose reconfiguration method some problems have been left aside. We do not take into account questions concerning heterogeneity of the system (see [14]). Our agents are allowed to be dissimilar and to have different problem solving perspectives, reflected in local expertise of domain solvers and global / strategic expertise of DPS organizers. However, all of them need to communicate, and some to cooperate and coordinate, despite their differences. We agree with [14] that agents' diversity should be restricted by possessing common ontologies, in this case with respect to action effects and motivational attitudes.

Future work will be to formalize the framework and to prove desired properties like efficiency of replanning. Next, the reconfiguration algorithm may be implemented for some applications.

Acknowledgments

We would like to thank professor Gerard Renardel de Lavalette for his helpful comments. This work is supported by the grant ESPRIT-CRIT2 No 20288.

References

[1] M. Bratman. *Intension, Plans, and Practical Reason.* Harvard University Press, Cambridge (MA), 1987.

[2] C. Castelfranchi. Commitments: From individual intentions to groups and organizations. In V. Lesser, editor, *Proceedings First International Conference on Multi-Agent Systems*, pages 41–48, San Francisco, 1995. AAAI-Press and MIT Press.

[3] D. Chapman. Planning for conjunctive goals. *Artificial Intelligence*, 32:333–377, 1987.

[4] F. Dignum and R. Conte. Intentional agents and goal formation: Extended abstract. In M.P. Singh, A.S. Rao, and M.J. Wooldridge, editors, *Preproceedings Fourth International Workshop on Agent Theories, Architectures and Languages*, pages 219–231, Providence, Rhode Island, 1997.

[5] 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), 1996. AAAI-Press.

[6] B. Dunin-Kęplicz and R. Verbrugge. A methodology for maintaining collective motivational attitudes during teamwork. In F. Garijo and Ch. Lemaitre, editors, *Multi Agent Systems Models Architectures and Applications: Proceedings of the Second Iberoamerican Workshop on DAI and MAS, October 1-2 1998*, pages 45–60, Toledo, Spain, 1998.

[7] A.E. Emerson. Temporal and modal logic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 995–1072. Elsevier and MIT-Press, Amsterdam and Cambridge (MA), 1990.

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

[9] D. Kinny, M. Ljungberg, A.S. Rao, E.A. Sonenberg, G. Tidhar, and E. Werner. Planned team activity. In *Artificial Social Systems*, volume 890 of *LNAI*, Berlin, 1992. Springer-Verlag.

[10] 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.

[11] 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, 1992. Elsevier.

[12] A.S. Rao and M.P. Georgeff. Formal models and decision procedures for multi-agent systems. Technical Report Technical Note 61, Australian Artificial Intelligence Institute, Carlton, Victoria, 1995.

[13] J.S. Rosenschein and G. Zlotkin. *Rules of Encounter: Designing Conventions for Automated Negotiation Among Computers.* MIT-Press, Cambridge (MA), 1994.

[14] R. Weihmayer, R. Brandau, and Hong Shinn. Modes of diversity: Issues in cooperation among dissimilar agents. In *Proceedings 10th International Workshop on Distributed AI (AAAI/MCC)*, Bandera, Texas, October 1990.

[15] M. Wooldridge and N.R. Jennings. Towards a theory of collective problem solving. In J.W. Perram and J.P. Muller, editors, *Distributed Software Agents and Applications*, volume 1069 of *Lecture Notes in Artificial Intelligence*, pages 40–53. Springer Verlag, Berlin, 1996.