Theoretical and Computational Properties of Preference-based Argumentation

Yannis Dimopoulos[1] and **Pavlos Moraitis**[2] and **Leila Amgoud**[3]

Abstract.

During the last years, argumentation has been gaining increasing interest in modeling different reasoning tasks of an agent. Many recent works have acknowledged the importance of incorporating preferences or priorities in argumentation. However, relatively little is known about the theoretical and computational implications of preferences in argumentation.

In this paper we introduce and study an abstract preference-based argumentation framework that extends Dung's formalism by imposing a preference relation over the arguments. Under some reasonable assumptions about the preference relation, we show that the new framework enjoys desirable properties, such as coherence. We also present theoretical results that shed some light on the role that preferences play in argumentation. Moreover, we show that although some reasoning problems are intractable in the new framework, it appears that the preference relation has a positive impact on the complexity of reasoning.

1 Introduction

Argumentation has become an Artificial Intelligence keyword for the last fifteen years, especially in sub-fields such as non monotonic reasoning [8] and agent technology (e.g. [4]). Argumentation is a promising reasoning model based on the interaction of different arguments for and against some statement. This interaction between arguments is typically based on a notion of *attack*, which can take different forms according to the form that the arguments have. For example, when an argument takes the form of a logical proof, arguments for and against a statement can be put across and in this case the attack relation expresses logical inconsistency. Argumentation can therefore be considered as a reasoning process implying construction and evaluation of interacting arguments.

Several interesting argumentation frameworks have been proposed in the literature (see e.g. [3, 14, 12]). The majority of these systems is based on the abstract argumentation framework of Dung [8], where no assumption is made about the nature of arguments or the properties of the attack relation (i.e. the attack relation can be any binary relation on the set of arguments).

Some recent works have proposed argumentation systems (see e.g. [2, 1, 5]) that are based on a *defeat* relation (corresponding to the attack relation in Dung's framework), that is composed from a *conflict relation* on the set of arguments and a *preference* relation between arguments, reflecting the fact that arguments may not have equal strengths. However till now, relatively little is known about the theoretical and computational properties of abstract preference-based argumentation systems.

This paper is an attempt towards understanding the effects of a preference relation on an argumentation system. More precisely, it investigates the impact of the preference relation between arguments within a new abstract argumentation framework. The attack relation is the composition of a conflict relation with the preference relation, both defined on the set of arguments. The framework is abstract and general in the sense that the only assumptions made are that the *conflict relation* is *symmetric* and *irreflexive*, and the preference relation is a *partial pre-order* (i.e. *reflexive* and *transitive*). Under these reasonable and general assumptions, we show that the new framework enjoys desirable properties for an argumentation system, such as *coherence*. It turns out that the preference relation on the arguments translates into a preference relation on the powerset of these arguments. Moreover, the stable extensions of the preference-based argumentation theories correspond to the most preferred sets of arguments that are conflict-free.

We also investigate the computational properties of the new framework and demonstrate that a transitive preference relation on the set of arguments can mitigate the computational burden of some reasoning tasks. Indeed, computing a stable extension of a preference-based argumentation theory can be performed in polynomial time. Furthermore, enumerating all stable extensions of such a theory without incomparability between arguments can be carried out with *polynomial delay*. Moreover, if in addition the theory does not contain indifferent arguments, finding its unique stable extension is also a polynomial computation. On the negative side, some other reasoning tasks are intractable. More specifically, deciding whether an argument is a *credulous conclusion* of a preference-based argumentation theory is *NP-hard*, while deciding whether it is a *skeptical* one is *coNP-hard*.

The paper is organized as follows. We first review the basics of argumentation as introduced in [8]. Then, we present the abstract preference-based argumentation framework we propose, and investigate some of its properties. We then present algorithms for reasoning in the new framework, along with some complexity results. The last section concludes with some remarks and perspectives.

2 Basics of argumentation

Argumentation is a reasoning model based on the following main steps: i) constructing *arguments* and counter-arguments, ii) defining the *strengths* of those arguments, and iii) concluding or defining the *justified conclusions*. Argumentation systems are built around an underlying logical language and an associated notion of logical consequence, defining the notion of argument. The argument construction is a monotonic process: new knowledge cannot rule out an argument

[1] University of Cyprus, 75 Kallipoleos Str. 1678, Nicosia-Cyprus
[2] Paris Descartes University, 45 rue des Saints-Pères 75270, Paris-France
[3] Paul Sabatier University, 118 route de Narbonne 31062, Toulouse-France

but only gives rise to new arguments which may interact with the first argument. Arguments may be conflicting for different reasons.

Definition 1 (Argumentation system [8]) *An* argumentation system *is a pair T = $(\mathcal{A}, \mathcal{R})$. $\mathcal{A}$ is a set of arguments and $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is an attack relation. We say that an argument a attacks an argument b iff $(a, b) \in \mathcal{R}$.*

Among all the arguments, it is important to know which arguments to keep for inferring conclusions. In [8], different acceptability semantics have been proposed. The basic idea behind these semantics is the following: for a rational agent, an argument a_i is acceptable if he can defend a_i against all attacks. All the arguments acceptable for a rational agent will be gathered in a so-called *extension*. An extension must satisfy a consistency requirement and must defend all its elements.

Definition 2 (Conflict-free, Defence [8]) *Let $\mathcal{B} \subseteq \mathcal{A}$, and $a_i \in \mathcal{A}$.*

- *$\mathcal{B}$ is* conflict-free *iff $\nexists\ a_i, a_j \in \mathcal{B}$ s.t. $(a_i, a_j) \in \mathcal{R}$.*
- *$\mathcal{B}$* defends *a_i iff $\forall\ a_j \in \mathcal{A}$, if $(a_j, a_i) \in \mathcal{R}$, then $\exists\ a_k \in \mathcal{B}$ s.t. $(a_k, a_j) \in \mathcal{R}$.*

The main semantics introduced by Dung are summarized in the following definition.

Definition 3 (Acceptability semantics [8]) *Let $\mathcal{B}$ be a conflict-free set of arguments.*

- *$\mathcal{B}$ is* admissible *iff it defends any argument in $\mathcal{B}$.*
- *$\mathcal{B}$ is a* preferred *extension iff it is a maximal (w.r.t $\subseteq$) admissible extension.*
- *$\mathcal{B}$ is a* stable *extension iff it is a preferred extension that attacks any argument in $\mathcal{A} \backslash \mathcal{B}$.*

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

Definition 4 (Argument status) *Let $T = (\mathcal{A}, \mathcal{R})$ be an argumentation system, and $\mathcal{E}_1, \ldots, \mathcal{E}_x$ its stable extensions. Let $a \in \mathcal{A}$.*

- *a is* skeptical conclusion *of T iff $a \in \mathcal{E}_i$, $\forall \mathcal{E}_{i=1,\ldots,x} \neq \emptyset$.*
- *a is* credulous conclusion *of T iff $\exists \mathcal{E}_i$ such that $a \in \mathcal{E}_i$.*

3 A Preference-based Argumentation Framework

In [1] the basic argumentation framework of Dung has been extended into *preference-based argumentation theory (PBAT)*. The basic idea of a PBAT is to consider two binary relations between arguments:

1. A *conflict* relation, denoted by $\mathcal{C}$, that is based on the logical links between arguments.
2. A *preference* relation, denoted by $\succeq$, that captures the idea that some arguments are stronger than others. Indeed, for two arguments $a, b \in \mathcal{A}$, $a \succeq b$ means that a is at least as good as b. The relation $\succeq$ is assumed to be a partial pre-order (that is *reflexive* and *transitive*). The relation $\succ$ denotes the corresponding strict relation. That is, $a \succ b$ iff $a \succeq b$ and $b \not\succeq a$.

The two relations are combined into a unique attack relation, denoted by $\mathcal{R}$, and the Dung's semantics are applied on the resulting framework. In what follows, we will study a particular class of PBATs, where the conflict relation $\mathcal{C}$ is *irreflexive* and *symmetric*.

Definition 5 (Preference-based Argumentation Theory (PBAT)) *Given an irreflexive and symmetric conflict relation $\mathcal{C}$ and a preference relation $\succeq$ on a set of arguments $\mathcal{A}$, a* preference-based argumentation theory *(*PBAT*) on $\mathcal{A}$ is an argumentation system $T = (\mathcal{A}, \mathcal{R})$, where $(a, b) \in \mathcal{R}$ iff $(a, b) \in \mathcal{C}$ and $b \not\succ a$.*

It follows directly from the definition that if $(a, b) \in \mathcal{C}$ and $a \succeq b$ and $b \not\succeq a$, then $(a, b) \in \mathcal{R}$. Moreover, if $(a, b) \in \mathcal{C}$ and a, b are either indifferent or incompatible in $\succeq$, then $(a, b) \in \mathcal{R}$ and $(b, a) \in \mathcal{R}$. Also note that if $(a, b) \in \mathcal{C}$, then either $(a, b) \in \mathcal{R}$ or $(b, a) \in \mathcal{R}$. Finally, if $(a, b) \in \mathcal{R}$ and $(b, a) \notin \mathcal{R}$, then $a \succ b$. The following example illustrates some features of PBATs.

Example 1 *Let $\mathcal{A} = \{a, b, c, d\}$ be a set of arguments, and $\mathcal{C}$ the conflict relation on $\mathcal{A}$ defined as $\mathcal{C} = \{(a, b), (b, a), (b, c), (c, b), (c, d), (d, c)\}$. Moreover, let the preference relation $\succeq$ contain transitive closure of the set of pairs $a \succeq b$, $b \succeq c$, $c \succeq d$, and $d \succeq c$. The corresponding PBAT is $T = (\mathcal{A}, \mathcal{R})$, where $\mathcal{R} = \{(a, b), (b, c), (c, d), (d, c)\}$. Theory T has two stable extensions, $E_1 = \{a, c\}$ and $E_2 = \{a, d\}$.*

We note here that, although it seems that combining the conflict and preference relations can be done in many different ways other than the one proposed in definition 5, all of these combinations lead to counterintuitive results and properties. A detailed analysis of these possibilities will appear in an extended version of this paper.

4 Basic Properties of PBATs

In this section we present some basic properties of PBATs. To facilitate the discussion and the presentation of the results of this section as well as those of other part in the remainder of this paper, we use some basic notions from graph theory. Indeed, as with every binary relation on a set, an argumentation system T is associated with a directed graph (digraph) $\mathcal{G}_T$ whose nodes are the different arguments, and the edges represent the attack relation defined on them. The identification of graph theoretical structures has led to useful results regarding the properties of argumentation systems (e.g. [9]).

Let $G = (N, E)$ be a digraph and $n \in N$ a node of G. The in-degree of n in G is the number of nodes n' of G such that $(n', n) \in E$. A *(strongly connected) component* C of a digraph G is a maximal subgraph C of G such that for every pair of nodes $x, y \in C$, there is a path from x to y in C. If each component of a digraph G is contracted to a single node, the resulting graph is a directed acyclic one, and is called the *components graph* of G. A *top* component of a digraph G is one that has in-degree 0 in the components graph of G. Our first result characterizes the cycles of the graph of a PBAT.

Proposition 1 *Let $\mathcal{G}_T$ be the graph associated with a PBAT T = $(\mathcal{A}, \mathcal{R})$. Every cycle of $\mathcal{G}_T$ has at least two symmetric edges.*

Proof *We prove by case analysis that a cycle of $\mathcal{G}_T$ cannot have no or one symmetric edges. Let $a_1, a_2, \ldots, a_n$ be a cycle of $\mathcal{G}_T$. This means that $\forall i < n$, $(a_i, a_{i+1}) \in \mathcal{R}$ and $(a_n, a_1) \in \mathcal{R}$.*

Let us assume that this cycle has no symmetric edges, ie. $\forall i < n$, $(a_{i+1}, a_i) \notin \mathcal{R}$ and $(a_1, a_n) \notin \mathcal{R}$. Since $\forall i < n$, $(a_i, a_{i+1}) \in \mathcal{R}$ and $(a_{i+1}, a_i) \notin \mathcal{R}$, it holds that $\forall i < n$, $a_i \succeq a_{i+1}$. By transitivity, $a_1 \succeq a_n$, meaning $(a_1, a_n) \in \mathcal{R}$, contradiction.
Assume now that $a_1, a_2, \ldots, a_n$ is a cycle of $\mathcal{G}_T$ such that (a_n, a_1) is the only symmetric edge of the cycle. Assume first that the two arguments a_n, a_1 are incomparable wrt the underlying preference relation $\succeq$. The transitivity of the preference relation requires that

$a_1 \succeq a_n$, *which contradicts the incomparability of the two arguments. Assume now that* $a_1 \succeq a_n$ *and* $a_n \succeq a_1$. *Since* $a_n \succeq a_1$ *and* $a_1 \succeq a_2$, *by transitivity* $a_n \succeq a_2$. *On the other hand we have* $a_2 \succeq a_3, \ldots, a_{n-1} \succeq a_n$, *and by transitivity* $a_2 \succeq a_n$. *Hence the cycle must also contain a symmetric edge between* a_2 *and* a_n. *Therefore every cycle of* $\mathcal{G}_T$ *has at least two symmetric edges.* ■

Doutre [6] has shown that the kernels of the associated graph of an argumentation theory correspond exactly to its stable extensions. A *kernel* of a directed graph $G = (N, E)$ is a set of nodes $K \subseteq N$ such that (a) K is an *independent set*, that is, there is no pair of nodes $n_i, n_j \in K$ s.t. $(n_i, n_j) \in E$ or $(n_j, n_i) \in E$ (b) for all $n \in N \setminus K$ there is a node $n' \in K$ s.t. $(n', n) \in E$. Moreover, Duchet [7] proved that every graph with at least two symmetric edges has a kernel. By combining these two results we obtain the following theorem.

Theorem 1 *Every PBAT has a stable extension.*

We show now that the graph associated with a PBAT has no *elementary cycles* of length greater than 2. The notion of elementary cycle is defined as follows.

Definition 6 (Elementary cycle) *Let* $T = (\mathcal{A}, \mathcal{R})$ *be a PBAT and* $X = \{a_1, \ldots, a_n\}$ *be a set of arguments of* $\mathcal{A}$. X *is an* elementary cycle *of* T *iff:*

1. $\forall i \leq n-1$, $(a_i, a_{i+1}) \in \mathcal{R}$ *and* $(a_n, a_1) \in \mathcal{R}$
2. $\nexists X' \subset X$ *such that* X' *satisfies condition 1.*

Proposition 2 *Let* $T = (\mathcal{A}, \mathcal{R})$ *be a PBAT on an underlying preorder* $\succeq$. *Then,* $\mathcal{R}$ *has no elementary cycle of length greater than 2.*

Proof *Let* $a_1, \ldots, a_n$ *be arguments of* $\mathcal{A}$, *with* $n > 2$, *and assume that they form an elementary cycle, i.e.* $\forall i \leq n$, $(a_i, a_{i+1}) \in \mathcal{R}$, *and* $(a_n, a_1) \in \mathcal{R}$. *Since the cycle is elementary, then* $\nexists a_i, a_{i+1}$ *such that* $(a_i, a_{i+1}) \in \mathcal{R}$ *and* $(a_{i+1}, a_i) \in \mathcal{R}$. *Thus,* $a_i \succ a_{i+1}$, $\forall i < n$. *Therefore,* $a_1 \succ a_2 \succ \ldots a_n \succ a_1$, *contradiction.* ■

A direct consequence of the above property is that PBATs do not have elementary *odd-length* cycles. By the results of [10], this implies that PBATs are coherent, i.e., their preferred and stable extensions coincide.

Theorem 2 *Every PBAT is coherent.*

In the remaining of this section we investigate the impact of the preference relation on an argumentation system. We first define a relation $\rhd$ on the powerset of the arguments of a PBAT $T = (\mathcal{A}, \mathcal{R})$ (we denote by $\mathcal{P}(\mathcal{A})$ the powerset of $\mathcal{A}$), and then show that the stable extensions of T correspond to the most preferred elements of $\mathcal{P}(\mathcal{A})$ wrt this relation.

Definition 7 *Let* $T = (\mathcal{A}, \mathcal{R})$ *be a PBAT built on an underlying preorder* $\succeq$. *If* $A_1, A_2 \in \mathcal{P}(\mathcal{A})$, *with* $A_1 \neq A_2$, *then* $A_1 \rhd A_2$ *iff one of following holds:*

- $A_1 \supset A_2$
- *for all* a, b *such that* $a \in A_1 \setminus A_2$ *and* $b \in A_2 \setminus A_1$, *it holds that* $a \succ b$

The following result states the relation between $\rhd$ and stable extensions, and hence sheds some light on the connection between preference and argumentation.

Theorem 3 *Let* $T = (\mathcal{A}, \mathcal{R})$ *be a PBAT built on an underlying preorder* $\succeq$ *and a conflict relation* $\mathcal{C}$. E *is a stable extension of* T *iff there are no arguments* $a, b \in E$ *s.t.* $(a, b) \in \mathcal{C}$, *and for all* $A \in \mathcal{P}(\mathcal{A})$ *such that* $A \rhd E$, *there are* $a_1, a_2 \in A$ *such that* $(a_1, a_2) \in \mathcal{C}$.

Proof *Let* E *be a stable extension of* T. *Then, by definition, it contains no pair of arguments* a, b *s.t.* $(a, b) \in \mathcal{R}$. *Hence,* E *can not contain arguments* a, b *s.t.* $(a, b) \in \mathcal{C}$. *We prove by case analysis that for all* $A \in \mathcal{P}(\mathcal{A})$ *such that* $A \rhd E$ *there exists a pair of arguments* $a_1, a_2 \in A$ *s.t.* $(a_1, a_2) \in \mathcal{C}$.

Assume first a set A *with* $A \supset E$. *Since* E *is a stable extension, for all* $a \in A \setminus E$, *there is* $b \in E$, *and because* $A \supset E$, $b \in A$ *s.t.* $(b, a) \in \mathcal{R}$. *Therefore there exist* $a, b \in A$, *s.t.* $(a, b) \in \mathcal{C}$.

Assume now that $A \rhd E$ *and* $A \not\supset E$. *Again, for all* $a \in A \setminus E$, *there is* $b \in E$ *s.t.* $(b, a) \in \mathcal{R}$. *Since* $A \rhd E$, *by definition 7 follows that for all* $a \in A \setminus E$ *and* $c \in E \setminus A$, *it holds* $a \succ c$ *and hence* $(c, a) \notin R$. *Therefore, it must be the case that* $b \in E \cap A$, *which means that* A *contains a pair* a, b *such that* $(b, a) \in \mathcal{R}$, *and therefore* $(a, b) \in \mathcal{C}$.

Let now E *be a set of arguments that contains no pair of elements* a, b *s.t.* $(a, b) \in \mathcal{C}$, *and for all* $A \in \mathcal{P}(\mathcal{A})$ *such that* $A \rhd E$, *there are* $a_1, a_2 \in A$ *such that* $(a_1, a_2) \in \mathcal{C}$. *We prove that* E *is a stable extension. We show first that* E *is admissible. Observe that since* E *contains no pair of elements* a, b *s.t.* $(a, b) \in \mathcal{C}$, *it can not contain a pair* a, b *s.t.* $(a, b) \in \mathcal{R}$. *Assume that there exist* $a \in E$ *and* $b \in \mathcal{A} \setminus E$ *s.t.* $(b, a) \in \mathcal{R}$ *and there is no* $c \in E$ *such that* $(c, b) \in \mathcal{R}$. *Hence* $b \succ a$. *Then define* $D(b) = \{d | (b, d) \in \mathcal{R}$ *and* $d \in E\}$, *and construct the set* $E' = (E \setminus D(b)) \cup \{b\}$. *Then, it is the case that* $E' \rhd E$ *and furthermore there is no pair* $a_1, a_2 \in E'$ *such that* $(a_1, a_2) \in \mathcal{R}$, *and therefore* $(a_1, a_2) \in \mathcal{C}$, *contradiction.*
Assume now that there exists $b \in \mathcal{A} \setminus E$ *s.t. for all* $a \in E$ *it holds that* $(a, b) \notin \mathcal{R}$. *Clearly,* $(b, a) \notin \mathcal{R}$, *because otherwise* E *is not admissible. Then again,* $E \cup \{b\} \rhd E$ *and furthermore there is no pair* $a_1, a_2 \in E \cup \{b\}$ *such that* $(a_1, a_2) \in \mathcal{C}$, *contradiction.* ■

The example below highlights the link between the relation $\rhd$ and the stable extensions.

Example 2 *Let* $T = (\mathcal{A}, \mathcal{R})$ *be a PBAT with* $\mathcal{A} = \{a, b, c\}$ *and* $\mathcal{R}$ *composed from the conflict relation* $\mathcal{C} = \{(a, b), (b, a)(a, c), (c, a)\}$ *and preference relation that contains the pairs* $a \succeq b$ *and* $a \succeq c$. *The relation* $\rhd$ *on* $\mathcal{P}(\mathcal{A})$ *induced by* $\succeq$ *is depicted in figure 1. Since the sets* $\{a, b, c\}$, $\{a, b\}$, $\{a, c\}$ *are ruled out by* $\mathcal{C}$, *the set* $E = \{a\}$ *is the stable extension of* T.

5 Reasoning in PBATs

This section contains a preliminary investigation of the computational properties of the new argumentation framework. We start by presenting below the algorithm *stable extension* that computes a stable extension of a PBAT in *polynomial time*. Recall that finding a stable extension of a general argumentation system is an intractable task (see eg. [9]).

stable extension$(\mathcal{A}, \mathcal{R})$
$A' = \mathcal{A}$; $E = \emptyset$
While ($A' \neq \emptyset$) do
  Compute a top component C of theory $(\mathcal{A}, \mathcal{R})$
  Select a node $n \in C$ such that for all $n' \in \mathcal{A}$ with
  $(n', n) \in \mathcal{R}$ it holds that $(n, n') \in \mathcal{R}$
  $E = E \cup \{n\}$;

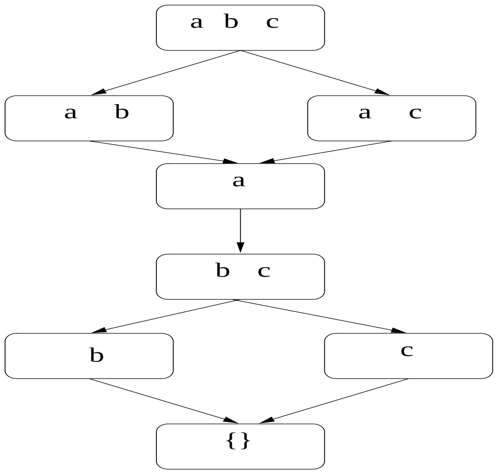

Figure 1. Ranking relation where an edge from A to B means that $A \triangleright B$.

$A' = A' - (\{n\} \cup \{n' | (n, n') \in \mathcal{R}\})$
end do
Return E

Notice that by construction the set E returned by the above algorithm does not contain two elements x, y such that $(x, y) \in \mathcal{R}$. Moreover, again by construction, for each element $x \in \mathcal{A}$ that is not included in E, there must by some element $y \in E$ such that $(y, x) \in \mathcal{R}$. Therefore, the set E returned by the algorithm is a stable extension of the input theory $(\mathcal{A}, \mathcal{R})$.

The key point of the *stable extension* algorithm is that at each iteration it finds a node n from a top component of the input theory such that for all $n' \in \mathcal{A}$ for which $(n', n) \in \mathcal{R}$, it holds that $(n, n') \in \mathcal{R}$. An informal justification of the existence of such elements is the following. Assume that the algorithm reaches a point where there is a top component C of the theory that contains no node with the above property. This means that for every node $n \in C$ there exists some other node $n' \in C$ such that $(n', n) \in \mathcal{R}$ and $(n, n') \notin \mathcal{R}$. Remove from C all symmetric edges (the edge $(x, y) \in R$ is symmetric if $(y, x) \in \mathcal{R}$ also holds). Then, in the resulting graph all nodes of C must have an incoming edge, which means that C contains a cycle with no symmetric edges, which contradicts proposition 1.

Although computing a stable extension of a PBAT can be performed in polynomial time, we prove below that credulous and skeptical reasoning in the new framework are intractable.

Theorem 4 *Let $T = (\mathcal{A}, \mathcal{R})$ be a PBAT and $a \in \mathcal{A}$. Deciding whether a is a credulous conclusion of T is NP-hard.*

Proof *We prove the claim by a reduction from 3SAT. Let $S = \{c_1, \ldots c_n\}$ be a 3SAT theory on a set of clauses $c_1, \ldots c_n$. From S we construct a PBAT $S_T = (\mathcal{A}, \mathcal{R})$. The set of arguments $\mathcal{A}$ of S_T contains the following elements:*

- *An argument l_i for each literal l_i that appears in S.*
- *An argument c_j for each clause c_j of S, $1 \leq j \leq n$.*
- *An additional argument t that corresponds to the whole theory S.*

The underlying conflict relation $\mathcal{C}$ of S_T contains the following (symmetric) pairs:

- *$(l_i, \neg l_i)$, for each argument l_i that corresponds to a literal l_i of S*
- *(l_i, c_j), if literal l_i appears in clause c_j.*
- *(c_i, t), for $1 \leq i \leq n$.*

Finally, the underlying preference relation $\succeq$ of S_T, is defined as $\succeq = \{(a, b) | a, b \in A, a \neq b\} - \{(t, c_i) | c_i$ is the argument that corresponds to clause $c_i\}$, that is, each argument that corresponds to clauses is preferred to the argument that corresponds to the theory, whereas all other arguments are indifferent to each other. Therefore, R coincides with its underlying conflict relation, with the only difference that it does not contain the pairs (t, c_i), for $1 \leq i \leq n$.

We now prove that S is satisfiable iff S_T has a stable (admissible) extension that contains argument t.
Let M be a satisfying truth assignment of S. We show that the set of arguments $E = M \cup \{t\}$ is an extension of S_T. First note that for any pair of arguments $a_i, a_j \in E$, it holds that $(a_i, a_j) \notin \mathcal{R}$. Furthermore, it holds that for each $c_i \in A$ that corresponds to a clause of S, there must be some argument $l_j \in E$ that corresponds to some literal of S such that $(l_j, c_i) \in \mathcal{R}$ (otherwise M is not satisfying). Therefore, E is a stable extension of S_T.

Let now E be a stable extension of S_T such that $t \in E$. We prove that the assignment that corresponds to the arguments of E is a satisfying one for S. This assignment does not contain any pairs of complementary literals because these pairs of literals belong to $\mathcal{R}$. Furthermore, since $t \in E$, it must be the case that $c_i \notin E$ for $1 \leq i \leq n$. Therefore it must be the case that for each clause c_i of S at least one of its literals must belong to E, therefore the assignment that corresponds to E is satisfying. ■

Proposition 3 *Let $T = (\mathcal{A}, \mathcal{R})$ be a PBAT and $a \in \mathcal{A}$. Deciding whether a is a skeptical conclusion of T is coNP-hard.*

Proof *Given a propositional theory S we construct a PBAT $T_S = (\mathcal{A}, \mathcal{R})$ in a way similar to that of the previous proof with the difference that $\mathcal{A}$ contains an additional argument t' such that pair $(t, t') \in \mathcal{C}$, $(t', t) \in \mathcal{C}$, and $t \succeq t'$, $t' \not\succeq t$. It is not difficult to prove that t' is a skeptical conclusion of T_S iff S is unsatisfiable.* ■

6 Theories without incomparability

In this section we turn our attention to PBATs without incomparability, i.e. theories $T = (\mathcal{A}, \mathcal{R})$ such that for each pair of arguments $a_i, a_j \in \mathcal{A}$, either $a_i \succeq a_j$ or $a_j \succeq a_i$. More specifically we present an algorithm that enumerates *all* stable extensions of a theory in this class with *polynomial delay*. An algorithm that enumerates the elements of a set S is said to be a polynomial delay one, if it computes the first element of the set within polynomial time in the size of the input, and furthermore the time taken by the algorithm between computing two consecutive elements of this set is also bounded by some polynomial in the size of the input.

The key property of PBATs without incomparability that is exploited by the stable extensions computation algorithm, is that the strongly connected components of the graph G_T of such a theory T contain only symmetric edges, and therefore these components are essentially undirected (sub)graphs. This useful property is proved in the following result.

Proposition 4 *Let $T = (\mathcal{A}, \mathcal{R})$ be a PBAT without incomparability, and G_T its associated digraph. If $a, b \in \mathcal{A}$ are arguments that belong to the same component of G_T and $(a, b) \in \mathcal{R}$, then $(b, a) \in \mathcal{R}$.*

Proof *Let $a, b \in \mathcal{A}$ be arguments that belong to the same component of G_T and $(a, b) \in \mathcal{R}$. Therefore $(b, a) \in \mathcal{C}$, and $a \succeq b$. Since*

a, b belong to same component there must be a path from b to a. Since there is no incomparability, by transitivity we get that $b \succeq a$. From this and the fact $(b, a) \in \mathcal{C}$ we conclude that $(b, a) \in \mathcal{R}$. ■

The kernels (recall that kernels correspond to stable extensions) of a graph that contains only symmetric edges are exactly its *maximal (w.r.t. set inclusion) independent sets (MISs)*. To see this, note that it follows from the definition, that every kernel is an MIS. On the other hand, since in this case all edges are symmetric, an MIS is also a kernel. This connection between stable models, kernels and MISs, allows us to employ well-known procedures that enumerate all maximal independent sets of a graph with polynomial delay [11].

Algorithm *all stable extensions*, that is presented below, enumerates the stable extensions of the input theory by traversing the theory from its top components downwards. Singleton components are handled separately by the first iteration of the algorithm. To enumerate the elements that belong to stable extensions and at the same time to components with more than one nodes, the algorithm utilizes a procedure that performs MISs computation with polynomial delay.

all stable extensions$(\mathcal{A}, \mathcal{R})$
$A' = \mathcal{A}$; $E = \emptyset$
While ($A' \neq \emptyset$) do
  While (A' has nodes with in-degree 0) do
    $E = E \cup \{a | a \in A'$ and has in-degree 0 $\}$
    $A' = A' - (E \cup \{a' | a \in E$ and $(a, a') \in \mathcal{R}\})$
  end do
  Select a top component C of $(\mathcal{A}, \mathcal{R})$
  For each MIS M of C computed with polynomial delay do
    $E = E \cup M$;
    $A' = A' - (M \cup \{a' | a \in M$ and $(a, a') \in \mathcal{R}\})$
    call *stable extension*$(A', \mathcal{R})$
  end do
end do
Return E

It is known [13] that the number of MISs of a graph with n nodes is at most $n^{n/3}$. Therefore, if a PBAT has m components each of which has at least 2 nodes and at most k nodes, then the theory has at most $n^{mk/3}$ stable extensions. Hence, the run time of the algorithm is exponential in mk. For "small" values of m and k, the above algorithm can be also used to perform credulous and skeptical reasoning. The idea is to simply enumerate all stable extensions of the input theory, and terminate as soon as the given argument belongs (credulous reasoning) or does not belong (skeptical reasoning) to one of the stable extensions.

Consider now a PBAT $T = (\mathcal{A}, \mathcal{R})$ where the underlying preference $\succeq$ relation contains neither incomparability nor indifference. Therefore, for all pairs of arguments $a_i, a_j \in \mathcal{A}$, either $a_i \succeq a_j$ or $a_j \succeq a_i$ holds, but not both. In this case the graph of T is acyclic and T has exactly one stable extension. The first iteration of the algorithm *all stable extensions* above computes this unique stable extension in polynomial time. Obviously, the same procedure can be used for credulous and skeptical reasoning in this restricted class of PBATs.

7 Conclusion and Future Work

In this paper we presented an abstract preference-based argumentation framework. Although other works in the literature (see e.g. [2, 1, 5]) have also acknowledged the importance of incorporating preferences in argumentation systems, very little have been said about the theoretical and computational properties of such systems.

This paper is a work in the direction of filling this gap by proposing a new preference-based argumentation framework and studying its basic properties. We have shown that the theories of the new framework have always stable extensions and are coherent. We also characterized the structure of preference-based argumentation theories by extending previous works that attempted to link argumentation and graph theory (see eg. [9] for a recent example). Moreover, it seems that the transitivity of the underlying preference relation imposes a strong structure on the preference-based argumentation theories that can be exploited computationally. Indeed, some computational problems become easier in the new framework, whereas others remain intractable.

There are many directions for future research. We plan to investigate more deeply the structural properties of PBATs and further extend the link with graph theory. Moreover, we intend to study the properties of the $\triangleright$ relation and identify its effects on argumentation. Finally, the computational properties of the new framework will be explored more fully in the future.

ACKNOWLEDGEMENTS

We thank one of the reviewers for many helpful comments.

REFERENCES

[1] L. Amgoud and C. Cayrol, 'On the acceptability of arguments in preference-based argumentation framework', in *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, pp. 1–7, (1998).

[2] L. Amgoud, Y. Dimopoulos, and P. Moraitis, 'A Unified and General Framework for Argumentation-based Negotiation', in *Proc. 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pp. 963–970. ACM Press, (2007).

[3] L. Amgoud and H. Prade, 'Explaining qualitative decision under uncertainty by argumentation', in *21st National Conference on Artificial Intelligence, AAAI'06*, pp. 16 – 20, (2006).

[4] T. Bench-Capon and P. Dunne, 'Argumentation in artificial intelligence', *Artif. Intell.*, **171**(10-15), 619–641, (2007).

[5] T. J. M. Bench-Capon, 'Persuasion in practical argument using value-based argumentation frameworks', *Journal of Logic and Computation*, **13**(3), 429–448, (2003).

[6] S. Doutre, *Autour de la sémantique préférée des systèmes d'argumentation*, PhD thesis, Université Paul Sabatier, Toulouse – France., 2002.

[7] P. Duchet, *Représentations, Noyaux en Théorie des Graphes et Hypergraphes*, PhD thesis, 1979.

[8] 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).

[9] P. Dunne, 'Computational properties of argument systems satisfying graph-theoretic constraints', *Artif. Intell.*, **171**(10-15), 701–729, (2007).

[10] P. Dunne and T. Bench Capon, 'Coherence in finite argument systems', *Artificial Intelligence*, **141 (1–2)**, 187–203, (2002).

[11] D. S. Johnson, C. H. Papadimitriou, and M. Yannakakis, 'On generating all maximal independent sets', *Inf. Process. Lett.*, **27**(3), 119–123, (1988).

[12] A. Kakas and P. Moraitis, 'Argumentation based decision making for autonomous agents', in *Proc. 2nd International Joint Conference on Autonomous Agents and Multi-Agents Systems*, pp. 883–890, (2003).

[13] J. Moon and L. Moser, 'On cliques in graphs', *Israel Journal of Mathematics*, **3**, 23–28, (1965).

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