An Abstract Framework for Argumentation-based Negotiation

Leila Amgoud, Yannis Dimopoulos, Pavlos Moraïtïs

An Abstract Framework for Argumentation-based Negotiation

Leila Amgoud
amgoud@irit.fr

Yannis Dimopoulos
yannis@cs.ucy.ac.cy

Pavlos Moraitis
pavlos@math-info.univ-paris5.fr

IRIT–CNRS
118, Route de Narbonne
31062 Toulouse cedex 09, FRANCE

Univ. of Cyprus
75 Kallipoleos Str.
PO Box 20537, Cyprus

Paris Descartes Univ.
45 rue des Saints-Pères
75270 Paris, France

Résumé :
Le papier propose un cadre abstrait pour la négociation à base d'argumentation, dans lequel le rôle de l'argumentation est formellement analysé, et les résultats d'une telle négociation sont étudiés. Il formalise la notion d'un *accord* dans une négociation. Le papier montre aussi comment cet accord est lié aux théories des agents et quand il peut être atteint. Il définit aussi la notion de *concession* et montre dans quelle situation un agent en fera une.

Mots-clés : Négociation, Argumentation

Abstract:
This paper proposes an *abstract* framework for argumentation-based negotiation, in which the role of argumentation is formally analyzed. The framework makes it possible to study the outcomes of an argumentation-based negotiation. It shows what an *agreement* is, how it is related to the theories of the agents, when it is possible, and how this can be attained by the negotiating agents in this case. It defines also the notion of *concession*, and shows in which situation an agent will make one, as well as how it influences the evolution of the dialogue.

Keywords: Negotiation, Argumentation

1 Introduction

Roughly speaking, negotiation is a process aiming at finding some compromise or consensus between two or several agents about some matters of collective agreement, such as pricing products, allocating resources, or choosing candidates.

Integrating argumentation theory in negotiation provides a good means for supplying additional information and also helps agents to convince each other by adequate arguments during a negotiation dialogue. Indeed, an offer supported by a good argument has a better chance to be accepted by an agent, and can also make him reveal his goals or give up some of them. The basic idea is that by exchanging arguments, the theories of the agents (i.e. their mental states) may evolve, and consequently, the status of offers may change. For instance, an agent may reject an offer because it is not acceptable for it. However, the agent may change its mind if it receives a strong argument in favor of this offer.

Several proposals have been made in the literature for modeling such an approach [1, 2, 6, 7, 11]. However, the work is still preliminary. Some researchers have mainly focused on relating argumentation with protocols. They have shown how and when arguments in favor of offers can be computed and exchanged. Others have emphasized on the decision making problem. In [2, 6], the authors argued that selecting an offer to propose at a given step of the dialogue is a decision making problem. They have thus proposed an argumentation-based decision model, and have shown how such a model can be related to the dialogue protocol.

In existing works, there is no formal analysis on the role of argumentation in negotiation dialogues. It is not clear how ar-

gumentation can influence the outcome of the dialogue. Moreover, basic concepts in negotiation such as *concession* and *agreement* (i.e. *optimal solutions*, or *compromise*) are neither defined nor studied.

This paper aims at proposing an *abstract* framework for argumentation-based negotiation, in which the role of argumentation is formally analyzed, and where the existing systems can be restated. In this framework, a negotiation dialogue takes place between two agents on a set $\mathcal{O}$ of offers, whose structure is not known. The goal of a negotiation is to find among elements of $\mathcal{O}$, an offer that satisfies more or less the preferences of both agents. Each agent is supposed to have a *theory* represented in an abstract way. A theory consists of a set $\mathcal{A}$ of arguments whose structure and origin are not known, a function specifying for each possible offer in $\mathcal{O}$, the arguments of $\mathcal{A}$ that support it, a non specified conflict relation among the arguments, and finally a preference relation between the arguments. The status of each argument is defined using Dung's acceptability semantics. Consequently, the set of offers is partitioned into four subsets: *acceptable*, *rejected*, *negotiable* and *non-supported* offers. We show how an agent's theory may evolve during a negotiation dialogue. We define formally the notions of concession, compromise, and optimal solution. Then, we propose a protocol that allows agents i) to exchange offers and arguments, and ii) to make concessions when necessary. We show that dialogues generated under such a protocol terminate, and even reach optimal solutions when they exist.

2 The logical language

In what follows, $\mathcal{L}$ will denote a logical language, and $\equiv$ is an equivalence relation associated with it.

From $\mathcal{L}$, a set $\mathcal{O} = \{o_1, \ldots, o_n\}$ of n *offers* is identified, such that $\nexists o_i, o_j \in \mathcal{O}$ such that $o_i \equiv o_j$. This means that the offers are

different. Offers correspond to the different alternatives that can be exchanged during a negotiation dialogue. For instance, if the agents try to decide the place of their next meeting, then the set $\mathcal{O}$ will contain different towns.

Different *arguments* can be built from $\mathcal{L}$. The set $\texttt{Args}(\mathcal{L})$ will contain all those arguments. By argument, we mean a *reason* in believing or of doing something. In [2], it has been argued that the selection of the best offer to propose at a given step of the dialogue is a decision problem. In [3], it has been shown that in an argumentation-based approach for decision making, two kinds of arguments are distinguished: arguments supporting choices (or decisions), and arguments supporting beliefs. Moreover, it has been acknowledged that the two categories of arguments are formally defined in different ways, and they play different roles. Indeed, an argument in favor of a decision, built both on an agent's beliefs and goals, tries to justify the choice; whereas an argument in favor of a belief, built only from beliefs, tries to destroy the decision arguments, in particular the beliefs part of those decision arguments. Consequently, in a negotiation dialogue, those two kinds of arguments are generally exchanged between agents. In what follows, the set $\texttt{Args}(\mathcal{L})$ is then divided into two subsets: a subset $\texttt{Args}_o(\mathcal{L})$ of arguments supporting offers, and a subset $\texttt{Args}_b(\mathcal{L})$ of arguments supporting beliefs. Thus, $\texttt{Args}(\mathcal{L})$ = $\texttt{Args}_o(\mathcal{L}) \cup \texttt{Args}_b(\mathcal{L})$. As in [4], in what follows, we consider that the structure of the arguments is not known.

Since the knowledge bases from which arguments are built may be inconsistent, the arguments may be conflicting too. In what follows, those conflicts will be captured by the relation $\mathcal{R}_{\mathcal{L}}$, thus $\mathcal{R}_{\mathcal{L}} \subseteq \texttt{Args}(\mathcal{L}) \times \texttt{Args}(\mathcal{L})$. Three assumptions are made on this relation: First the arguments supporting different offers are conflicting. The idea behind this assumption is that since offers are exclusive, an agent has to choose

only one at a given step of the dialogue. Note that, the relation $\mathcal{R}_{\mathcal{L}}$ is not necessarily symmetric between the arguments of $\text{Args}_b(\mathcal{L})$. The second hypothesis says that arguments supporting the same offer are also conflicting. The idea here is to return the strongest argument among these arguments. The third condition does not allow an argument in favor of an offer to attack an argument supporting a belief. This avoids wishful thinking. Formally:

Definition 1 $\mathcal{R}_{\mathcal{L}} \subseteq \text{Args}(\mathcal{L}) \times \text{Args}(\mathcal{L})$ *is a* conflict relation *among arguments s.t:*
- $\forall a, a' \in \text{Args}_o(\mathcal{L})$, *s.t.* $a \neq a'$, $a \, \mathcal{R}_{\mathcal{L}} \, a'$
- $\nexists \, a \in \text{Args}_o(\mathcal{L})$ *and* $a' \in \text{Args}_b(\mathcal{L})$ *s.t* $a \, \mathcal{R}_{\mathcal{L}} \, a'$

Note that the relation $\mathcal{R}_{\mathcal{L}}$ is not symmetric. This is due to the fact that arguments of $\text{Args}_b(\mathcal{L})$ may be conflicting but not necessarily in a symmetric way. In what follows, we assume that the set $\text{Args}(\mathcal{L})$ of arguments is finite, and each argument is attacked by a finite number of arguments.

3 Negotiating agents theories and reasoning models

In this section we define formally the negotiating agents, i.e. their theories, as well as the reasoning model used by those agents in a negotiation dialogue.

3.1 Negotiating agents theories

Agents involved in a negotiation dialogue, called negotiating agents, are supposed to have theories. In this paper, the theory of an agent will not refer, as usual, to its mental states (i.e. its beliefs, desires and intentions). However, it will be encoded in a more abstract way in terms of the arguments owned by the agent, a conflict relation among those arguments, a preference relation between the arguments, and a function that specifies which arguments

support offers of the set $\mathcal{O}$. We assume that an agent is aware of all the arguments of the set $\text{Args}(\mathcal{L})$. The agent is even able to express a preference between any pair of arguments. This does not mean that the agent will use all the arguments of $\text{Args}(\mathcal{L})$, but it encodes the fact that when an agent receives an argument from another agent, it can interpret it correctly, and it can also compare it with its own arguments. Similarly, each agent is supposed to be aware of the conflicts between arguments. This also allows us to encode the fact that an agent can recognize whether the received argument is in conflict or not with its arguments. However, in its theory, only the conflicts between its own arguments are considered.

Definition 2 (Negotiating agent theory)
Let $\mathcal{O}$ be a set of n offers. A negotiating agent theory is a tuple $\langle \mathcal{A}, \mathcal{F}, \succeq, \mathcal{R}, \text{Def} \rangle$:

- $\mathcal{A} \subseteq \text{Args}(\mathcal{L})$.

- $\mathcal{F}: \mathcal{O} \rightarrow 2^{\mathcal{A}}$ *s.t* $\forall i, j$ *with* $i \neq j$, $\mathcal{F}(o_i) \cap \mathcal{F}(o_j) = \emptyset$. *Let $\mathcal{A}_{\mathcal{O}} = \cup \mathcal{F}(o_i)$ with $i = 1, \ldots, n$.*

- $\succeq \, \subseteq \text{Args}(\mathcal{L}) \times \text{Args}(\mathcal{L})$ *is a partial preorder denoting a preference relation between arguments.*

- $\mathcal{R} \subseteq \mathcal{R}_{\mathcal{L}}$ *s.t* $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$

- $\text{Def} \subseteq \mathcal{A} \times \mathcal{A}$ *s.t* $\forall \, a, b \in \mathcal{A}$, *a defeats b, denoted $a \, \text{Def} \, b$ iff $a \, \mathcal{R} \, b$, and not ($b \succeq a$).*

The function $\mathcal{F}$ returns the arguments supporting offers in $\mathcal{O}$. We assume that an argument cannot support two distinct offers. However, $\mathcal{F}(o_i)$ may be empty.

Example 1 *Let $\mathcal{O} = \{o_1, o_2, o_3\}$.*

- $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$
- $\mathcal{F}(o_1) = \{a_1\}$, $\mathcal{F}(o_2) = \{a_2\}$, $\mathcal{F}(o_3) = \emptyset$. *Thus, $\mathcal{A}_o = \{a_1, a_2\}$*

- $\succeq = \{(a_1, a_2), (a_2, a_1), (a_3, a_2), (a_4, a_3)\}$

- $\mathcal{R} = \{a_1, a_2), (a_2, a_1), (a_3, a_2), (a_4, a_3)\}$

- $\text{Def} = \{(a_4, a_3), (a_3, a_2)\}$

3.2 The reasoning model

From the theory of an agent, one can define the argumentation system used by that agent for reasoning about the offers and the arguments, i.e. for computing the status of the different offers and arguments.

Definition 3 (Argumentation system)
Let $\langle \mathcal{A}, \mathcal{F}, \succeq, \mathcal{R}, \text{Def} \rangle$ be the theory of an agent. The argumentation system *of that agent is the pair $\langle \mathcal{A}, \text{Def} \rangle$.*

In [4], different acceptability semantics have been introduced for computing the status of arguments. These are based on two basic concepts, *defence* and *conflict-free*, defined as follows:

Definition 4 (Defence/conflict-free) *Let $S \subseteq \mathcal{A}$.*

- S defends *an argument a iff each argument that defeats a is defeated by some argument in S.*

- S is conflict-free *iff there exist no a, a' in S such that a Def a'.*

Definition 5 (Acceptability semantics)
Let S be a conflict-free set of arguments, and let $\mathcal{T}: 2^{\mathcal{A}} \to 2^{\mathcal{A}}$ be a function such that $\mathcal{T}(S) = \{a \mid a$ is defended by $S\}$.

- S is a complete *extension iff $S = \mathcal{T}(S)$.*

- S is a preferred *extension iff S is a maximal (w.r.t set $\subseteq$) complete extension.*

- S is a grounded *extension iff it is the smallest (w.r.t set $\subseteq$) complete extension.*

Let $\mathcal{E}_1, \ldots, \mathcal{E}_x$ denote the different extensions under a given semantics.

Note that there is only one grounded extension. It contains all the arguments that are not defeated, and those arguments that are defended directly or indirectly by non-defeated arguments.

Theorem 1 *Let $\langle \mathcal{A}, \text{Def} \rangle$ the argumentation system defined as shown above.*

1. *It may have $x \geq 1$ preferred extensions.*

2. *The grounded extensions is $\mathcal{S} = \bigcup^{i \geq 1} \mathcal{T}(\emptyset)$.*

Note that when the grounded extension (or the preferred extension) is empty, this means that there is no acceptable offer for the negotiating agent.

Example 2 *In example 1, there is one preferred extension, $\mathcal{E} = \{a_1, a_2, a_4\}$.*

Now that the acceptability semantics is defined, we are ready to define the status of any argument.

Definition 6 (Argument status) *Let $\langle \mathcal{A}, \text{Def} \rangle$ be an argumentation system, and $\mathcal{E}_1, \ldots, \mathcal{E}_x$ its extensions under a given semantics. Let $a \in \mathcal{A}$.*

1. a is accepted *iff $a \in \mathcal{E}_i$, $\forall \mathcal{E}_i$ with $i = 1, \ldots, x$.*

2. a is rejected *iff $\nexists \mathcal{E}_i$ such that $a \in \mathcal{E}_i$.*

3. a is undecided *iff a is neither accepted nor rejected. This means that a is in some extensions and not in others.*

Note that $\mathcal{A} = \{a|a \text{ is accepted}\} \cup \{a|a \text{ is rejected}\} \cup \{a|a \text{ is undecided}\}$.

Example 3 *In example 1, the arguments a_1, a_2 and a_4 are accepted, whereas the argument a_3 is rejected.*

As said before, agents use argumentation systems for reasoning about offers. In a negotiation dialogue, agents propose and accept offers that are acceptable for them, and reject bad ones. In what follows, we will define the status of an offer. According to the status of arguments, one can define four statuses of the offers as follows:

Definition 7 (Offers status) *Let $o \in \mathcal{O}$.*

- *The offer o is* acceptable *for the negotiating agent iff $\exists\ a \in \mathcal{F}(o)$ s.t a is accepted. $\mathcal{O}_a = \{o_i \in \mathcal{O}, \text{ s.t } o_i \text{ is acceptable}\}$.*

- *The offer o is* rejected *for the negotiating agent iff $\forall\ a \in \mathcal{F}(o)$, a is rejected. $\mathcal{O}_r = \{o_i \in \mathcal{O}, \text{ s.t } o_i \text{ is rejected}\}$.*

- *The offer o is* negotiable *iff $\forall\ a \in \mathcal{F}(o)$, a is undecided. $\mathcal{O}_n = \{o_i \in \mathcal{O}, \text{ s.t } o_i \text{ is negotiable}\}$.*

- *The offer o is* non-supported *iff it is neither acceptable, nor rejected or negotiable. $\mathcal{O}_{ns} = \{o_i \in \mathcal{O}, \text{ s.t } o_i \text{ is non-supported offers}\}$.*

Example 4 *In example 1, the two offers o_1 and o_2 are acceptable since they are supported by accepted arguments, whereas the offer o_3 is non-supported since it has no argument in its favor.*

From the above definitions, the following results hold:

Property 1 *Let $o \in \mathcal{O}$.*

- $\mathcal{O} = \mathcal{O}_a \cup \mathcal{O}_r \cup \mathcal{O}_n \cup \mathcal{O}_{ns}$.

- *The set $\mathcal{O}_a$ may contain more than one offer.*

From the above partition of the set $\mathcal{O}$ of offers, a preference relation between offers is defined. Let $\mathcal{O}_x$ and $\mathcal{O}_y$ be two subsets of $\mathcal{O}$. $\mathcal{O}_x \rhd \mathcal{O}_y$ means that any offer in $\mathcal{O}_x$ is preferred to any offer in the set $\mathcal{O}_y$. We can write also for two offers o_i, o_j, $o_i \rhd o_j$ iff $o_i \in \mathcal{O}_x$, $o_j \in \mathcal{O}_y$ and $\mathcal{O}_x \rhd \mathcal{O}_y$.

Definition 8 (Preference between offers) *Let $\mathcal{O}$ be a set of offers, and $\mathcal{O}_a$, $\mathcal{O}_r$, $\mathcal{O}_n$, $\mathcal{O}_{ns}$ its partition. $\mathcal{O}_a \rhd \mathcal{O}_n \rhd \mathcal{O}_{ns} \rhd \mathcal{O}_r$.*

Example 5 *In example 1, we have $o_1 \rhd o_3$, and $o_2 \rhd o_3$. However, o_1 and o_2 are indifferent.*

4 Argumentation-based Negotiation

In this section, we define formally a protocol that generates argumentation-based negotiation dialogues between two negotiating agents P and C. The two agents negotiate about an object whose possible values belong to a set $\mathcal{O}$. This set $\mathcal{O}$ is supposed to be known and the same for both agents. For simplicity reasons, we assume that this set does not change during the dialogue. The agents are equipped with theories denoted respectively $\langle \mathcal{A}^P$, $\mathcal{F}^P$, $\succeq^P$, $\mathcal{R}^P$, $\text{Def}^P \rangle$, and $\langle \mathcal{A}^C$, $\mathcal{F}^C$, $\succeq^C$, $\mathcal{R}^C$, $\text{Def}^C \rangle$. Note that the two theories may be different in the sense that the agents may have different sets of arguments, and different preference relations. Worst yet, they may have different arguments in favor of the same offers. Moreover, these theories may *evolve* during the dialogue.

4.1 Evolution of the theories

Before defining formally the evolution of an agent's theory, let us first introduce the

notion of moves.

Definition 9 (Move) *A* move *is a tuple* m_i = $\langle p_i, a_i, o_i, t_i \rangle$ *such that:*

- $p_i \in \{P, C\}$

- $a_i \in \mathrm{Args}(\mathcal{L}) \cup \theta^1$

- $o_i \in \mathcal{O} \cup \theta$

- $t_i \in \mathcal{N}^*$ *is the target of the move, such that* $t_i < i$

The function Player *(resp.* Argument, Offer, Target*) returns the player of the move (i.e.* p_i*) (resp. the argument of a move, i.e* a_i*, the offer* o_i*, and the target of the move,* t_i*). Let* $\mathcal{M}$ *denote the set of all the moves that can be built from* $\langle\{P, C\}, \mathrm{Arg}(\mathcal{L}), \mathcal{O}\rangle$.

Note that the set $\mathcal{M}$ is finite since $\mathrm{Arg}(\mathcal{L})$ and $\mathcal{O}$ are assumed to be finite. Let us now see how an agent's theory evolves and why. The idea is that if an agent receives an argument from another agent, it will add the new argument to its theory. Moreover, since an argument may bring new information for the agent, thus new arguments can emerge.

Example 6 *Suppose that an agent* P *has the following propositional knowledge base:* $\Sigma_P = \{x, y \rightarrow z\}$. *From this base one cannot deduce* z. *Let's assume that this agent receives the following argument* $\{a, a \rightarrow y\}$ *that justifies* y. *It is clear that now* P *can build an argument, say* $\{a, a \rightarrow y, y \rightarrow z\}$ *in favor of* z.

In a similar way, if a received argument is in conflict with the arguments of the agent i, then those conflicts are also added to its relation $\mathcal{R}^i$. Note that new conflicts may arise between the original arguments

---
[1] In what follows θ denotes the fact that no argument, or no offer is given

of the agent and the ones that emerge after adding the received arguments to its theory. Those new conflicts should also be considered. As a direct consequence of the evolution of the sets $\mathcal{A}^i$ and $\mathcal{R}^i$, the defeat relation Def^i is also updated.

The initial theory of an agent i, (i.e. its theory before the dialogue starts), is denoted by $\langle \mathcal{A}_0^i, \mathcal{F}_0^i, \succeq_0^i, \mathcal{R}_0^i, \mathrm{Def}_0^i \rangle$, with $i \in \{P, C\}$. Besides, in this paper, we suppose that the preference relation $\succeq^i$ of an agent does not change during the dialogue.

Definition 10 (Theory evolution) *Let* $m_1, \ldots, m_t, \ldots, m_j$ *be a sequence of moves. The theory of an agent* i *at a step* $t > 0$ *is:* $\langle \mathcal{A}_t^i, \mathcal{F}_t^i, \succeq_t^i, \mathcal{R}_t^i, \mathrm{Def}_t^i \rangle$ *such that:*

- $\mathcal{A}_t^i = \mathcal{A}_0^i \cup \{a_i, \ i = 1, \ldots, t, \ a_i = \mathrm{Argument}(m_i)\} \cup \mathcal{A}'$ *with* $\mathcal{A}' \subseteq \mathrm{Args}(\mathcal{L})$

- $\mathcal{F}_t^i = \mathcal{O} \rightarrow 2^{\mathcal{A}_t^i}$

- $\succeq_t^i = \succeq_0^i$

- $\mathcal{R}_t^i = \mathcal{R}_0^i \cup \{(a_i, a_j) \mid a_i = \mathrm{Argument}(m_i), a_j = \mathrm{Argument}(m_j), i, j \leq t,$ *and* $a_i \ \mathcal{R}_\mathcal{L} \ a_j\} \cup \mathcal{R}'$ *with* $\mathcal{R}' \subseteq \mathcal{R}_\mathcal{L}$

- $\mathrm{Def}_t^i \subseteq \mathcal{A}_t^i \times \mathcal{A}_t^i$

The above definition captures the *monotonic* aspect of an argument. Indeed, an argument cannot be removed. However, its status may change. An argument that is accepted at step t of the dialogue by an agent may become rejected at step $t + i$. Consequently, the status of offers also change. Thus, the sets $\mathcal{O}_a$, $\mathcal{O}_r$, $\mathcal{O}_n$, and $\mathcal{O}_{ns}$ may change from one step of the dialogue to another. That means for example that some offers could move from the set $\mathcal{O}_a$ to the set $\mathcal{O}_r$ and vice-versa. Note that in the definition of $\mathcal{R}_t$, the relation $\mathcal{R}_\mathcal{L}$ is used to denote a conflict between exchanged arguments. The reason is that, such a conflict may not be in the set $\mathcal{R}^i$ of the agent

i. Thus, in order to recognize such conflicts, we have supposed that the set $\mathcal{R}_{\mathcal{L}}$ is known to the agents. This allows us to capture the situation where an agent is able to prove an argument that it was unable to prove before, by incorporating in its beliefs some information conveyed through the exchange of arguments with another agent. This, unknown at the beginning of the dialogue argument, could give to this agent the possibility to defeat an argument that it could not by using its initial arguments. This could even lead to a change of the status of these initial arguments and this change would lead to the one of the associated offers' status.

In what follows, $\mathcal{O}_{t,x}^i$ denotes the set of offers of type x, where $x \in \{a, n, r, ns\}$, of the agent i at step t of the dialogue. In some places, we can use for short the notation $\mathcal{O}_t^i$ to denote the partition of the set $\mathcal{O}$ at step t for agent i. Note that we have: not($\mathcal{O}_{t,x}^i \subseteq \mathcal{O}_{t+1,x}^i$).

4.2 The notion of agreement

As said in the introduction, negotiation is a process aiming at finding an *agreement* about some matters. By agreement, one means a solution that satisfies to the largest possible extent the preferences of both agents. In case there is no such solution, we say that the negotiation fails. In what follows, we will discuss the different kinds of solutions that may be reached in a negotiation. The first one is the *optimal* solution. An optimal solution is the best offer for both agents. Formally:

Definition 11 (Optimal solution) *Let* $\mathcal{O}$ *be a set of offers, and* $o \in \mathcal{O}$. *The offer* o *is an* optimal *solution at a step* $t \geq 0$ *iff* $o \in \mathcal{O}_{t,a}^P \cap \mathcal{O}_{t,a}^C$

Such a solution does not always exist since agents may have conflicting preferences.

Thus, agents make *concessions* by proposing/accepting less preferred offers.

Definition 12 (Concession) *Let* $o \in \mathcal{O}$. *The offer* o *is a* concession *for an agent* i *iff* $o \in \mathcal{O}_x^i$ *such that* $\exists \mathcal{O}_y^i \neq \emptyset$, *and* $\mathcal{O}_y^i \rhd \mathcal{O}_x^i$.

During a negotiation dialogue, agents exchange first their most preferred offers, and if these last are rejected, they make concessions. In this case, we say that their best offers are no longer *defendable*. In an argumentation setting, this means that the agent has already presented all its arguments supporting its best offers, and it has no counter argument against the ones presented by the other agent. Formally:

Definition 13 (Defendable offer) *Let* $\langle \mathcal{A}_t^i, \mathcal{F}_t^i, \succeq_t^i, \mathcal{R}_t^i, \mathtt{Def}_t^i \rangle$ *be the theory of agent* i *at a step* $t > 0$ *of the dialogue. Let* $o \in \mathcal{O}$ *such that* $\exists j \leq t$ *with* $\mathtt{Player}(m_j) = i$ *and* $\mathtt{offer}(m_j) = o$. *The offer* o *is* defendable *by the agent* i *iff:*

- $\exists a \in \mathcal{F}_t^i(o)$, *and* $\nexists k \leq t$ *s.t.* $\mathtt{Argument}(m_k) = a$, *or*

- $\exists a \in \mathcal{A}^t \backslash \mathcal{F}_t^i(o)$ *s.t.* $a \, \mathtt{Def}_t^i \, b$ *with*
  - $\mathtt{Argument}(m_k) = b$, $k \leq t$, *and* $\mathtt{Player}(m_k) \neq i$
  - $\nexists l \leq t$, $\mathtt{Argument}(m_l) = a$

The offer o *is said* non-defendable *otherwise and* ND_t^i *is the set of non-defendable offers of agent* i *at a step* t.

4.3 Negotiation dialogue

Now that we have shown how the theories of the agents evolve during a dialogue, we are ready to define formally an argumentation-based negotiation dialogue. For that purpose, we need to define first the notion of a legal continuation.

Definition 14 (Legal move) *A move m is a* legal continuation *of a sequence of moves $m_1, \ldots, m_l$ iff $\nexists j, k < l$, such that:*

- Offer(m_j) = Offer(m_k), *and*

- Player$(m_j) \neq$ Player(m_k)

The idea here is that if the two agents present the same offer, then the dialogue should terminate, and there is no longer possible continuation of the dialogue.

Definition 15 (Negotiation) *A negotiation dialogue d between two agents P and C is a non-empty sequence of moves $m_1, \ldots, m_l$ s.t:*

- $p_i = P$ *iff i is even, and $p_i = C$ iff i is odd*

- Player$(m_1) = P$, Argument$(m_1) = \theta$, Offer$(m_1) \neq \theta$, *and* Target$(m_1) = 0^2$

- $\forall \ m_i,$ *if* Offer$(m_i) \neq \theta$, *then* Offer$(m_i) \rhd o_j, \ \forall \ o_j \in \mathcal{O} \backslash (\mathcal{O}_{i,r}^{\texttt{Player}(m_i)} \cup ND_i^{\texttt{Player}(m_i)})$

- $\forall i = 1, \ldots, l,$ m_i *is a legal continuation of $m_1, \ldots, m_{i-1}$*

- Target$(m_i) = m_j$ *such that $j < i$ and* Player$(m_i) \neq$ Player(m_j)

- *If* Argument$(m_i) \neq \theta$, *then:*
    - *if* Offer$(m_i) \neq \theta$ *then* Argument$(m_i) \in \mathcal{F}($Offer$(m_i))$
    - *if* Offer$(m_i) = \theta$ *then* Argument(m_i) Def$_i^{\texttt{Player}(m_i)}$ Argument$($Target$(m_i))$

- $\nexists \ i, j \leq l$ *such that $m_i = m_j$*

- $\nexists \ m \in \mathcal{M}$ *such that m is a legal continuation of $m_1, \ldots, m_l$*

---

²The first move has no target.

Let $\mathcal{D}$ be the set of all possible dialogues.

The first condition says that the two agents take turn. The second condition says that agent P starts the negotiation dialogue by presenting an offer. Note that, in the first turn, we suppose that the agent does not present an argument. This assumption is made for strategical purposes. Indeed, arguments are exchanged as soon as a conflict appears. The third condition ensures that agents exchange their best offers, but never the rejected ones. This condition takes also into account the concessions that an agent will have to make if it was established that a concession is the only option for it at the current state of the dialogue. Of course, as we have shown in a previous section, an agent may have several good or acceptable offers. In this case, the agent chooses one of them randomly. The fourth condition ensures that the moves are legal. This condition allows to terminate the dialogue as soon as an offer is presented by both agents. The fifth condition allows agents to backtrack. The sixth condition says that an agent may send arguments in favor of offers, and in this case the offer should be stated in the same move. An agent can also send arguments in order to defeat arguments of the other agent. The next condition prevents repeating the same move. This is useful for avoiding loops. The last condition ensures that all the possible legal moves have been presented.

The outcome of a negotiation dialogue is computed as follows:

Definition 16 (Dialogue outcome) *Let $d = m_1, \ldots, m_l$ be a argumentation-based negotiation dialogue. The outcome of this dialogue, denoted* Outcome, *is* Outcome$(d) =$ Offer(m_l) *iff $\exists j < l$ s.t.* Offer$(m_l) =$ Offer(m_j), *and* Player$(m_l) \neq$ Player(m_j). *Otherwise,* Outcome$(d) = \theta$.

Note that when $\texttt{Outcome}(d) = \theta$, the negotiation fails, and no agreement is reached by the two agents. However, if $\texttt{Outcome}(d) \neq \theta$, the negotiation succeeds, and a solution that is either optimal or a compromise is found.

Theorem 2 $\forall d_i \in \mathcal{D}$, the argumentation-based negotiation d_i terminates.

The above result is of great importance, since it shows that the proposed protocol avoids loops, and dialogues terminate. Another important result shows that the proposed protocol ensures to reach an optimal solution if it exists. Formally:

Theorem 3 (Completeness) Let $d = m_1, \ldots, m_l$ be a argumentation-based negotiation dialogue. If $\exists t \leq l$ such that $\mathcal{O}_{t,a}^P \cap \mathcal{O}_{t,a}^C \neq \emptyset$, then $\texttt{Outcome}(d) \in \mathcal{O}_{t,a}^P \cap \mathcal{O}_{t,a}^C$.

We show also that the proposed dialogue protocol is sound in the sense that, if a dialogue returns a solution, then that solution is for sure a compromise. In other words, that solution is a "common agreement" at a given step of the dialogue. We show also that if the negotiation fails, then there is no possible solution.

Theorem 4 (Soundness) Let $d = m_1, \ldots, m_l$ be a argumentation-based negotiation dialogue.

1. If $\texttt{Outcome}(d) = o$, ($o \neq \theta$), then $\exists t \leq l$ such that $o \in \mathcal{O}_{t,x}^P \cap \mathcal{O}_{t,y}^C$, with $x, y \in \{a, n, ns\}$.

2. If $\texttt{Outcome}(d) = \theta$, then $\forall t \leq l$, $\mathcal{O}_{t,x}^P \cap \mathcal{O}_{t,y}^C = \emptyset$, $\forall x, y \in \{a, n, ns\}$.

A direct consequence of the above theorem is the following:

Property 2 Let $d = m_1, \ldots, m_l$ be a argumentation-based negotiation dialogue. If $\texttt{Outcome}(d) = \theta$, then $\forall t \leq l$,

- $\mathcal{O}_{t,r}^P = \mathcal{O}_{t,a}^C \cup \mathcal{O}_{t,n}^C \cup \mathcal{O}_{t,ns}^C$, and

- $\mathcal{O}_{t,r}^C = \mathcal{O}_{t,a}^P \cup \mathcal{O}_{t,n}^P \cup \mathcal{O}_{t,ns}^P$.

5 Illustrative examples

In this section we will present some examples that illustrate the framework.

Example 7 (No argumentation) Let $\mathcal{O} = \{o_1, o_2\}$, P and C are two agents equipped with the same theory: $\langle \mathcal{A}, \mathcal{F}, \succeq, \mathcal{R}, \texttt{Def} \rangle$ s.t. $\mathcal{A} = \emptyset$, $\mathcal{F}(o_1) = \mathcal{F}(o_2) = \emptyset$, $\succeq = \emptyset$, $\mathcal{R} = \emptyset$, $\texttt{Def} = \emptyset$. It is clear that the two offers o_1 and o_2 are non-supported. The proposed protocol (see Definition 15) will generate one of the following dialogues:

P: $m_1 = \langle P, \theta, o_1, 0 \rangle$

C: $m_2 = \langle C, \theta, o_1, 1 \rangle$

This dialogue ends with o_1 as a compromise. This solution is optimal since it is not an acceptable offer for the agents.

P: $m_1 = \langle P, \theta, o_1, 0 \rangle$

C: $m_2 = \langle C, \theta, o_2, 1 \rangle$

P: $m_3 = \langle P, \theta, o_2, 2 \rangle$

This dialogue ends with o_2 as a compromise.

P: $m_1 = \langle P, \theta, o_2, 0 \rangle$

C: $m_2 = \langle C, \theta, o_2, 1 \rangle$

This dialogue also ends with o_2 as a compromise. The last possible dialgue ends with o_1 as a compromise.

P: $m_1 = \langle P, \theta, o_2, 0 \rangle$

C: $m_2 = \langle C, \theta, o_1, 1 \rangle$

P: $m_3 = \langle P, \theta, o_1, 2 \rangle$

In the above example, the theories of both agents do not change since there is no exchange of arguments. Let us now consider the following example.

Example 8 (Static theories) *Let* $\mathcal{O} = \{o_1, o_2\}$ *be the set of all possible offers. The theory of agent* P *is* $\langle \mathcal{A}^P, \mathcal{F}^P, \succeq^P, \mathcal{R}^P, \text{Def}^P \rangle$ *such that:* $\mathcal{A}^P = \{a_1, a_2\}$, $\mathcal{F}^P(o_1) = \{a_1\}$, $\mathcal{F}^P(o_2) = \{a_2\}$, $\succeq^P = \{(a_1, a_2)\}$, $\mathcal{R}^P = \{(a_1, a_2), (a_2, a_1)\}$, $\text{Def}^P = \{a_1, a_2\}$. *The argumentation system* $\langle \mathcal{A}^P, \text{Def}^P \rangle$ *of this agent will return* a_1 *as an accepted argument, and* a_2 *as a rejected one. Consequently, the offer* o_1 *is acceptable and* o_2 *is rejected.*

The theory of agent C *is* $\langle \mathcal{A}^C, \mathcal{F}^C, \succeq^C, \mathcal{R}^C, \text{Def}^C \rangle$ *such that:* $\mathcal{A}^C = \{a_1, a_2\}$, $\mathcal{F}^C(o_1) = \{a_1\}$, $\mathcal{F}^C(o_2) = \{a_2\}$, $\succeq^C = \{(a_2, a_1)\}$, $\mathcal{R}^C = \{(a_1, a_2), (a_2, a_1)\}$, $\text{Def}^C = \{a_2, a_1\}$. *The argumentation system* $\langle \mathcal{A}^C, \text{Def}^C \rangle$ *of this agent will return* a_2 *as an accepted argument, and* a_1 *as a rejected one. Consequently, the offer* o_2 *is acceptable and* o_1 *is rejected. The first possible dialogue is:*

P: $m_1 = \langle P, \theta, o_1, 0 \rangle$

C: $m_2 = \langle C, \theta, o_2, 1 \rangle$

P: $m_3 = \langle P, a_1, o_1, 2 \rangle$

C: $m_4 = \langle C, a_2, o_2, 3 \rangle$

The second possible dialogue is:

P: $m_1 = \langle P, \theta, o_1, 0 \rangle$

C: $m_2 = \langle C, a_2, o_2, 1 \rangle$

P: $m_3 = \langle P, a_1, o_1, 2 \rangle$

C: $m_4 = \langle C, \theta, o_2, 3 \rangle$

Both dialogues end with failure. Note that in both dialogues, the theories of both agents do not change. The reason is that the exchanged arguments are already known to both agents. The negotiation fails because the agents have conflicting preferences.

Let us now consider an example in which argumentation will allow agents to reach an agreement.

Example 9 (Dynamic theories) *Let* $\mathcal{O} = \{o_1, o_2\}$ *be the set of all possible offers. The theory of agent* P *is* $\langle \mathcal{A}^P, \mathcal{F}^P, \succeq^P, \mathcal{R}^P, \text{Def}^P \rangle$ *such that:* $\mathcal{A}^P = \{a_1, a_2\}$, $\mathcal{F}^P(o_1) = \{a_1\}$, $\mathcal{F}^P(o_2) = \{a_2\}$, $\succeq^P = \{(a_1, a_2), (a_3, a_1)\}$, $\mathcal{R}^P = \{(a_1, a_2), (a_2, a_1)\}$, $\text{Def}^P = \{(a_1, a_2)\}$. *The argumentation system* $\langle \mathcal{A}^P, \text{Def}^P \rangle$ *of this agent will return* a_1 *as an accepted argument, and* a_2 *as a rejected one. Consequently, the offer* o_1 *is acceptable and* o_2 *is rejected.*

The theory of agent C *is* $\langle \mathcal{A}^C, \mathcal{F}^C, \succeq^C, \mathcal{R}^C, \text{Def}^C \rangle$ *such that:* $\mathcal{A}^C = \{a_1, a_2, a_3\}$, $\mathcal{F}^C(o_1) = \{a_1\}$, $\mathcal{F}^C(o_2) = \{a_2\}$, $\succeq^C = \{(a_1, a_2), (a_3, a_1)\}$, $\mathcal{R}^C = \{(a_1, a_2), (a_2, a_1), (a_3, a_1)\}$, $\text{Def}^C = \{(a_1, a_2), (a_3, a_1)\}$. *The argumentation system* $\langle \mathcal{A}^C, \text{Def}^C \rangle$ *of this agent will return* a_3 *and* a_2 *as accepted arguments, and* a_1 *as a rejected one. Consequently, the offer* o_2 *is acceptable and* o_1 *is rejected.*

The following dialogue may take place between the two agents:

P: $m_1 = \langle P, \theta, o_1, 0 \rangle$

C: $m_2 = \langle C, \theta, o_2, 1 \rangle$

P: $m_3 = \langle P, a_1, o_1, 2 \rangle$

C: $m_4 = \langle C, a_3, \theta, 3 \rangle$

C: $m_5 = \langle P, \theta, o_2, 4 \rangle$

At step 4 of the dialogue, the agent P receives the argument a_3 from P. Thus, its theory evolves as follows: $\mathcal{A}^P = \{a_1, a_2, a_3\}$, $\mathcal{R}^P = \{(a_1, a_2), (a_2, a_1), (a_3, a_1)\}$, $\text{Def}^P = \{(a_1, a_2), (a_3, a_1)\}$. At this step, the argument a_1 which was accepted will become rejected, and the argument a_2 which was at the beginning of the dialogue rejected will become accepted. Thus, the offer o_2 will be acceptable for the agent, whereas o_1 will become rejected. At this step 4, the offer o_2 is acceptable for both agents, thus it is an optimal solution. The dialogue ends by returning this offer as an outcome.

6 Related work

Argumentation has been integrated in negotiation dialogues at the early nineties by Sycara [11]. In that work, the author has emphasized the advantages of using argumentation in negotiation dialogues, and a specific framework has been introduced. In [7], the different types of arguments that are used in a negotiation dialogue, such as threats and rewards, have been discussed. Moreover, a particular framework for negotiation have been proposed. In [8], different other frameworks have been proposed. Even if all these frameworks are based on different logics, and use different definitions of arguments, they all have at their heart an exchange of offers and arguments. However, none of those proposals explain when arguments can be used within a negotiation, and how they should be dealt with by the agent that receives them. Thus the protocol for handling arguments was missing. Another limitation of the above frameworks is the fact that the argumentation frameworks they use are quite poor, since they use a very simple acceptability semantics. In [1] a negotiation framework that fills the gap has been suggested. A protocol that handles the arguments was proposed. However, the notion of concession is not modeled in that framework, and it is not clear what is the status of the outcome of the dialogue. Moreover, it is not clear how an agent chooses the offer to propose at a given step of the dialogue. In [2, 6], the authors have focused mainly on this decision problem. They have proposed an argumentation-based decision framework that is used by agents in order to choose the offer to propose or to accept during the dialogue. In that work, agents are supposed to have a beliefs base and a goals base.

Our framework is more general since it does not impose any specific structure for the arguments, the offers, or the beliefs. The negotiation protocol is general as well. Thus this framework can be instantiated in different ways by creating, in such manner, different specific argumentation-based negotiation frameworks, all of them respecting the same properties. Our framework is also a unified one because frameworks like the ones presented above can be represented within this framework. For example the decision making mechanism proposed in [6] for the evaluation of arguments and therefore of offers, which is based on a priority relation between mutually attacked arguments, can be captured by the relation *defeat* proposed in our framework. This relation takes simultaneously into account the attacking and preference relations that may exist between two arguments.

7 Conclusions and Future Work

In this paper we have presented a unified and general framework for argumentation-based negotiation. Like any other argumentation-based negotiation framework, as it is evoked in (e.g. [9]), our framework has all the advantages that argumentation-based negotiation approaches present when related to the negotiation approaches based either on game

theoretic models (see e.g. [10]) or heuristics ([5]). This work is a first attempt to formally define the role of argumentation in the negotiation process. More precisely, for the first time, it formally establishes the link that exists between the status of the arguments and the offers they support, it defines the notion of concession and shows how it influences the evolution of the negotiation, it determines how the theories of agents evolve during the dialogue and performs an analysis of the negotiation outcomes. It is also the first time where a study of the formal properties of the negotiation theories of the agents as well as of an argumentative negotiation dialogue is presented.

Our future work concerns several points. A first point is to relax the assumption that the set of possible offers is the same to both agents. Indeed, it is more natural to assume that agents may have different sets of offers. During a negotiation dialogue, these sets will evolve. Arguments in favor of the new offers may be built from the agent theory. Thus, the set of offers will be part of the agent theory. Another possible extension of this work would be to allow agents to handle both arguments PRO and CONS offers. This is more akin to the way human take decisions. Considering both types of arguments will refine the evaluation of the offers status. In the proposed model, a preference relation between offers is defined on the basis of the partition of the set of offers. This preference relation can be refined. For instance, among the acceptable offers, one may prefer the offer that is supported by the strongest argument. In [3], different criteria have been proposed for comparing decisions. Our framework can thus be extended by integrating those criteria. Another interesting point to investigate is that of considering negotiation dialogues between two agents with different profiles. By profile, we mean the criterion used by an agent to compare its offers.

References

[1] L. Amgoud, S. Parsons, and N. Maudet. Arguments, dialogue, and negotiation. In *Proc. of the 14th ECAI*, 2000.

[2] L. Amgoud and H. Prade. Reaching agreement through argumentation: A possibilistic approach. In *Proc. of the 9 th KR*, 2004.

[3] L. Amgoud and H. Prade. Explaining qualitative decision under uncertainty by argumentation. In *Proc. of the 21st AAAI*, pages 16–20, 2006.

[4] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[5] N. R. Jennings, P. Faratin, A. R. Lumuscio, S. Parsons, and C. Sierra. *Automated negotiation: Prospects, methods and challenges*. International Journal of Group Decision and Negotiation, 2001.

[6] A. Kakas and P. Moraitis. Adaptive agent negotiation via argumentation. In *Proc. of the 5th AAMAS*, pages 384–391, 2006.

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

[8] S. Parsons and N. R. Jennings. Negotiation through argumentation—a preliminary report. In *Proc. of the 2nd ICMAS*, pages 267–274, 1996.

[9] I. Rahwan, S. D. Ramchurn, N. R. Jennings, P. McBurney, S. Parsons, and E. Sonenberg. Argumentation-based negotiation. *Knowledge Engineering Review*, 18 (4):343–375, 2003.

[10] J. Rosenschein and G. Zlotkin. *Rules of Encounter: Designing Conventions for Automated Negotiation Among Computers,*. MIT Press, Cambridge, Massachusetts, 1994.

[11] K. Sycara. Persuasive argumentation in negotiation. *Theory and Decision*, 28:203–242, 1990.