Argument-based Negotiation among BDI Agents

Sonia V. Rueda Alejandro J. García Guillermo R. Simari
svr@cs.uns.edu.ar ajg@cs.uns.edu.ar grs@cs.uns.edu.ar
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
8000 Bahía Blanca, Argentina

Abstract

Negotiation is a basic mechanism for interaction that allows the members in a Multiagent System to coordinate their actions and to reach a favorable agreement. When agents are collaborative, the negotiation process progresses through a dialogue in which proposals and counter-proposals are exchanged in a common effort to advance towards a mutual agreement.

An *Interaction Protocol* regulates communication and gives structure to the dialog. Most interaction protocols designed to regulate negotiation among agents are abstract models based in some real world negotiation practice (*e.g.* auctions). Here we propose a deliberative mechanism for negotiation among *BDI* agents based in Argumentation.

Keywords: Multiagent Systems, Autonomous Agents, Argumentation, Negotiation, BDI Model, Collaborative Interaction.

1 Introduction

Negotiation is a basic mechanism for interaction that allows the members in a Multiagent System to coordinate their actions and to reach a favorable agreement. When agents are collaborative, the negotiation process progresses through a dialogue in which proposals and counter-proposals are exchanged in a common effort to advance towards a mutual agreement.

An *Interaction Protocol* regulates communication and gives structure to the dialog. Most interaction protocols designed to regulate negotiation among agents are abstract models based in some real world negotiation practice (*e.g.* auctions). Here we propose a deliberative mechanism for negotiation among *BDI* agents based on Argumentation.

In a BDI agent, *mental attitudes* are used to model its cognitive capabilities. These mental attitudes include ***B****eliefs*, ***D****esires* and ***I****ntentions* among others such as *preferences*, *obligations*, *commitments*, *et cetera*. These attitudes represent motivations of the agent and its informational and deliberative states which are used to determine its behaviour.

In this paper, the mental attitude of an agent will change as a consequence of its actions and the interaction in which it will engage with other agents. Each member in that group will develop plans trying to reach its committed goals. Some of those goals are proposals, or requests of collaboration, coming from other members of the group.

Agents will use a formalism based in argumentation in order to obtain plans for their goals represented by literals. They will begin by trying to construct a warrant for the goal. That might not be possible because some needed literals are not available. The agent will try to obtain those missing literals, regarded as subgoals, by executing the actions it has available. When no action can achieve the subgoals the agent will request collaboration. In making that request, given the social obligation of being collaborative, the agent will change the intentions of the receiver. The agent that receives the proposal uses its beliefs and abilities to respond to the request and, when the negotiation is successful, with an agreement. In that case, it is the mental attitude of the agent that started the negotiation that gets modified by the addition of the beliefs needed to achieve its goals.

When the receiver cannot collaborate, it should inform whether the reason is that lacks the capabilities to do it or the reason is that there exists a conflict with its own plans or intentions. In the latter case, the agent which made the original request will try to change its plans. If that is not possible, it will engage in dialog again and it will *insist* in its request. Thus, the agent which receives the request will try to revise its own plans and intentions in an attempt to collaborate.

In our model, agents cannot use their abilities in order to change the intentions of other member of the group. Nevertheless, they can achive that through a dialog. In that sense, every illocution can be thought as an special kind of action that can modify the mental attitude of the receiver. Therefore, the rol of language is fundamental. In what follows, we propose a set of primitives that support argumentation-based negotiation among BDI agents.

The paper is organized as follows. Next, in section 2 and 3, we will briefly describe BDI agents and Planning with Defeasible Argumentation. In section 4 we define a collaborative Multiagent System and we introduce some illustrative examples of the basic interaction model. Finally, in section 5 we propose an interaction protocol and in section 6 we offer some conclusions and future work directions.

2 The Construction of a BDI Agent's Plan

An agent is a computational entity which is capable of perceiving, reacting, and acting upon an environment. A rational agent adds a certain level of *practical reasoning* to its perceptive, reactive, and effectoric capabilities. This sort of reasoning [16] allows the agent to select the best possible action based on knowledge about the goal and the state of the environment it is in. That is to say, the agent's cognitive capability is what will allow it to behave adequately in each particular context.

Practical reasoning involves two fundamental processes: decide **what** goals are going to be pursued, and choose a plan on **how** to achieve them. The decision process requires the consideration of all the alternative options, the selection of some of these options, and the commitment to them. The selected options will make up the agent's *intentions*; they will also have an influence on its actions, restrict future practical reasoning, and persist (in some way) in time. The planning process consists of selecting a set of actions that will allow to satisfy the intentions.

An agent will modify its intentions when it decides that it will not be able to achieve them, or the reasons that supported their selection among other alternatives are no longer valid. Thus, an agent's intentions are linked to its *beliefs* about the world that surrounds it.

The belief, desire, and intentions model (BDI), based on practical reasoning, possesses the necessary elements for the representation of the mental attitude of a rational agent that acts in a dynamic environment, subject to sudden and frequent changes. Beliefs make up the agent's knowledge about the environment that it occupies. Its desires and intentions refer to the state that it wishes to achieve, and represent its motivations and commitments.

A BDI agent elaborates a plan considering its beliefs about the world, its intentions with respect to the desired state, and its abilities to transform it. Abilities are associated with actions that have *preconditions* and *consequences*. An agent can consider an action only if its beliefs tell it that the preconditions are satisfied. The application of an action creates consequences that may modify the set of beliefs, creating a dynamic and continuous interaction. A plan is a sequence of actions that, upon execution, provoke changes in the knowledge about the environment.

The individual agent's plan is built considering only its beliefs, intentions, and abilities. When an agent is part of a community, it can ask for collaboration and make use of the beliefs and abilities of the members of the system in which it participates.

A social agent's beliefs are built from its perception of the world that surrounds it, but they also include its knowledge about the knowledge of other agents in the group. Some cognitive elements will be shared by all the members of the system. Each individual will know these elements, but will also reason knowing that the others also know them, and know that they are shared.

Each agent's knowledge will then gather its own beliefs and the beliefs shared with other members of the group. This set of beliefs is consistent; nevertheless, the same may not be applicable among the sets of different agent's proper beliefs. If an agent's actions are going to modify the shared beliefs, all of the members of the group must agree.

3 Planning and Argumentation

As mentioned above, a rational agent has certain level of *practical reasoning*. This kind of reasoning allows the agent to select actions to interact with its environment. In this section we will introduce the necessary definitions of an argumentation-based formalism to construct plans. For a more detailed account see [1].

Definition 3.1 [Argumentative BDI agent]
An argumentative BDI agent **a** will be denoted with the tuple $\mathbf{a} = \langle \mathcal{B}, \mathcal{D}, \mathcal{I}, \Gamma \rangle$, where $\mathcal{B}$ represents the agent beliefs, $\mathcal{D}$ its desires, $\mathcal{I}$ the agent intentions and Γ a set of actions that the agent is able to execute. □

The agents desires $\mathcal{D}$ will be represented by a set of literals that will also be called *goals*. A subset of $\mathcal{D}$ will represent a set of *committed goals* and will be referred to as the agent intentention, denoted $\mathcal{I}$.

The agent's beliefs $\mathcal{B}$ will be represented by a a restricted *Defeasible Logic Program* (Φ, Δ), where Φ will be a consistent set of facts, and Δ a set of defeasible rules. A brief description of Defeasible Logic Programming (DeLP) follows. For further details of DeLP see [2], and for its restricted version see [3].

In DeLP, a literal h is *warranted* if there exists a non-defeated *argument* $\mathcal{A}$ supporting h. An argument structure $\mathcal{A}$ for a literal h (denoted $\langle \mathcal{A}, h \rangle$) is a minimal and consistent set of defeasible rules that allows to infer h. In order to establish whether $\langle \mathcal{A}, h \rangle$ is a non-defeated argument, *argument rebuttals* or *counter-arguments* that could be *defeaters* for $\langle \mathcal{A}, h \rangle$ are considered, *i.e.*, counter-arguments that by some criterion, are preferred to $\langle \mathcal{A}, h \rangle$. Since counter-arguments are arguments, there may exist defeaters for them, and defeaters for these defeaters, and so on. Thus, a sequence of arguments *called argumentation line* may appear, where each argument defeats its predecessor in the line. Usually, each argument has more than one defeater and more than one argumentation line exists. Therefore, a tree of arguments called *dialectical tree* is constructed, where the root is $\langle \mathcal{A}, h \rangle$ and each path from the root to a leaf is an argumentation line. A *dialectical analysis* of this tree is used for deciding whether h is warranted.

Besides its beliefs, desires and intentions, an agent will have a set of actions Γ that it may use to change its world. The formal definitions that were introduced in [1] are recalled below.

Definition 3.2 [**Action**] An action A is an ordered triple $\langle \mathsf{P}, \mathsf{X}, \mathsf{C} \rangle$, where P is a set of literals representing preconditions for A, X is a consistent set of literals representing consequences of executing A, and C is a set of constraints of the form *not* L, where L is a literal. We will denote actions as follows:

$$\{X_1, \ldots, X_n\} \xleftarrow{A} \{P_1, \ldots, P_m\}, not\ \{C_1, \ldots, C_k\}$$

Notice that the notation *not* $\{C_1, \ldots, C_k\}$ represents $\{not\ C_1, \ldots, not\ C_k\}$. □

Definition 3.3 **[Applicable Action]** Let $\mathcal{B} = (\Phi, \Delta)$ be an agent's beliefs. Let Γ be the set of actions available to that agent. An action A in Γ, defined as before, is applicable if every precondition P_i in P has a warrant built from (Φ, Δ) and every constraint C_i in C fails to be warranted from (Φ, Δ). $\Box$

Definition 3.4 **[Action Effect]** Let $\mathcal{B} = (\Phi, \Delta)$ be an agent's beliefs. Let Γ be the set of actions available to that agent. Let A be an applicable action in Γ defined by:

$$\{X_1, \ldots, X_n\} \xleftarrow{A} \{P_1, \ldots, P_m\}, not\ \{C_1, \ldots, C_k\}$$

The effect of executing A is the revision of Φ by X, i.e. $\Phi^{*\mathsf{X}} = \Phi^{*\{X_1,\ldots,X_n\}}$.

Revision will consist of removing any literal in Φ that is complementary of any literal in X and then adding X to the resulting set. Formally:

$$\Phi^{*\mathsf{X}} = \Phi^{*\{X_1,\ldots,X_n\}} = (\Phi - \overline{\mathsf{X}}) \cup \mathsf{X}$$

where $\overline{\mathsf{X}}$ represents the set of complements of members of X. $\Box$

It has been shown in [1] that the interaction between actions and the defeasible argumentation formalism is twofold. On one hand, as stated by Definition 3.3, defeasible argumentation is used for testing preconditions and constraints through the construction of the necessary warrants. On the other hand, actions may be used by agents in order to change the world and then obtain a warrant for a literal h that is not warranted from its current beliefs (Φ, Δ).

When an agent $\mathbf{a}$ selects an intention g from $\mathcal{I}$, and g is not warranted from (Φ, Δ), agent $\mathbf{a}$ may look for an action A in Γ that modifies the set of beliefs in such a way that an argument without defeaters can be obtained to support g. If A cannot be executed because its preconditions are not warranted, $\mathbf{a}$ can look for a sequence of actions that will allow it to establish them. The complete sequence, including A, will make up a plan for g. Each action can modify the set of beliefs adding literals that will allow the construction of new arguments.

In the construction of a plan actions are chosen so that, when executed, will allow to introduce literals necessary to obtain warrants. Warrants are also obtained to support preconditions that will allow the execution of actions. Action selection is not a trivial task, and the classic planning problems are reflected in the argumentative schema. In a sequence of actions $[A_1,A_2]$, A_2 may be applicable according to the initial state Φ, but not after A_1 is executed. Therefore, the consequences of A_1 may modify Φ in such a way that the preconditions of A_2 are not warranted. The execution of A_1 may add literals that allow the construction of new defeaters for the preconditions of A_2 and eliminate literals that allow the construction of warrants for the preconditions of A_2. An analysis on Progression and Regression Planning in the context of this formalism can be found in [1].

When planning is performed as a collaborative process, an agent which forms part of a group could select an action that may interfere with the rest of the members' plans. As its behavior should be collaborative, it will ask for permission to introduce changes that may affect others. Next, we will explore some of the situations that appear during the interaction of this type of agents.

4 Collaborative Agents

The behavior of an isolated agent is determined by its individual motivations, its beliefs about the world, and its own abilities. This characterization is insufficient for the modeling of an agent that interacts in a social context with a cooperative attitude.

A social agent's beliefs are built from the agent's perception of the environment and there will be some shared cognitive elements among all the members of the system. Each individual will know these elements, and will reason knowing that the rest also knows them, and know that they are shared [5].

Definition 4.1 **[Collaborative MAS]**
A collaborative Multi-Agent System will be a pair $[\Phi_s, \mathbf{A}]$, where $\mathbf{A}$ is a set of argumentative BDI agents, and Φ_s is the set of shared beliefs. $\Box$

The knowledge of each member of the system is incomplete and frequently is insufficient for the deduction of a specific fact from it. This fact could be deduced by gathering all the distributed knowledge in the group. The group is heterogenous, and therefore its members have different abilities. Each of them builds plans from its own sets of actions. When it asks for collaboration, another member of the group may execute actions to satisfy the requirement.

In this work, we will reduce the cardinality of set $\mathbf{A}$ to two agents. This simplification will allow us to concentrate on the essential problems of interaction. It is clear that the introduction of more members to the group will bring new complications, and these will be the subject of future work.

Example 4.1 Let $[\Phi_s, \mathbf{A}]$ be a group integrated by two collaborative agents, $\mathbf{a}_1 = \langle \mathcal{B}_1, \mathcal{D}_1, \mathcal{I}_1, \Gamma_1 \rangle$ and $\mathbf{a}_2 = \langle \mathcal{B}_2, \mathcal{D}_2, \mathcal{I}_2, \Gamma_2 \rangle$ defined as:
$\Phi_s = \{a, b\}$
$\mathcal{B}_1 = (\Phi_1, \Delta_1)$ where $\Phi_1 = \{c, \sim d, e\}$ and
$\Delta_1 = \{(p \prec c, h, q), (q \prec b, \sim d), (r \prec c, s), (s \prec f)\}$
$\Gamma_1 = \{\{f\} \xleftarrow{A_1} \{a, e\}, not\ \{\}\}$
$\mathcal{B}_2 = (\Phi_2, \Delta_2)$ where $\Phi_2 = \{h, \sim i\}$ and
$\Delta_2 = \{(u \prec a, v), (v \prec b, j), (w \prec h)\}$
$\Gamma_2 = \{\{j, \sim a\} \xleftarrow{A_2} \{b, \sim i\}, not\ \{\}\}$
Suppose that $\mathcal{D}_1 = \{p, r\}$ is $\mathbf{a}_1$'s set of goals and $\mathcal{I}_1 = \{p\}$ is the set of committed goals. Agent $\mathbf{a}_1$ is not capable of constructing a warrant for p using its own beliefs and the shared beliefs, but it could do it if it gathered all of its beliefs with those of $\mathbf{a}_2$. $\Box$

The aggregation of the whole group's beliefs in one set would invalidate the advantages of working with a community of agents. The whole model could in that case colapse to just one agent. A more interesting alternative, which also respects the paradigm, is that agents interact asking for cooperation when the set of beliefs they can use is insufficient.

The group may be heterogenous and all the members of the group can benefit if they cooperate. Agents have a collaborative attitude; they offer their beliefs and abilities to the rest of the group and they can, on the other hand, request the knowledge necessary to carry out an action to build a warrant by themselves.

Social contact requires some capability for resolving conflicts that may be introduced by the interaction. In example 4.1, if $\mathbf{a}_2$ has the intention of constructing a warrant for v, it may decide to execute A_2, and j will become a fact and v will be warranted. A_2's side-effect will be to change the shared beliefs, removing a.

$\Phi'_s = \{\sim a, b\}$ $\quad \Phi'_2 = \{h, \sim i, j\}$

The decision of $\mathbf{a}_2$ may be in conflict with $\mathbf{a}_1$'s plan in the case that this agent has the intention of executing A_1 to build a warrant that will support r.

Agents are autonomous, and this quality is reflected in the fact that none of them will be able to execute actions that can directly modify other agents' beliefs. On the other hand, all of the members of the group can modify the shared beliefs. Moreover, each agent maintains consistency between its own beliefs and the shared ones, but it cannot control consistency between other agents' beliefs and the shared ones. Social interaction requires that each agent communicate with the rest before executing an action that may affect the group's beliefs.

4.1 Negotiating Beliefs

An agent $\mathbf{a}_1$ can ask for collaboration when it has committed to an intention g of $\mathcal{I}_1$ and, to build an argument to support it, it requires a literal l. Agent $\mathbf{a}_1$ cannot find a warrant for l using belief sets $\mathcal{B}_1$ and Φ_s nor can it elaborate a sequence of actions that allow it to build a warrant. Therefore, it asks for assistance from another member of the group. An agent $\mathbf{a}_2$ then receives a *proposal* from $\mathbf{a}_1$ with respect to l, and the literal l is added to its set of intentions.

The following example will allow us to illustrate different situations, shown below as different cases produced by varying the set of committed intentions.

Example 4.2 Let $[\Phi_s, \mathbf{A}]$ be a group composed of two collaborative agents, $\mathbf{a}_1 = \langle \mathcal{B}_1, \mathcal{D}_1, \mathcal{I}_1, \Gamma_1 \rangle$ and $\mathbf{a}_2 = \langle \mathcal{B}_2, \mathcal{D}_2, \mathcal{I}_2, \Gamma_2 \rangle$ where:

$\Phi_s = \{a, b\}$

$\mathcal{B}_1 = (\Phi_1, \Delta_1)$ where $\Phi_1 = \{c, \sim d, e\}$ and

$\Delta_1 = \{(p \prec a, h), (q \prec c, f), (r \prec b, \sim d), (s \prec e, \sim d, g), (t \prec a, k)\}$

$\Gamma_1 = \{(\{f\} \xleftarrow{A_1} \{e, r, h\}, \textit{not}\ \{\}), (\{g\} \xleftarrow{A_2} \{\sim d, j\}, \textit{not}\ \{\})\}$

$\mathcal{B}_2 = (\Phi_2, \Delta_2)$ where $\Phi_2 = \{h, \sim i\}$ and

$\Delta_2 = \{(u \prec a, v), (v \prec b), (\sim v \prec w), (w \prec h), (\sim w \prec \sim i), (x \prec h, \sim i, j)\}$

$\Gamma_2 = \{(\{j\} \xleftarrow{A_3} \{a, h, u\}, \textit{not}\ \{\}), (\{k\} \xleftarrow{A_4} \{b, x\}, \textit{not}\ \{\})\}$

□

In the simplest case, $\mathbf{a}_2$ *knows* l and the request can be taken directly adding l to the set of committed beliefs and notifying $\mathbf{a}_1$. More complex cases are introduced below.

Case 1 below shows that when the group is conformed by only two agents, and one of them can satisfy a proposal without executing any actions, the change in Φ_s will not introduce any conflicts, because the only literal that is being added is exactly the one that was required. If the group were larger, a conflict may arise with the beliefs of the rest of the members.

Case 1 Let $\mathcal{D}_1 = \{p, q, s, t\}$ be $\mathbf{a}_1$'s set of goals, and $\mathcal{I}_1 = \{p\}$ the set of committed goals. In order to build a warrant for p, $\mathbf{a}_1$ needs h and makes a request to $\mathbf{a}_2$. The literal is in the beliefs of $\mathbf{a}_2$, so the response is immediate. and the beliefs are modified accordingly:

$\Phi'_s = \{a, b, h\}$ $\quad \Phi'_2 = \{\sim i\}$

Now, $\mathbf{a}_1$ can build an argument $\mathcal{A}^1\mathbf{a}_1 = \{ p \prec a, h \}$ that supports p, since $\mathcal{A}^1\mathbf{a}_1$ has no counterarguments that defeats it, therefore p is warranted. □

The following case shows a different situation where an agent may request a literal because that literal is part of the preconditions of an action that agent needs to execute.

Case 2 Let $\mathcal{D}_1 = \{p, q, s, t\}$ be $\mathbf{a}_1$'s set of goals and $\mathcal{I}_1 = \{q\}$ the set of committed goals. In order to build a warrant for q, $\mathbf{a}_1$ needs f and its reasoning indicates that if it were to execute A_1 the literal f would be added to Φ_1. Nevertheless, one of A_1's preconditions is not satisfied. The literal e belongs to Φ_1, the argument $\mathcal{A}^1\mathbf{a}_1 = \{r \prec b, \sim d\}$ supports r, but its not possible for $\mathbf{a}_1$ to build a warrant for the precondition h. Therefore, it makes a proposal to $\mathbf{a}_2$ requesting h. The literal is in the beliefs of $\mathbf{a}_2$, so the response is immediate and the beliefs are modified as before:

$\Phi'_s = \{a, b, h\}$ $\quad \Phi'_2 = \{\sim i\}$

Now, $\mathbf{a}_1$ can execute A_1, generate $\Phi_1{}' = \{c, \sim d, e, f\}$ and build an argument $\mathcal{A}^1\mathbf{a}_1 = \{q \prec c, f\}$ that supports q. $\mathcal{A}^1\mathbf{a}_1$ has no counterarguments that defeat it, so q is warranted. □

In case 3 $\mathbf{a}_1$ requests l but the literal does not belong to Φ_2, so $\mathbf{a}_2$ tries to execute an action that will allow it to add l to the set of shared beliefs.

Case 3 Let $\mathcal{D}_1 = \{p, q, s, t\}$ be $\mathbf{a}_1$'s set of beliefs and $\mathcal{I}_1 = \{s\}$ the set of committed goals. In order to build a warrant for s, it needs g and its reasoning indicates that if it were to execute A_2, the literal g would be added to Φ_1. Notwithstanding, one of A_2's preconditions is not satisfied; its knowledge does not allow it to warrant j. Then, $\mathbf{a}_1$ makes a proposal to $\mathbf{a}_2$ requiring j.

The literal j is not in $\mathbf{a}_2$'s beliefs, but its intention is to collaborate with $\mathbf{a}_1$, so it decides to execute A_3 and then modify the shared beliefs into $\Phi'_s = \{a, b, j\}$.

Note that A_3's preconditions are the literals a, h and u. The first two belong to Φ_s and Φ_2 respectively, while u is supported by $\mathcal{A}^1\mathbf{a}_2 = \{(u \prec a, v), (v \prec b)\}$. But $\mathcal{A}^1\mathbf{a}_2$ is defeated by $\mathcal{A}^2\mathbf{a}_2 = \{(\sim v \prec w), (w \prec h)\}$ that attacks v. $\mathcal{A}^2\mathbf{a}_2$ is defeated by $\mathcal{A}^3\mathbf{a}_2 = \{\sim w \prec \sim i\}$. Then, u is warranted because it is supported by the argument $\mathcal{A}^1\mathbf{a}_2$ that only has one defeater $\mathcal{A}^2\mathbf{a}_2$ which is defeated by $\mathcal{A}^3\mathbf{a}_2$ and this last argument has no defeaters.

Now, $\mathbf{a}_1$ can execute A_2, generate $\Phi_1{}' = \{c, \sim d, e, g\}$ and build and argument $\mathcal{A}^1\mathbf{a}_1 = \{s \prec e, \sim d, g\}$ that supports s. □

Frequently, $\mathbf{a}_2$'s work cannot be limited to finding a simple action that allows it to add l to Φ_s. To satisfy $\mathbf{a}_1$'s proposal, the agent may need to find a plan in which l is one of the consequences of the last action in the sequence. This situation is shown below.

Case 4 Let $\mathcal{D}_1 = \{p, q, s, t\}$ be $\mathbf{a}_1$'s set of goals $\mathbf{a}_1$ and $\mathcal{I}_1 = \{t\}$ the set of shared beliefs. In order to build an argument with no defeaters that supports t, $\mathbf{a}_1$ needs k. Because the literal is not part of its knowledge, it asks for collaboration from $\mathbf{a}_2$. $\mathbf{a}_2$'s reasoning leads it to the decision that it will not be able to collaborate with $\mathbf{a}_1$ by executing only one action, but it can elaborate a plan in which the first action will allow it to add to its beliefs the preconditions of the second one.

Then, $\mathbf{a}_2$ decides to execute A_4, which has x among its preconditions. In order to obtain a warrant for x, $\mathbf{a}_2$ needs j. As before, the literal j is not in its beliefs, but it can execute A_3, now to modify its own beliefs obtaining $\Phi'_2 = \{h, \sim i, j\}$

As a result $\mathbf{a}_2$ can execute A_4 and modify the shared knowledge $\Phi'_s = \{a, b, k\}$. Then $\mathbf{a}_1$ builds an argument with no defeaters for its committed goal t. $\square$

4.2 Proposals and Counterproposals

Suppose that $\mathbf{a}_1$ requests a literal l to $\mathbf{a}_2$ and that $\mathbf{a}_2$ can satisfy the request but to do so it will need another literal m. In this case, $\mathbf{a}_2$ will make a counterproposal, asking m to $\mathbf{a}_1$. If $\mathbf{a}_1$ is capable of satisfying that request for m, it wll add m to the set of shared beliefs. In any case, $\mathbf{a}_1$ will notify $\mathbf{a}_2$ the result of the interaction (see Section 5).

Example 4.3 Let $[\Phi_s, \mathbf{A}]$ be a group conformed by two collaborative agents, $\mathbf{a}_1 = \langle \mathcal{B}_1, \mathcal{D}_1, \mathcal{I}_1, \Gamma_1 \rangle$ and $\mathbf{a}_2 = \langle \mathcal{B}_2, \mathcal{D}_2, \mathcal{I}_2, \Gamma_2 \rangle$ as defined below:

$\Phi_s = \{a, b\}$

$\mathcal{B}_1 = (\Phi_1, \Delta_1)$ where $\Phi_1 = \{c, \sim d, e\}$ and

$\Delta_1 = \{(p \prec a, q), (q \prec c, f), (r \prec b, e), (s \prec b, c)\}$

$\Gamma_1 = \{(\{f\} \xleftarrow{A_1} \{e, j\}, not\ \{\}), (\{g\} \xleftarrow{A_2} \{\sim d\}, not\ \{\})\}$

$\mathcal{D}_1 = \{p, s\}$ is $\mathbf{a}_1$'s set of goals and $\mathcal{I}_1 = \{p\}$ the set of committed goals.

$\mathcal{B}_2 = (\Phi_2, \Delta_2)$ where $\Phi_2 = \{h, \sim i\}$ and

$\Delta_2 = \{(u \prec a, \sim i), (v \prec j)\}$

$\Gamma_2 = \{\{j\} \xleftarrow{A_3} \{b, h, u, g\}, not\ \{\}\}$

To reach p, $\mathbf{a}_1$ needs to build an argument that supports it, and its reasoning capability leads it to decide that it must execute action A_1. In order to satisfy this action's preconditions, it makes a proposal to $\mathbf{a}_2$ requiring j. This agent can only add j if it executes A_3, but in order to do so it needs g.

Agent $\mathbf{a}_2$ does not know g and it isn't capable of executing an action that will allow it to reach it. Nevertheless, it can ask for collaboration from $\mathbf{a}_1$ by means of a counterproposal. When $\mathbf{a}_1$ receives the request, it decides to execute A_2, it adds g to the set of shared beliefs, and it communicates with $\mathbf{a}_2$ to signal the completion.

When $\mathbf{a}_2$ receives $\mathbf{a}_1$'s response, it executes A_3, modifies the shared beliefs, and communicates one again. At the end of the process, the set of shared beliefs will be $\Phi'_s = \{a, b, j, g\}$. $\square$

In each of the cases analyzed so far, the negotiation ended successfully. During the search process, the agents have been able to find actions favorable to both. Evidently, this is not the only possible situation.

The negotiation can fail because of several reasons. It could be that in trying to collaborate one of the agents will introduce unexpected side-effects. Another reason is that the knowledge is not enough to build a plan, or because there is conflict among the agents' beliefs. In the following sections we will show examples that illustrate these cases.

4.3 Side-effects

The actions executed in order to change the world can modify the shared beliefs and thus produce unplanned side-effects. The example below illustrate that problem.

Example 4.4 Let $[\Phi_s, \mathbf{A}]$ be a group conformed by two collaborative agents, $\mathbf{a}_1 = \langle \mathcal{B}_1, \mathcal{D}_1, \mathcal{I}_1, \Gamma_1 \rangle$ and $\mathbf{a}_2 = \langle \mathcal{B}_2, \mathcal{D}_2, \mathcal{I}_2, \Gamma_2 \rangle$ as defined below:

$\Phi_s = \{a, b, \sim c\}$

$\mathcal{B}_1 = (\Phi_1, \Delta_1)$ where $\Phi_1 = \{\sim d, e\}$

$\Delta_1 = \{(p \prec b, j), (q \prec e, f)\}$

$\Gamma_1 = \{\{f\} \xleftarrow{A_1} \{a, \sim d, \sim c, k\}, not\ \{\}\}$

$\mathcal{D}_1 = \{p, q\}$

$\mathcal{B}_2 = (\Phi_2, \Delta_2)$ where $\Phi_2 = \{h, \sim i\}$

$\Delta_2 = \{(u \prec a, v), (v \prec b, h)\}$

$\Gamma_2 = \{(\{j, \sim b\} \xleftarrow{A_2} \{\sim i\}, not\ \{\}),$

$(\{k, c\} \xleftarrow{A_3} \{u\}, not\ \{\})\}$

If $\mathcal{I}_1 = \{p\}$, in order to achieve its intention, $\mathbf{a}_1$ can ask for collaboration from $\mathbf{a}_2$, requiring j. $\mathbf{a}_2$ reasons and decides to execute A_2, and modifies the shared beliefs into $\Phi'_s = \{a, \sim b, \sim c, j\}$.

Agent $\mathbf{a}_1$ requested j in order to build a warrant for p, but it also needed b to do so, which is no longer part of the set of shared beliefs.

Another way that conflict may arise appears when $\mathcal{I}_1 = \{q\}$. The agent $\mathbf{a}_1$ decides to execute A_1 in order to build a warrant for q. The action requires k as a precondition, and $\mathbf{a}_1$ asks for collaboration from $\mathbf{a}_2$. $\mathbf{a}_2$ is not capable of directly constructing an argument that supports k either, but it can execute A_2 and add k to Φ_s. After executing A_2, the shared beliefs are modified, because k and c are added.

Note that agent $\mathbf{a}_1$ had requested k with the intention of executing A_1, but now the precondition $\sim c$ is no longer part of the set of shared beliefs. $\square$

To avoid harming $\mathbf{a}_1$'s plans with its actions, $\mathbf{a}_2$ can communicate to $\mathbf{a}_1$ its intention of adding or removing a literal from Φ_s, besides of the one specified in the proposal. The request tends to assure mutual benefit.

When $\mathbf{a}_1$ receives the request, it analyzes if the modification that $\mathbf{a}_2$ proposes provokes any conflicts with its own knowledge, or with the action that it intends to execute. In any case, it communicates with $\mathbf{a}_2$ to accept or reject the request. If $\mathbf{a}_1$ accepts, $\mathbf{a}_2$ executes the sequence of actions that allow the addition of l to the shared knowledge, and it communicates again to indicate that its task was successful.

4.4 Failure in the Negotiation

The simplest case of failure is produced when an agent receives a proposal and it cannot satisfy it because of lack of knowledge for the execution of appropriate actions.

Example 4.5 Let $[\Phi_s, \mathbf{A}]$ be a group conformed of two collaborative agents, $\mathbf{a}_1 = \langle \mathcal{B}_1, \mathcal{D}_1, \mathcal{I}_1, \Gamma_1 \rangle$ and $\mathbf{a}_2 = \langle \mathcal{B}_2, \mathcal{D}_2, \mathcal{I}_2, \Gamma_2 \rangle$ where:
$\Phi_s = \{a, b\}$
$\mathcal{B}_1 = (\Phi_1, \Delta_1)$ where $\Phi_1 = \{c, \sim d\}$ and
$\Delta_1 = \{(p \prec c, q), (q \prec b, e), (r \prec c, s), (s \prec f, j)\}$
$\Gamma_1 = \{\{f\} \xleftarrow{A_1} \{a, \sim d\}, not\ \{\}\}$
$\mathcal{D}_1 = \{p, r\}$
$\mathcal{B}_2 = (\Phi_2, \Delta_2)$ where $\Phi_2 = \{h, \sim i\}$ and
$\Delta_2 = \{(u \prec a, v), (v \prec b, j), (w \prec h)\}$
$\Gamma_2 = \{\{j\} \xleftarrow{A_2} \{b, \sim i, g\}, not\ \{\}\}$
If $\mathbf{a}_1$'s set of committed goals is $\mathcal{I}_1 = \{p\}$, $\mathbf{a}_1$ can make a proposal to $\mathbf{a}_2$ in reference to e. The agent $\mathbf{a}_2$ cannot accept the proposal, and it cannot elaborate a counterproposal either. It then communicates to end the negotiation.

A somewhat more complicated situation presents itself when the agent that receives the proposal can elaborate counterproposal, but it is rejected. If $\mathcal{I}_1 = \{r\}$, the agent $\mathbf{a}_1$ makes a proposal to $\mathbf{a}_2$ referred to j. The agent $\mathbf{a}_2$ decides that it can execute an action that will allow it to add j to the shared knowledge, but it needs g in order to do so. It then makes a counterproposal, but $\mathbf{a}_1$ cannot accept the proposal, and can't elaborate a counterproposal either. It then communicates to inform the failure and ends the negotiation. □

The negotiation can fail, as in the previous examples, because of $\mathbf{a}_2$'s lack of knowledge and capacity to elaborate a successful plan for l. Another reason for failure is produced when there is inconsistency among $\mathbf{a}_1$'s proposal with respect to l, to achieve its intention g, and $\mathbf{a}_2$'s knowledge, committed goals, or proper plans. In any case, $\mathbf{a}_1$'s plan fails, and it must try to elaborate a new one. If all of the plans for g require l, agent $\mathbf{a}_1$ may *insist* with its proposal, forcing $\mathbf{a}_2$ to revise its own plans.

In order to be able to collaborate with $\mathbf{a}_1$, agent $\mathbf{a}_2$ has to revise its own plans too, trying to eliminate the source of conflict. If it achieves this, it accepts the offer, but if it fails, it notifies its failure again. If this is not possible, it can *demand* a revision of intentions from $\mathbf{a}_2$. It is clear that the dialogue requires a sufficiently expressive language so that each agent can express its attitude in the negotiation.

5 Communication Languages

The success of an application developed from the agent model depends considerably on the agents' ability to communicate. In this sense, the role of the *communication language*, ACL, is a fundamental one because it allow them to interact and share knowledge. Even though technically the communication takes place by messages over a network that uses a low level protocol, on a conceptual level agents do not exchange messages, but they maintain *conversations* based on its purposes. The specification of an ACL comprises of three levels [15]:

- An Interaction Protocol
- An Interaction Language
- A Language for representing Shared Beliefs

Each agent's *interaction protocol* is a *conversation pattern* that governs its interaction with other agents and allows to structure the communication. The *interaction language* is the medium through which agents communicate attitudes and exchange knowledge. The interaction language must allow agents to assist other agents, demand services, relay tasks and control its execution, commit to or reject the execution of some tasks, report its evolution and the final success or failure.

The *language for the representation of shared knowledge* should guarantee the preservation of the semantics. That is to say, a concept, object, or entity will have a uniform meaning in different applications, even when it is referenced through different names. An alternative for maintaining a body of shared knowledge about a domain is through an *ontology*. Our proposal for an interaction protocol considers that the participants in that interaction will exchange proposals and counterproposals trying to reach a general agreement.

5.1 An Interaction Protocol

Negotiation can be thought of as a distributed search process over a space consisting of potential agreements. In the majority of the cases, only a portion of the search space will satisfy each particular agent's desires. An agent makes a proposal within the space of acceptable agreements. Another member of the group receives it and may accept it, reject it, or make a counterproposal. In this last case, the first agent analyzes it again and may once again accept it, reject it, or make a counterproposal.

Acceptance indicates that a point of shared agreement has been reached. Rejection implies that the negotiation ended without success. In the process, it is possible that one or both participants will have to yield and compromise; nevertheless, if a state of acceptance is reached, the agreement must be favorable for both. Under this metaphor, the negotiation requires that the agents have certain minimal capabilities. They must be able to:

- Make a proposal within the space of acceptable situations
- Accept or reject a proposal
- Elaborate a counterproposal

The words used may have an influence in the effectiveness of a message. The use of expressions like *insist* or *demand* imply greater vehemence. Notwithstanding, the acceptance will not only depend on the words, but also on the mental attitude of the receiver with respect to the request.

The dialogue is then initiated with proposal in while an agent $\mathbf{a}_1$ *requires* from another agent $\mathbf{a}_2$ a literal p so it can execute a plan for q. Agent $\mathbf{a}_2$ can *accept* the proposal, attend to the requirement and it adds p to the shared knowledge. Alternatively, $\mathbf{a}_2$ can *reject* the proposal because p is in conflict with its own plan to achieve r. In that case, $\mathbf{a}_1$ tries to elaborate another plan for q that does not require

p, but if it cannot achieve that it can *insist* with its requirement. In the face of this insistence, $\mathbf{a}_2$ seeks another plan for r that is not in conflict with p. If it finds it, it accepts the requirement and adds p to the shared knowledge. In the other case, it rejects the proposal again.

Agent $\mathbf{a}_1$ then revises its intentions, but if it decides to persist in q, it can *demand* p to $\mathbf{a}_2$. This agent then revises its own intentions, and tries to abandon its reasons for trying to reach r, committing to a new goal s for which it can elaborate a plan without conflicts wit p.

5.2 Interaction Language

An interaction language for BDI agents should be able to express mental attitudes like beliefs, desires, and intentions. In this way, the essential functions offered by an interaction language should retain much of the aspects that allow humans to communicate.

Philosophers and linguists have developed a formal model for human communication known as *Speech Act Theory* [17, 18]. This model has served as a base for the development of languages that are oriented towards communication in Multiagent Systems [19, 20, 21].

Interaction languages based on *Speech Act Theory* capture the essential characteristics of human communication and transport them to an adequate model for the development of artificial agents. The idea is to recognize all the internal aspects of an artificial autonomous entity, considering in particular the changes that are made when it interacts with others.

An essential idea in the theory is that a communication is a special class of *action*. When an agent communicates, it doesn't only express sentences that may be true or false, but it also executes actions such as requirements, suggestions, promises, etc. The *speech act* is the minimum unit of communication. Every expression is a speech act and denotes an action at the same time.

An interaction language among agents, based on the theory of speech acts, is comprised of a set of communication *primitives*. Generally, it is possible to define a set of primitives that captures the essential aspects of communication, independent from the application domain. The semantic description of this set yields a *communication model*.

The selection of a set of primitives is related to the knowledge representation language. If every system is developed from one content specification language, the set of interaction language primitives can be reduced. If, on the other hand, different languages are used in the communication of contents, the set of primitives should be larger. In the same way, if an interaction language is going to be used in the implementation of applications in different domains (in which different content specification languages are used), the set should be versatile enough to cover all of this diversity.

5.3 Negotiation Primitives

During the negotiation, a dialogue is established in which the agents exchange messages that express proposals and counter-proposals. These messages are actions that try to modify the mental attitude of the other participant in the dialogue. Therefore, we say *communicative acts* to express the fact that communication is a particular type of action. The simplest negotiation primitives are:

- **Requests_Add**(s, h, p) Agent s will modify the shared knowledge, but in order to do so it needs authorization from h.
- **Authorize_Add**(s, h, p) Agent s received a request from h, which needs to add p to the shared knowledge; s does not find conflict, so it authorizes the request.
- **Require**(s, h, p) Agent s needs p and requests to h the addition to the shared knowledge.
- **Insist**(s, h, p) Agent s needs p and request to h the addition to the shared knowledge, even when it must reformulate its own plans.
- **Demand**(s, h, p) Agent s needs p and request to h the addition to the shared knowledge, even when it must modify its intentions.
- **Accept**(s, h, p) Agent s received a requirement from h, asking it to add p to the shared knowledge, s is capable of doing it, so it accepts the proposal.
- **Reject**(s, h, p) Agent s received a requirement from h, asking it to add p to the shared knowledge, s is not capable of doing it, so it rejects the proposal because it is in conflict with its own intentions.
- **Unable**(s, h, p) Agent s received a requirement from h, asking it to add p to the shared knowledge, s is not capable of doing it because it has not enough knowledge.

Both participants are interested in reaching an agreement ad have a cooperative attitude. The more information a counterproposal has, the faster the process (and the negotiation) will be. If $\sim q$ is part of the shared knowledge, and s needs p as part of a plan to reach q, it can communicate its intention to h, avoiding future conflicts. Some of the primitives are extended to include the intention of the agent in the message:

- **Require_for**(s, h, p, q)
- **Insist_for**(s, h, p, q)
- **Demand_for**(s, h, p, q)

Each agent has partial knowledge about the world and partial knowledge about the mental state of the rest of the agents. When an agent includes an argument in its proposal or counterproposal it is in some way showing part of its mental state.

6 Conclusions and Future Work

In this work we proposed a negotiation protocol between two BDI agents, in which both participants maintain a conversation while trying to make their initially divergent interests converge. If the process is successful, the communication ends when a shared agreement is reached.

Initially, an agent makes a request asking for another agent's knowledge in order to obtain a goal of its own. The receiver has a collaborative attitude, so it tries to cooperate. If it is successful, it modifies the set of shared beliefs. If there is a conflict, the dialogue continues, and each of the participants shows part of its mental attitude and tries to influence the other agent's plans and preferences. The communicative acts are actions that in some way provoke changes in the environment's state.

The negotiation's success depends considerably on the agents' ability for communicating and exchanging knowledge. In this sense, the language's role is fundamental; its expresiveness determines the sender's capacity to specify the message's emphasis, and the receiver's capacity to choose a course of action that is compatible with its function.

This work proposes specific negotiation primitives that can have an influence on the effectiveness of the communicative acts that are exchanged in a dialogue. The mental attitude of the receiver will be considered the each proposal, and so will the type of communicative act used in the request.

The interaction between the agents could be considerably enhanced if the context provides some *norms* that would allow the establish general agreements about the behavior. The norms don't only force the individuals to act in a specific way, but they also condition their social behavior and structure their interaction with the rest of the agents. The proposed model will be extended to reflect the impact of the norms, roles, and relationships in the negotiation process.

References

[1] G. R. Simari, A. J. García. *Actions and Arguments: Preliminaries and Examples*, in Proceedings of the VIII Congreso Argentino en Ciencias de la Computación, Universidad de Buenos Aires, Argentina, 2002. Accepted for publication.

[2] A. J. García, G. R. Simari. *Defeasible Logic Programming: An Argumentative Approach.* Theory and Practice of Logic Programming, 2002. Accepted for publication.

[3] M. Capobianco, C. I. Chesñevar, G. R. Simari. *An argumentative formalism for implementing rational agents*, Proceedings of II Workshop en Agentes y Sistemas Inteligentes (WASI), VII CACIC, Universidad Nacional de la Patagonia San Juan Bosco, El Calafate, Argentina, pp. 1051–1062, 2001.

[4] A. J. García, G. R. Simari, C. I. Chesñevar. *An Argumentative Framework for Reasoning with Inconsistent and Incomplete Information*, Workshop on Practical Reasoning and Rationality, 13th European Conference on Artificial Intelligence (ECAI-98), England, pp. 13–20, 1998.

[5] R. Fagin, J. Halpern, Y. Moses, M. Vardi. *Reasoning About Knowledge*, The MIT Press, Cambridge, MA 1995.

[6] M. Wooldridge, N. R. Jennings. *Intelligent agents: Theory and Practice.*, The Knowledge Engineering Review **10**, pp.115-152, 1995.

[7] N. R. Jennings, S. Parsons, P. Noriega, C. Sierra. *On Argumentation-Based Negotiation*, Proceedings of the International Workshop on Multi-Agent Systems, Boston, USA. 1998.

[8] N. R. Jennings, K. Sycara, M. Wooldridge. *A Roadmap of Agent Research and Development*, in Autonomous Agents and Multi-Agent Systems, **1**(1), pp. 7-38, 1998.

[9] A. S. Rao, M. P. Georgeff. *BDI Agents: From Theory to Practice*, in Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95), San Francisco, USA, pp. 312–319, 1995.

[10] M. Wooldridge. *Intelligent Agents*, in Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, G. Weiss (Ed.), The MIT Press, pp. 27–78,1999.

[11] M. Huhns, L. Stephens. *Multiagent Systems and Societies of Agents*, in Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, G. Weiss (Ed.), The MIT Press, pp. 79–120, 1999.

[12] J. Müller. *Negotiation Principles*, in Foundations of Distributed Artificial Intelligence, G. M. P. O'Hare, N. R. Jennings (Eds.), John Wiley and Sons, pp. 211–230, 1996.

[13] S. Kalenka, N. R. Jennings. *Socially Responsible Decision Making by Autonomous Agents*, in Proceedings of the 5th International Colloquium on Cognitive Science, Korta, K., Sosa, E., Arrazola (Eds.) , X. pp. 153–169, 1999.

[14] J. E. Doran, N. Franklin, N. R. Jennings, T. Norman. *On Cooperation in Multi-Agent Systems.* The Knowledge Engineering Review, **12**(3), pp. 309–314, 1996.

[15] Y. Labrou. *Semantics for an Agent Communication Language.* Ph.D. Thesis, University of Maryland, 1996.

[16] M. E. Bratman, D. J. Israel, M. E. Pollack. *Plans and Resource-Bounded Practical Reasoning.* Computational Intelligence, **4**, pp. 349–355, 1988.

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

[18] J .L. Austin. *How to Do Things with Words.* Harvard University Press, 1969.

[19] H. J. Levesque, P. R. Cohen, J. H. T. Nunes. *On Acting Together*, in Proceedings of the Eight National COnference on Artificial Intelligence, pp. 94–99, 1990.

[20] P. R. Cohen, H. J. Levesque. *Intention is Choice with Commitment.* Artificial Intelligence, **42**, pp. 213–261, 1990.

[21] P. R. Cohen, H. J. Levesque. *Rational Interaction as the Basis for Communication.* In P. R. Cohen, J. Morgan, M. E. Pollack (Eds.) *Intentions in Communication*, pp 221–256. The MIT Press, 1990.