"Minimal defence": a refinement of the preferred semantics for argumentation frameworks

C. Cayrol
IRIT - UPS
118, rte de Narbonne
31062 Toulouse, France
{ccayrol@irit.fr}

S. Doutre
IRIT - UPS
118, rte de Narbonne
31062 Toulouse, France
{doutre@irit.fr}

M.-C. Lagasquie-Schiex
IRIT - UPS
118, rte de Narbonne
31062 Toulouse, France
{lagasq@irit.fr}

J. Mengin
IRIT - UPS
118, rte de Narbonne
31062 Toulouse, France
{mengin@irit.fr}

Abstract

Dung's abstract framework for argumentation enables a study of the interactions between arguments based solely on an "attack" binary relation on the set of arguments. Various ways to solve conflicts between contradictory pieces of information have been proposed in the context of argumentation, non-monotonic reasoning or logic programming, and can be captured by appropriate semantics within Dung's framework. A common feature of these semantics is that one can always maximize in some sense the set of acceptable arguments. We propose in this paper to extend Dung's framework in order to allow for the representation of what we call "restricted" arguments: these arguments should only be used if absolutely necessary, that is, in order to support other arguments that would otherwise be defeated. We modify Dung's preferred semantics accordingly: a set of arguments becomes acceptable only if it contains a minimum of restricted arguments, for a maximum of unrestricted arguments.

1 INTRODUCTION

Argumentation is a form of reasoning that has recently captured the interest of many researchers in the Artificial Intelligence community, among others. An abstract framework for studying it has been proposed by Dung [Dung, 1995]. It is based on the assumption that the process of argumentation is built from two types of jobs: the formation of arguments, associated with the interactions that exist between them; and the evaluation of the validity of the arguments. A form of interaction between arguments that is often difficult to ignore is contradiction, since a usual basic measure of the validity of a piece of reasoning is the level of self-contradiction it contains. In Dung's approach, an argumentation framework consists of two elements: a set of arguments, and a binary relation between arguments, so that one can represent an "attack", or contradiction, relation between arguments. Thus, the set of arguments and the attack relation form a graph.

If one accepts that the graph contains all needed information, it is possible to define which arguments are valid, or acceptable, by considering solely this graph. Our intuition tells us that an argument which is not attacked at all should be acceptable. More generally, one can define that an acceptable argument is one which is defended, that is, whose attackers are all attacked by arguments which are themselves acceptable.

When the graph contains cycles, this definition is circular, and one needs a more global approach. A solution is to define acceptable sets of arguments. Arguments in such a set do usually not attack each other, and defend each other: every attacker of an argument in the set is itself attacked by an argument in the set. This definition leads to Dung's "admissible" sets of arguments. Maximal sets of arguments that are acceptable are then called "preferred extensions" of the framework.

These definitions of acceptability make no distinction between arguments, in the sense that defenders are all equally desirable: if an argument can be independently defended by two different arguments, then one may choose any of the two, or even the two arguments together if they do not attack each other, to defend the first one. However, there are problems where a distinction between several levels of desirability of defenders is needed. Suppose for example that our arguments "belong" to two agents: one, the proponent, is trying to make a point, while the other one, the opponent, is trying to refute it. The proponent may have

arguments that should only be disclosed if absolutely necessary (like very personal information).

We propose below to refine Dung's argumentation framework with a distinction between two types of arguments: we consider *unrestricted* arguments, that one can maximize, and *restricted* arguments, the use of which will be kept to a minimum. Restricted arguments should only be used if absolutely necessary, that is, in order to support other arguments that would otherwise be defeated.

Notice that the distinction between restricted and unrestricted arguments is at a defensive level. Other frameworks have been designed in order to authorize varying strengths for the arguments. In the frameworks described in [Amgoud and Cayrol, 1998] and [Prakken and Sartor, 1997], a preference relation among arguments is taken into account when the interactions between arguments are studied: the idea is that in order to be a serious menace, an argument must be preferred to the argument it attacks. In this case, there is no need to minimize the number of arguments of a given strength, except maybe for a purpose of conciseness or of efficiency.

We also propose to allow for a distinction between the arguments of the proponent and those of the opponent: this enables us to ensure, when this is needed, that the arguments of the opponent are never used to defend an argument of the proponent. Note that this distinction is independent from the previous one, between restricted and unrestricted arguments.

The paper is built as follows: our refined argumentation framework is presented in the next section. In particular, we detail the partition of the set of arguments, and provide a new definition for admissibility; we then define a new type of extensions for our refined argumentation framework. We illustrate our definitions with three particular argumentation frameworks in section 3. Finally, we conclude and sketch future works in section 4.

2 FORMALIZATION

We recall the definition of [Dung, 1995]'s argumentation framework.

Definition 1 *[Dung, 1995] An* argumentation framework *is a pair $\mathcal{AF} = (A, R)$ where A is a set of arguments and R is an attack relation between arguments ($R \subseteq A \times A$).*

An argumentation framework is well-founded *if and only if there is no infinite sequence*

$$a_0,\ a_1,\ \ldots,\ a_n,\ \ldots$$

such that $\forall i, a_i \in A$ and $a_{i+1} R a_i$[1].

An argumentation framework can easily be represented as a directed graph, where vertices are the arguments and edges correspond to the elements of the relation R.

Example 1 *Let $\mathcal{AF}_1 = (A, R)$ be an argumentation framework such that:*

$A = \{u_1, u_2, u_3, u_4, u_5, r_1, r_2, r_3, r_4, o_1, o_2, o_3, o_4, o_5\}$ *and*

$R = \{(o_1, u_1), (o_2, u_2), (u_3, o_2), (o_3, u_4), (u_5, o_4), (r_1, o_2), (r_2, o_3), (o_4, r_3), (o_5, r_4)\}$.

The graph representation of $\mathcal{AF}_1$ is depicted on figure 1.

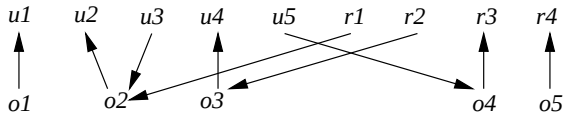

Figure 1: Graph representation of the argumentation framework $\mathcal{AF}_1$

Partitions of the set of arguments of an argumentation framework Let $\mathcal{AF} = (A, R)$ be an argumentation framework. In order to refine some existing semantics for argumentation and considering that A represents the arguments of different agents, the meaning of a partition of A may be the following:

- A simple case is to take into account one agent against the others. So we have a simple partition of A, called `type-1-partition`: $A = F \cup (A \setminus F)$ for a given set of arguments F (corresponding to the arguments of the selected agent);

  the argumentation framework $\mathcal{AF}$ `type-1-` partitioned, given a set of arguments F, is denoted by $\mathcal{AF}_F$;

- the other type of a partition corresponds to the distinction between restricted arguments (which should not be used if possible) and unrestricted arguments. So, we have a special partition of A, called `type-2-partition`, which is a refinement of a `type-1-partition`: given a set F such that $F = F_u \cup F_r$ with $F_u \cap F_r = \varnothing$, $A = F_u \cup F_r \cup (A \setminus (F_u \cup F_r))$ (F_u, F_r are respectively the set

[1] If R is finite, an argumentation framework is well-founded if and only if it does not contain any cycle.

of the unrestricted arguments and the set of the restricted arguments of the selected agent);

the argumentation framework $\mathcal{AF}$ `type-2-` partitioned, is denoted by $\mathcal{AF}_{F_{ru}}$.

Note that, although every `type-2-partition` has an underlying `type-1-partition`, the latter can be a vacuous one: it is possible to set F to be the entire set A.

2.1 ADMISSIBILITY

Let $\mathcal{AF} = (A, R)$ be an argumentation framework, we have the following definitions.

Definition 2 *[Dung, 1995] A set $S \subseteq A$ is* conflict-free *if and only if $\nexists a, b \in S$ such that aRb.*

[Dung, 1995] defines the notion of defence. Defence can be of two kinds: collective or individual. Collective defence is achieved by a set of arguments, individual defence by one argument.

Definition 3 *[Dung, 1995] Let $S \subseteq A$, $a \in A$. S* defends (collectively) *a if and only if $\forall b \in A$, if $bRa, \exists c \in S$ such that cRb. S* defends all its elements *if and only if $\forall a \in S$, if $\exists b \in A$ such that bRa then $\exists c \in S$ such that cRb.*

The following definition is inspired by [Dung, 1995]:

Definition 4 *Let $a, c \in A$. c* is an individual defender of *a if and only if there is a finite sequence $x_0, \ldots, x_{2n}$ such that:*

1. *$a = x_0$ and $c = x_{2n}$, and*
2. *$\forall i, 0 \leq i \leq (2n-1), x_{i+1} R x_i$.*

If $n = 1$, c is a direct individual defender *of a.*

Then, [Dung, 1995] proposes the admissible semantics:

Definition 5 *[Dung, 1995] A set $S \subseteq A$ is* admissible *if and only if S is conflict-free and S defends all its elements.*

Every argumentation framework has at least one admissible set because the empty set is admissible.

Example 2 *The sets $\emptyset$, $\{o_1\}$, $\{o_1, u_2, u_3\}$, $\{o_1, u_2, u_3, u_4, u_5, r_1, r_2, r_3, o_5\}$ are admissible sets of the argumentation framework $\mathcal{AF}_1$ represented on figure 1.*

We refine this admissible semantics in the context of a `type-2-partition` of A in order to accept restricted arguments only if their presence is justified, that is when they defend at least one unrestricted argument. This refinement is called the *restricted admissibility*.

Let $\mathcal{AF}_{F_{ru}}$ be a `type-2-`partitioned argumentation framework, we have the following notations and definition:

Notations Given $F \subseteq A$ such that $F = F_u \cup F_r$ and $F_u \cap F_r = \varnothing$ and a set $S \subseteq A$, S_u denotes $S \cap F_u$ and S_r denotes $S \cap F_r$. We say that S_u is the *unrestricted* part of S, and S_r is the *restricted* part of S.

Definition 6 *Let $S \subseteq F$. S is* restrictedly admissible *if and only if:*

1. *S is admissible and*
2. *$\forall x \in S_r$, $\exists y \in S_u$ such that x is an individual defender of y.*

Example 3 *The set of arguments of the argumentation framework $\mathcal{AF}_1$ depicted on figure 1, can be partitioned as follows: let $F = F_u \cup F_r$ with $F_u = \{u_1, u_2, u_3, u_4, u_5\}$, $F_r = \{r_1, r_2, r_3, r_4\}$ and $A \setminus F = \{o_1, o_2, o_3, o_4, o_5\}$. The resulting* `type-1-` *(resp.* `type-2-`*) partitioned argumentation framework is called $\mathcal{AF}_2$ (resp. $\mathcal{AF}_3$). $\mathcal{AF}_2$ and $\mathcal{AF}_3$ are depicted on figure 2.*

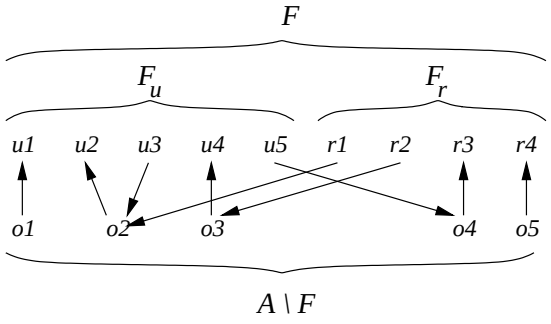

Figure 2: Argumentation frameworks $\mathcal{AF}_2$ and $\mathcal{AF}_3$

$\mathcal{AF}_3$ has several restrictedly admissible sets. For example:
$\emptyset$, $\{u_2, u_3\}$, $\{u_2, u_3, u_4, u_5, r_2\}$, $\{u_2, u_3, u_4, u_5, r_1, r_2\}$.

The set $S = \{u_5, r_3\}$ is not restrictedly admissible since r_3 does not defend any argument of the unrestricted part of S (that is $\{u_5\}$).

2.2 EXTENSIONS

Let $\mathcal{AF} = (A, R)$ be an argumentation framework. [Dung, 1995] has also defined the preferred semantics in which an acceptable set of arguments is called a *preferred extension.*

Definition 7 *[Dung, 1995] A set $S \subseteq A$ is a* preferred extension *of $\mathcal{AF}$ if and only if S is maximal for set-inclusion among the admissible sets of $\mathcal{AF}$.*

We recall:

Property 1 *[Dung, 1995]*

1. *Every admissible set of $\mathcal{AF}$ is included in a preferred extension of $\mathcal{AF}$.*
2. *Every argumentation framework has at least one preferred extension.*
3. *Every well-founded argumentation framework has one and only one preferred extension.*

Example 4 *$\mathcal{AF}_1$, $\mathcal{AF}_2$ and $\mathcal{AF}_3$ have only one preferred extension:*

$$\{o_1, u_2, u_3, u_4, u_5, r_1, r_2, r_3, o_5\}.$$

In order to privilege some arguments of A, and to facilitate the computation of the extensions under the new semantics, we define a *preferred extension on* a given set X. Note that it implies the use of a `type-1-partition` of A: $A = X \cup (A \setminus X)$.

Definition 8 *Let $\mathcal{AF} = (A, R)$ be an argumentation framework and let $X \subseteq A$. A* preferred extension on *X of $\mathcal{AF}$ is an admissible set $S \subseteq X$ such that S is $\subseteq$-maximal on X, i.e. $\forall Y \subseteq (X \setminus S)$, $Y \cup S$ is not admissible.*

Note that :

- The preferred extensions on A in the sense of definition 8 are the preferred extensions in the sense of definition 7.
- A preferred extension on X is not simply the intersection of X and of a preferred extension of $\mathcal{AF}$. For example, let $\mathcal{AF} = (\{a, b, c\}, \{(c, b), (b, a)\})$ be an argumentation framework and $X = \{a\}$, the preferred extension of $\mathcal{AF}$ is $\{c, a\}$, its intersection with X is $\{a\}$ and the only preferred extension on X is $\varnothing$!

The following results hold:

Property 2 *Let $\mathcal{AF} = (A, R)$ be an argumentation framework and let $X \subseteq A$.*

1. *Every admissible set included in X is included in a preferred extension on X.*
2. *Every preferred extension on X is included in a preferred extension of $\mathcal{AF}$.*
3. *Let (X, R') be the restriction of $\mathcal{AF}$ to X[2]. If (X, R') is a well-founded argumentation framework then $\mathcal{AF}$ has one and only one preferred extension on X.*

The proof of this property (and the detailed proofs of all the properties of this paper, except the property 1 which is given in [Dung, 1995]) is given in [Cayrol et al., 2002].

Example 5 *$\mathcal{AF}_2$ has one preferred extension on F:* $\{u_2, u_3, u_4, u_5, r_1, r_2, r_3\}$.

Preferred extensions are interesting in several respects:

- they provide a good "summary" of the admissible sets, since every subset of a preferred extension can be completed to an admissible set, and every admissible set is subset of at least one preferred extension;
- the intersection of the preferred extensions can be interpreted as being the set of arguments that cannot be defeated.

Similar remarks can be made for other types of extensions of argumentation frameworks, like stable extensions for example. Similarly, we want to define a family of "best" restrictedly admissible sets. In the context of a `type-2-`partitioned argumentation framework, the idea is to select the restrictedly admissible sets which contain a minimum of restricted arguments for a maximum of unrestricted arguments. In order to do so, we define the following relation:

Definition 9 *Let $\mathcal{AF}_{F_{do}}$ be a* `type-2-`*partitioned argumentation framework. Let S_1 and S_2 be two subsets of F. S_2 is $\prec$-*better than *S_1 (denoted by $S_1 \prec S_2$) if and only if $S_{1u} \subset S_{2u}$, or $S_{1u} = S_{2u}$ and $S_{2r} \subseteq S_{1r}$.*

The relation $\prec$ is a partial order on the set of the subsets of F, which is clearly different from set-inclusion. The meaning of this relation is the following: $S_1 \prec S_2$ if and only if S_2 contains more unrestricted arguments

[2] $R' = \{(a, b) | aRb$ and $a \in X, b \in X\}$.

than S_1, or S_1 and S_2 have the same unrestricted arguments but S_2 contains less restricted arguments than S_1.

The idea of the new semantics is to have as few restricted arguments as possible in a restrictedly admissible set. In other words, the set of restricted defenders must be minimal. That is why we call the new semantics the *minimal defence semantics*, or, for short, the *min-def* semantics. The acceptable sets under this semantics are called *min-def extensions*.

Definition 10 *S is a* min-def extension *of $\mathcal{AF}_{F_{ru}}$ if and only if S is $\prec$-maximal among the restrictedly admissible sets of $\mathcal{AF}_{F_{ru}}$.*

Note that the restricted arguments in an acceptable set under the new semantics are such that, not only they defend an unrestricted argument x, but they are essential defenders since there is no unrestricted argument which defends x. Remark also that not every preferred extension is a min-def extension.

Example 6 *$\mathcal{AF}_3$ has only one min-def extension: $S = \{u_2, u_3, u_4, u_5, r_2\}$ (which is different of its preferred extension). S does not contain the restricted argument r_1 whereas r_1 defends u_2 against o_2. The reason is that the unrestricted argument u_3 defends u_2, therefore r_1 is not useful.*

We have proved a number of properties. The first one establishes the link between the admissibility and the min-def semantics (note that the admissible sets defined in $\mathcal{AF}$ are the sames in $\mathcal{AF}_{F_{ru}}$):

Property 3 *Let $S \subseteq F$. S is a min-def extension of $\mathcal{AF}_{F_{ru}}$ if and only if S is $\prec$-maximal among the admissible sets of $\mathcal{AF}$.*

Proof (Sketch) *It is a consequence of the fact that every admissible set which is $\prec$-maximal is restrictedly admissible.* $\diamondsuit$

The following property shows that every finite argumentation framework has at least one min-def extension:

Property 4 *Suppose that A is finite. Then for every admissible set G included in F, there is a min-def extension S of $\mathcal{AF}_{F_{ru}}$ such that $G \prec S$.*

Proof (Sketch) *Given Adm_F the set of the admissible sets of $\mathcal{AF}$ included in F and the relation $\prec$, we prove that every chain in Adm_F has an upper bound in Adm_F.* $\diamondsuit$

The next two properties explain the link between preferred extensions on a given set and min-def extensions. This link is very interesting in a computational perspective (see section 4).

Property 5 *If S is a min-def extension of $\mathcal{AF}_{F_{ru}}$, then there is a preferred extension E on F such that $E_u = S_u$ and $E_r \supseteq S_r$, so that $E \prec S$.*

Proof (Sketch) *The proof relies upon property 3 and the following result:*

If G is a $\prec$-maximal admissible set included in F and if H is an admissible set which contains G, then $H_u = G_u$.

Since every admissible set included in F is included in a preferred extension on F (see property 2), there exists a preferred extension E on F such that $G \subseteq E$ and then $G_u = E_u$. $\diamondsuit$

Property 6 *Let E be a preferred extension on F. There is a min-def extension S of $\mathcal{AF}_{F_{ru}}$ such that $E_u \subseteq S_u$.*

Proof (Sketch) *We show that if G is an admissible set included in F, then there exists an admissible set H included in F maximal for the relation $\prec$ such that $G_u \subseteq H_u$. The property is a consequence of this result and of property 3.* $\diamondsuit$

It can be noticed that the preferred extension on F of example 5 is such that its unrestricted part $\{u_2, u_3, u_4, u_5\}$ is equal to the unrestricted part of the min-def extension of $\mathcal{AF}_3$, and its restricted part $\{r_1, r_2, r_3\}$ contains the restricted part $\{r_2\}$ of the min-def extension of $\mathcal{AF}_3$.

3 ILLUSTRATIONS

Let us give three illustrations of the formal framework presented in section 2. The first two are examples in a dialogical context, and the last one shows that the framework can be applied to other contexts.

3.1 A DISCUSSION BETWEEN FRIENDS

Denis and Theo have a discussion. Denis has a point of view and gives arguments to support it. Denis has a friend, Olivia, who agrees with him. Theo does not agree with Denis and Olivia's point of view and gives arguments attacking their arguments. In this discussion, Denis can be viewed as the proponent and Theo as the opponent. Denis has several solutions to defend his arguments against Theo's attacks:

1. either Denis uses his own arguments only;

2. or Denis also accepts Olivia's arguments and considers that Olivia's arguments have the same importance as his: he uses indiscriminately his arguments or hers to defend his arguments;

3. or Denis accepts Olivia's arguments but he wants to privilege his own arguments: he takes into account Olivia's arguments only when they defend some of his arguments which he cannot defend with his own arguments.

The different ways Denis can accept Olivia's arguments are justified by the fact that Olivia's argument could contain, for example, very personal information on Denis. Therefore, one can imagine that if Denis does not really want these informations to be revealed, he accepts Olivia's arguments only if absolutely necessary (case 3) or he does not accept them at all (case 1). On the contrary, if Denis agrees to reveal these informations, he makes no distinction between Olivia's arguments and his (case 2).

The arguments and the attacks of the discussion constitute an argumentation framework where Denis' arguments are the unrestricted arguments, Olivia's arguments are the restricted arguments and Theo's arguments are the other arguments.

We assume that the acceptable sets of arguments supporting Denis' point of view are conflict-free sets which defend all their arguments against Theo's attacks and contain a maximum of Denis' arguments. Thus, in case 1, these sets are preferred extensions on Denis' arguments; in case 2, they are restrictedly admissible sets; in case 3, they are min-def extensions.

In a multi-agent framework, it is usually the case that the agents need to communicate to fulfil tasks or to share resources, so that they have to engage in dialogues. The works of [Parsons et al., 1998], [Reed, 1998] and [Amgoud et al., 2000] outline that argumentation can be used as a basis for dialogues. [Amgoud et al., 2000] studies different types of dialogue in argumentation theory, and mentions the strategies that an agent uses to choose an argument during a dialogue. This point is more precisely investigated in [Amgoud and Maudet, 2002]. The example above suggests that our work could be used as a support to implement the strategies. The unrestricted arguments of an agent would be the ones which satisfy the strategy, the restricted arguments would be the other arguments of the agent.

3.2 MAFIA ON TRIAL

Consider a criminal court case with the mafia on trial, and assume that some piece of witness testimony is very convincing evidence against the suspects, but that the prosecution still hesitates to use it, since doing so would inevitably disclose the witnesses' identity, and the mafia gang is known to be very violent against people who testify against them. In such a case, the prosecutor will try to avoid using this piece of evidence as long as possible, and try to win the case in other ways. Only if these other ways fail, the prosecutor may decide to use the witness testimony.

The parallel between this criminal court case and a `type-2-`argumentation framework is easy to do: the pieces of evidence which can be disclosed are the unrestricted arguments, and the witness testimony is a restricted argument. [3]

The set of arguments advanced by the prosecution is a conflict-free set, which defends all its elements and which contains the restricted argument only if it defends a piece of evidence which could not be defended otherwise. One can imagine that the set of pieces of evidence which can be disclosed is maximal, since the more pieces of evidence are justified, the better the suspects' guilt is ensured. Therefore, the set of arguments advanced by the prosecution is a min-def extension.

3.3 CONFIGURATION OF A WARPLANE

The following illustration shows that our work can be applied to non dialogical contexts, to a configuration problem for instance.

Assume that the army wants to configure its warplanes in order for them to defend against any attack of the enemy during the next mission. Among the functions which can be put on a plane, some are desirable and even essential, like flight, propulsion and communication functions; others are optional, like the jammers of radar. A maximum of desirable functions must be put on a plane, whereas a minimum of optional functions must be installed because of cost, speed and weight constraints. The enemy has materials which threaten some functions; for instance, their jammer of flight threatens the flight function.

A parallel can easily be established between this problem of configuration and a `type-2-`partitioned argumentation framework: we set A to be the set of functions and threats; F is then the set of functions, and

[3] It is this example, suggested by the reviewers, which prompted to us the expression "restricted argument", in the sense of "confidential argument".

F_u is the set of desirable functions whereas F_r is the set of optional functions; the attack relation R corresponds to the conflicts between functions and threats.

A good configuration of a plane is defined as a set of functions C such that:

1. C is conflict-free;
2. if a function in C is threatened, then C contains a function which attacks the threat;
3. C contains a maximum of desirable functions and a minimum of optional functions.

The parallel established with a `type-2`-partitioned argumentation framework shows that a good configuration is a set included in F which is conflict-free (1) and which defends all its arguments (2); therefore, it is an admissible set included in F. This set contains a maximum of unrestricted arguments (the desirable functions), and a minimum of restricted arguments (the optional functions) (3); in other words, this admissible set is $\prec$-maximal. Consequently, according to property 3, a good configuration is a min-def extension.

4 CONCLUSION AND FUTURE WORKS

In this paper, using a partition of the set of arguments, we propose some refinements of the argumentation framework defined by [Dung, 1995]. The aim of this partition is to privilege one subset of arguments and, in addition, to make a distinction between restricted and unrestricted arguments. In this context, we have defined different semantics and the associated acceptable sets:

- the *restrictedly admissible sets* are admissible sets of arguments in which the presence of each restricted argument is justified (a restricted argument belongs to a restrictedly admissible set only if it defends an unrestricted argument);
- the *min-def extensions* are restrictedly admissible sets which respect the following condition: they maximize the unrestricted arguments and minimize the restricted arguments.

Different properties are presented which show the relations between the different semantics.

Some of these results (reported in section 2) suggest that the min-def extensions can be computed from the set of the preferred extensions on F.

Actually, from property 5, it is clear that each min-def extension can be obtained from one of the preferred extensions on F, by minimizing its intersection with F_r (its restricted part). It follows from property 6 that it is sufficient to consider the preferred extensions on F such that their intersection with F_u is maximal (for set-inclusion).

[Doutre and Mengin, 2001] proposes an algorithm based on the technique of set-enumeration for computing all the preferred extensions of a given argumentation framework. The basic idea is to reduce the number of generated sets of arguments using properties of the preferred extensions. That algorithm can be adapted using the following additional properties: we are interested in subsets of F, which are preferred extensions on F and whose unrestricted part (intersection with F_u) is maximal for set-inclusion.

So, we propose a two-steps method for computing min-def extensions:

1. Compute X_1, X_2, ..., X_n the preferred extensions on F with a maximal unrestricted part.
2. From each X_i, compute the admissible sets by removing as many elements of X_{ir} (*i.e.* restricted elements) as possible.

The above method is correct and complete (see [Cayrol et al., 2002]).

In the particular case where there is no cycle in F, the first step in the above algorithm produces only one preferred extension on F.

We consider an extension of this model to the case where the set of restricted arguments would be stratified. This is motivated by several examples. For instance, among the optional functions which can be put on a plane, some can be more costly than others; one wants to choose the least expensive in priority. Moreover, we have seen that the restricted arguments of an agent can be those which contain personal information. These informations can be more or less personal, and then one prefers to disclose the arguments which contain the least personal information in priority. Consequently, we want to extend a `type-2`-partitioned argumentation framework with a priority relation over the restricted arguments. We will integrate this priority in the definition of the $\prec$-relation and of the min-def semantics in order to choose the restricted arguments which have the greatest priority (*i.e.* those which contain the least expensive functions or the least personal information) for the defence of the unrestricted arguments.

Acknowledgements

We would like to thank Hélène Fargier for fruitful preliminary discussions on this topic. We would also like to thank the referees for helpful comments; in particular, they suggested the "mafia example".

References

[Amgoud and Cayrol, 1998] Amgoud, L. and Cayrol, C. (1998). On the acceptability of arguments in preference-based argumentation. In *14th Conference on Uncertainty in Artificial Intelligence (UAI'98), Madison, Wisconsin*, pages 1–7, San Francisco, California. Morgan Kaufmann.

[Amgoud and Maudet, 2002] Amgoud, L. and Maudet, N. (2002). Strategical considerations for argumentative agents (preliminary report). In *NMR'2002 special session on Argument, Dialogue and Decision.*

[Amgoud et al., 2000] Amgoud, L., Maudet, N., and Parsons, S. (2000). Modelling dialogues using argumentation. In *Fourth International Conference on MultiAgent Systems (ICMAS'2000), Boston, MA, USA*, pages 31–38.

[Cayrol et al., 2002] Cayrol, C., Doutre, S., Lagasquie-Schiex, M., and Mengin, J. (2002). Défense minimale : un raffinement de la sémantique préférée pour les systèmes d'argumentation. Rapport de recherche 2002-03-R, Institut de Recherche en Informatique de Toulouse (I.R.I.T.), France.

[Doutre and Mengin, 2001] Doutre, S. and Mengin, J. (2001). Preferred Extensions of Argumentation Frameworks: Computation and Query Answering. In R. Goré, A. L. and Nipkow, T., editors, *IJCAR 2001*, volume 2083 of *LNAI*, pages 272–288. Springer-Verlag.

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

[Parsons et al., 1998] Parsons, S., Sierra, C., and Jennings, N. R. (1998). Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261–292.

[Prakken and Sartor, 1997] Prakken, H. and Sartor, G. (1997). Argument-based logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, 7:25–75.

[Reed, 1998] Reed, C. (1998). Dialogue frames in agent communication. In *Proceedings of the 3rd International Conference on Multi Agent Systems*, pages 246–253.