Handling inconsistency with preference-based argumentation

Leila Amgoud Srdjan Vesic

IRIT-CNRS, 118, route de Narbonne,
31062 Toulouse Cedex 4 France
{amgoud, vesic}@irit.fr

Abstract. Argumentation is a promising approach for handling inconsistent knowledge bases, based on the justification of plausible conclusions by arguments. Due to inconsistency, arguments may be attacked by counterarguments. The problem is thus to evaluate the arguments in order to select the most acceptable ones.

The aim of this paper is to make a bridge between the argumentation-based and the coherence-based approaches for handling inconsistency. We are particularly interested by the case where priorities between the formulas of an inconsistent knowledge base are available. For that purpose, we will use the rich preference-based argumentation framework (PAF) we have proposed in an earlier work. A rich PAF has two main advantages: i) it overcomes the limits of existing PAFs, and ii) it encodes two different roles of preferences between arguments (handling critical attacks and refining the evaluation of arguments). We show that there exist full correspondences between particular cases of these PAF and two well known coherence-based approaches, namely the preferred sub-theories and the democratic ones.

1 Introduction

An important problem in the management of knowledge-based systems is the handling of inconsistency. Inconsistency may be present for mainly three reasons:

- The knowledge base includes default rules. Let us consider for instance the general rules ‘birds fly’, ‘penguins are birds’ and the specific rule ‘penguins do not fly’. If we add the fact ‘Tweety is a penguin’, we may conclude that Tweety does not fly because it is a penguin, and also that Tweety flies because it is a bird.
- In model-based diagnosis, a knowledge base contains a description of the normal behavior of a system, together with observations made on this system. Failure detection occurs when observations conflict with the normal functioning mode of the system and the hypothesis that the components of the system are working well; that leads to diagnose which component fails;
- Several consistent knowledge bases pertaining to the same domain, but coming from different sources of information, are available. For instance, each source is a reliable specialist in some aspect of the concerned domain but is less reliable in other aspects. A straightforward way of building a global base Σ is to concatenate the knowledge bases Σ_i provided by each source. Even if each base Σ_i is consistent, it is unlikely that their concatenation will be consistent also.

Classical logic has many appealing features for knowledge representation and reasoning, but unfortunately when reasoning with inconsistent information, i.e. drawing conclusions from an inconsistent knowledge base, the set of classical consequences is trivialized. To solve this problem, two kinds of approaches have been proposed. The first one, called *coherence-based* approach and initiated in [10], proposes to give up some formulas of the knowledge base in order to get one or several consistent subbases of the original base. Then plausible conclusions may be obtained by applying classical entailment on these subbases. The second approach accepts inconsistency and copes with it. Indeed, it retains all the available information but prohibits the logic from deriving trivial conclusions. Argumentation is one of these approaches. Its basic idea is that each plausible conclusion inferred from the knowledge base is justified by some reason(s), called *argument*(s), for believing in it. Due to inconsistency, those arguments may be attacked by other arguments (called counterarguments). The problem is thus to evaluate the arguments in order to select the most acceptable ones.

In [7], it has been shown that the results of the coherence-based approach proposed in [10] can be recovered within Dung's argumentation framework [9]. Indeed, there is a full correspondence between the maximal consistent subbases of a given inconsistent knowledge base and the stable extensions of the argumentation system built over the same base. In [10], the formulas of the knowledge base are assumed to be equally preferred. This assumption has been discarded in [6] and in [8]. Indeed, in the former work, a knowledge base is equipped with a total preorder. Thus, instead of computing the maximal consistent subbases, *preferred sub-theories* are computed. These sub-theories are consistent subbases that privilege the most important formulas. In [8], the knowledge base is rather equipped with a partial reorder. The idea was to define a preference relation, called *democratic relation*, between the consistent subbases. The best subbases, called *democratic sub-theories*, wrt this relation are used for inferring conclusions from the knowledge base.

The aim of this paper is to investigate whether it is possible to recover the results of these two works within an argumentation framework. Since priorities are available, it is clear that we need a preference-based argumentation framework (PAF). Recently, we have shown in [3] that existing PAFs (developed in [2, 4]) are not appropriate since they may return unintended results, especially when the attack relation is asymmetric. Moreover, their results are not optimal since they may be refined by the available preferences between arguments. Consequently, we have proposed in the same paper (i.e. [3]) a new family of PAFs, called *rich PAF*, that encodes two distinct roles of preferences between arguments: handling critical attacks (that is an argument is stronger than its attacker) and refining the result of the evaluation of arguments using acceptability semantics. In this paper, we show that there is a full correspondence between the preferred sub-theories proposed in [6] and the stable extensions of an instance of this rich PAF, and also a full correspondence between the democratic sub-theories developed in [8] and another instance of the rich PAF. The two correspondences are obtained by choosing appropriately the main components of a rich PAF: the definition of an argument, the attack relation, the preference relation between arguments and the preference relation between subsets of arguments.

The paper is organized as follows: Sections 2 and 3 recall respectively the rich PAF if [3] and the two works of [6, 8]. Section 4 shows how instances of the rich PAF compute preferred and democratic sub-theories of a knowledge base. The last section is devoted to some concluding remarks.

2 Preference-based argumentation frameworks

In [9], Dung has developed the most abstract argumentation framework in the literature. It consists of a set of arguments and an attack relation between them.

Definition 1 (Argumentation framework [9]). *An* argumentation framework *(AF) is a pair $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, where $\mathcal{A}$ is a set of arguments and $\mathcal{R}$ is an attack relation ($\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$). The notation $a\mathcal{R}b$ means that the argument a* attacks *the argument b.*

In the above definition, the arguments and attacks are abstract entities since Dung's framework completely abstracts from the application. However, the two components can be defined as follows when handling inconsistency in a *propositional* knowledge base Σ.

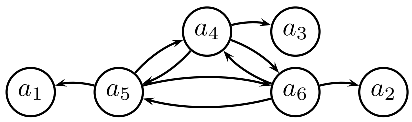
Definition 2 (Argument - Undercut). *Let Σ be a propositional knowledge base.*

- *An* argument *is a pair $\alpha = (H, h)$ such that:*
  - $H \subseteq \Sigma$
  - *H is consistent*
  - $H \vdash h$
  - *$\nexists H' \subset H$ such that H' is consistent and $H' \vdash h$.*
- *An argument (H, h)* undercuts *an argument (H', h') iff $\exists h'' \in H'$ s.t. $h \equiv \neg h''$.*

Example 1. Let $\Sigma = \{x, \neg y, x \rightarrow y\}$ be a propositional knowledge base. The following arguments are built from this base:

$a_1 : (\{x\}, x)$	$a_2 : (\{\neg y\}, \neg y)$
$a_3 : (\{x \rightarrow y\}, x \rightarrow y)$	$a_4 : (\{x, \neg y\}, x \wedge \neg y)$
$a_5 : (\{\neg y, x \rightarrow y\}, \neg x)$	$a_6 : (\{x, x \rightarrow y\}, y)$

The figure below depicts the attacks wrt "undercut".



Different *acceptability semantics* for evaluating arguments have been proposed in the same paper [9]. Each semantics amounts to define sets of acceptable arguments, called *extensions*. For the purpose of our paper, we only need to recall stable semantics.

Definition 3 (Conflict-free, Stable semantics [9]). *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF, $\mathcal{B} \subseteq \mathcal{A}$.*

- *$\mathcal{B}$ is* conflict-free *iff $\nexists$ $a, b \in \mathcal{B}$ such that $a\mathcal{R}b$.*

– $\mathcal{B}$ *is a* stable *extension iff it is conflict-free and attacks any element in* $\mathcal{A} \setminus \mathcal{B}$.

Example 1 (Cont): The argumentation framework of Example 1 has three stable extensions: $\mathcal{E}_1 = \{a_1, a_2, a_4\}$, $\mathcal{E}_2 = \{a_2, a_3, a_5\}$ and $\mathcal{E}_3 = \{a_1, a_3, a_6\}$.

The attack relation is the backbone of any acceptability semantics in [9]. An attack from an argument b towards an argument a always wins unless b is itself attacked by another argument. However, this assumption is very strong because some attacks cannot always 'survive'. Especially when the attacked argument is stronger than its attacker.

Throughout the paper, the relation $\geq\ \mathcal{A} \times \mathcal{A}$ is assumed to be a preorder (reflexive and transitive). For two arguments a and b, writing $a \geq b$ (or $(a, b) \in \geq$) means that a is at least as strong as b. The relation $>$ is the strict version of $\geq$. Indeed, $a > b$ iff $a \geq b$ and not ($b \geq a$). Examples of such relations are those based on the certainty level of the formulas of a propositional knowledge base Σ. The base Σ is equipped with a total preorder $\trianglerighteq$. For two formulas x and y, writing $x \trianglerighteq y$ means that x is at least as certain as y. In this case, the base Σ is stratified into $\Sigma_1 \cup \ldots \cup \Sigma_n$ such that formulas of Σ_i have the same certainty level and are more certain than formulas in Σ_j where $j > i$. The stratification of Σ enables to define a certainty level of each subset $\mathcal{S}$ of Σ. It is the highest number of stratum met by this subset. Formally:

$$\texttt{Level}(\mathcal{S}) = \max\{i \mid \exists\, x \in \mathcal{S} \cap \Sigma_i\} \text{ (with } \texttt{Level}(\emptyset) = 0).$$

The above certainty level is used in [5] in order to define a total preorder on the set of arguments that can be built from a knowledge base. The preorder is defined as follows:

Definition 4 (Weakest link principle [5]). *Let $\Sigma = \Sigma_1 \cup \ldots \cup \Sigma_n$ be a propositional knowledge. An argument (H, h) is preferred to (H', h'), denoted by $(H, h) \geq_{WLP} (H', h')$, iff* $\texttt{Level}(H) \leq \texttt{Level}(H')$.

Example 1 (Cont): Assume that $\Sigma = \Sigma_1 \cup \Sigma_2$ with $\Sigma_1 = \{x\}$ and $\Sigma_2 = \{x \to y, \neg y\}$. It holds that $\texttt{Level}(\{x\}) = 1$ while $\texttt{Level}(\{\neg y\}) = \texttt{Level}(\{x \to y\}) = \texttt{Level}(\{x, \neg y\}) = \texttt{Level}(\{\neg y, x \to y\}) = \texttt{Level}(\{x, x \to y\}) = 2$. Thus, $a_1 \geq_{WLP} a_2, a_3, a_4, a_5, a_6$ while the five other arguments are all equally preferred.

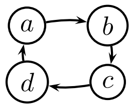
In [2, 4], Dung's argumentation framework has been extended by preferences between arguments. The idea behind those works is to remove *critical* attacks[1] and to apply Dung's semantics on the remaining attacks. Unfortunately, this solution does not work, in particular, when the attack relation is asymmetric.

Example 1 (Cont): The classical approaches of PAFs remove the critical attack from a_5 to a_1 (since $a_1 >_{WLP} a_5$) and get $\{a_1, a_2, a_3, a_5\}$ as a stable extension. Note that this extension, which intends to support a *coherent point of view*, is conflicting since it contains both a_1 and a_5. Consequently, the union of the supports of its arguments is an inconsistent set.

[1] An attack $(b, a) \in \mathcal{R}$ is *critical* iff $a \geq b$ and not $(b \geq a)$.

The approach followed in [2, 4] suffers from another problem. Its results may need to be *refined* by preferences between arguments as shown by the following example.

Example 2. Let us consider the AF depicted in the figure below.



Assume that $a > b$ and $c > d$. The corresponding PAF has two stable extensions: $\{a, c\}$ and $\{b, d\}$. Note that any element of $\{b, d\}$ is weaker than at least one element of the set $\{a, c\}$. Thus, it is natural to consider $\{a, c\}$ as better than $\{b, d\}$. Consequently, we may conclude that the two arguments a and c are "more acceptable" than b and d.

What is worth noticing is that a refinement amounts to *compare* subsets of arguments. In Example 2, the so-called *democratic* relation, $\succeq_d$, can be used for comparing the two sets $\{a, c\}$ and $\{b, d\}$. This relation is defined as follows:

Definition 5 (Democratic relation). *Let Δ be a set of objects and $\geq \subseteq \Delta \times \Delta$ be a partial preorder. For $\mathcal{X}, \mathcal{X}' \subseteq \Delta$, $\mathcal{X} \succeq_d \mathcal{X}'$ iff $\forall x' \in \mathcal{X}' \setminus \mathcal{X}$, $\exists x \in \mathcal{X} \setminus \mathcal{X}'$ such that $x > x'$.*

In [3], we have proposed a novel approach which palliates the limits of the existing ones. It follows two steps:

1. To repair the critical attacks by computing a new attack relation $\mathcal{R}_r$.
2. To refine the results of the framework $(\mathcal{A}, \mathcal{R}_r)$ by comparing its extensions using a refinement relation.

The idea behind the first step is to modify the graph of attacks in such a way that, for any critical attack, the preference between the arguments is taken into account and the conflict between the two arguments of the attack is represented. For this purpose, we *invert* the arrow of the critical attack. For instance, in Example 1, the arrow from a_5 to a_1 is replaced by another arrow emanating from a_1 towards a_5. The intuition behind this is that an attack between two arguments represents in some sense two things: i) an incoherence between the two arguments, and ii) a kind of preference determined by the direction of the attack. Thus, in our approach, the direction of the arrow represents a real preference between arguments. Moreover, the conflict is kept between the two arguments. Dung's acceptability semantics are then applied on the modified graph.

Definition 6 (PAF [3]). *A* preference-based argumentation framework *(PAF) is a tuple $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ where $\mathcal{A}$ is a set of arguments, $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is an attack relation and $\geq$ is a (partial or total) preorder on $\mathcal{A}$. The extensions of $\mathcal{T}$ under a given semantics are the extensions of the argumentation framework $(\mathcal{A}, \mathcal{R}_r)$, called* repaired framework*, under the same semantics with: $\mathcal{R}_r = \{(a, b) | (a, b) \in \mathcal{R}$ and not $(b > a)\} \cup \{(b, a) | (a, b) \in \mathcal{R}$ and $b > a\}$.*

This approach does not suffer from the drawback of the existing one. Indeed, it delivers conflict-free extensions of arguments.

Property 1. Let $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq)$ be a PAF and $\mathcal{E}_1, \ldots, \mathcal{E}_n$ its extensions under a given semantics. For all $i = 1, \ldots, n$, $\mathcal{E}_i$ is conflict-free wrt $\mathcal{R}$.

At the second step, the result of the above PAF is refined using a refinement relation. The two steps are captured in an abstract framework, called *rich preference-based argumentation framework*.

Definition 7 (Rich PAFs [3]). *A* rich PAF *is a tuple $\mathcal{T} = (\mathcal{A}, \mathcal{R}, \geq, \succeq)$ where $\mathcal{A}$ is a set of arguments, $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is an attack relation, $\geq \subseteq \mathcal{A} \times \mathcal{A}$ is a (partial or total) preorder and $\succeq \subseteq \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A})$[2] is a refinement relation. The extensions of $\mathcal{T}$ under a given semantics are the elements of* $\mathtt{Max}(\mathcal{S}, \succeq)$[3] *where $\mathcal{S}$ is the set of extensions (under the same semantics) of the PAF $(\mathcal{A}, \mathcal{R}, \geq)$.*

Example 3. Let us consider the argumentation framework depicted in the left side of the following figure.

Assume that $a > b$, $c > d$ and $b > e$. The repaired framework corresponding to $(\mathcal{A}, \mathcal{R}, \geq)$ is depicted in the right side of the above figure. This latter has two stable extensions $\{a, c\}$ and $\{b, d\}$. According to the democratic relation $\succeq_d$, it is clear that the first extension is better than the second one. Thus, the set $\{a, c\}$ is the stable extension of the rich PAF $(\mathcal{A}, \mathcal{R}, \geq, \succeq_d)$.

In [3], we have studied deeply the properties of the rich PAF. However, for the purpose of this paper we do not need to recall them.

3 Coherence-based approach for handling inconsistency

Coherence-based approach for handling inconsistency in a propositional knowledge base Σ follows two steps: At the first step, some subbases of Σ are chosen. In [10], these subbases are the maximal (for set inclusion) consistent ones. At the second step, an inference mechanism is chosen. This later defines the inferences to be made from Σ. An example of inference mechanism is the one that infers a formula if it is a classical conclusion of all the chosen subbases.

Several works have been done on choosing the subbases, in particular when Σ is equipped with a (total or partial) preorder $\trianglerighteq$ ($\trianglerighteq \subseteq \Sigma \times \Sigma$). Recall that when $\trianglerighteq$ is total, Σ is stratified into $\Sigma_1 \cup \ldots \cup \Sigma_n$ such that $\forall i, j$ with $i \neq j$, $\Sigma_i \cup \Sigma_j = \emptyset$. Moreover, Σ_1 contains the most important formulas while Σ_n contains the least important ones.

In [6], the knowledge base Σ is equipped with a total preorder. The chosen subbases privilege the most important formulas.

[2] $\mathcal{P}(\mathcal{A})$ is the powerset of the set $\mathcal{A}$.

[3] $\mathtt{Max}(\mathcal{S}, \succeq) = \{s \in \mathcal{S} \mid \nexists s' \in \mathcal{S}$ s.t. $s' \succeq s$ and not $s \succeq s'\}$.

Definition 8 (Preferred sub-theory [6]). *Let Σ be stratified into $\Sigma_1 \cup \ldots \cup \Sigma_n$. A* preferred sub-theory *is a set $\mathcal{S} = \mathcal{S}_1 \cup \ldots \cup \mathcal{S}_n$ such that $\forall k \in [1, n]$, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_k$ is a maximal (for set inclusion) consistent subbase of $\Sigma_1 \cup \ldots \cup \Sigma_k$.*

Example 1 (Cont): The knowledge base $\Sigma = \Sigma_1 \cup \Sigma_2$ with $\Sigma_1 = \{x\}$ and $\Sigma_2 = \{x \to y, \neg y\}$ has two preferred sub-theories: $\mathcal{S}_1 = \{x, x \to y\}$ and $\mathcal{S}_2 = \{x, \neg y\}$.

It can be shown that the preferred sub-theories of a knowledge base Σ are maximal (wrt set inclusion) consistent subbases of Σ.

Property 2. Each preferred sub-theory of a knowledge base Σ is a maximal (for set inclusion) consistent subbase of Σ.

In [8], the above definition has been extended to the case where Σ is equipped with a partial preorder $\trianglerighteq$. The basic idea was to define a preference relation on the power set of Σ. The best elements according to this relation are the preferred theories , called also *democratic sub-theories.* The relation that generalizes preferred sub-theories is the democratic relation (see Definition 5). In this context, Δ is Σ and $\geq$ is the relation $\trianglerighteq$. In what follows, $\triangleright$ denotes the strict version of $\trianglerighteq$. Thus:

$$\text{Let } \mathcal{S}, \mathcal{S}' \subseteq \Sigma.\ \mathcal{S} \succeq_d \mathcal{S}' \text{ iff } \forall x' \in \mathcal{S}' \setminus \mathcal{S}, \exists x \in \mathcal{S} \setminus \mathcal{S}' \text{ such that } x \triangleright x'.$$

Definition 9 (Democratic sub-theory [8]). *Let Σ be propositional knowledge base and $\trianglerighteq \subseteq \Sigma \times \Sigma$ be a partial preorder. A* democratic sub-theory *is a set $\mathcal{S} \subseteq \Sigma$ such that $\mathcal{S}$ is consistent and $(\nexists \mathcal{S}' \subseteq \Sigma)$ s.t. $\mathcal{S}'$ is consistent and $\mathcal{S}' \succeq_d \mathcal{S}$.*

Example 4. Let $\Sigma = \{x, \neg x, y, \neg y\}$ be such that $\neg x \trianglerighteq y$ and $\neg y \trianglerighteq x$. Let $\mathcal{S}_1 = \{x, y\}$, $\mathcal{S}_2 = \{x, \neg y\}$, $\mathcal{S}_3 = \{\neg x, y\}$, and $\mathcal{S}_4 = \{\neg x, \neg y\}$. The three subbases $\mathcal{S}_2$, $\mathcal{S}_3$ and $\mathcal{S}_4$ are the democratic sub-theories of Σ. However, $\mathcal{S}_1$ is not a democratic sub-theory since $\mathcal{S}_4 \succeq_d \mathcal{S}_1$.

It is easy to show that the democratic sub-theories of a knowledge base Σ are maximal (for set inclusion) consistent.

Property 3. Each democratic sub-theory of a knowledge base Σ is a maximal (for set inclusion) consistent subbase of Σ.

4 Computing sub-theories with argumentation

This section shows how two instances of the rich PAF presented in Section 2 compute the preferred and the democratic sub-theories of a propositional knowledge base Σ. The two instances use all the arguments that can be built from Σ using Definition 2 (i.e. the set $\texttt{Arg}(\Sigma)$). Similarly, they both use the attack relation "Undercut" given also in Definition 2. However, as we will see next, they are grounded on distinct preference relations between arguments. The last component of a rich PAF is a preference relation on the power set of $\texttt{Arg}(\Sigma)$. Both instances will use the democratic relation $\succeq_d$. Thus, for recovering preferred and democratic sub-theories, we will use two instances of the rich PAF $(\mathtt{Arg}(\Sigma), \text{Undercut}, \geq, \succeq_d)$.

It can be shown that when the preference relation $\geq$ is a total preorder, then the stable extensions of the PAF $(\mathtt{Arg}(\Sigma), \text{Undercut}, \geq)$ are all incomparable wrt the democratic relation $\succeq_d$.

Property 4. Let $\mathcal{T} = (\mathtt{Arg}(\Sigma), \text{Undercut}, \geq)$ be a PAF. For all stable extensions $\mathcal{E}$ and $\mathcal{E}'$ of $\mathcal{T}$ with $\mathcal{E} \neq \mathcal{E}'$, if $\geq$ is a total preorder, then $\neg(\mathcal{E} \succeq_d \mathcal{E}')$.

From the previous property, it follows that the stable extensions of $(\mathtt{Arg}(\Sigma)$, Undercut, $\geq)$ coincide with those of the rich PAF $(\mathtt{Arg}(\Sigma)$, Undercut, $\geq$, $\succeq_d)$.

Property 5. If $\geq$ is a total preorder, then the stable extensions of $(\mathtt{Arg}(\Sigma)$, Undercut, $\geq$, $\succeq_d)$ are exactly the stable extensions of $(\mathtt{Arg}(\Sigma)$, Undercut, $\geq)$.

Let us start by introducing some useful notations.

Notations: Let $a = (H, h)$ be an argument (in the sense of Definition 2). The functions Supp and Conc return respectively the support H and the conclusion h of the argument a. For $\mathcal{S} \subseteq \Sigma$, $\mathtt{Arg}(\mathcal{S}) = \{(H, h) \mid (H, h)$ is an argument in the sense of Definition 2 and $H \subseteq \mathcal{S}\}$. Thus, $\mathtt{Arg}(\Sigma)$ denotes the set of all the arguments that can be built from the whole knowledge base Σ. For $\mathcal{B} \subseteq \mathtt{Arg}(\Sigma)$, $\mathtt{Base}(\mathcal{B}) = \bigcup \mathtt{Supp}(a)$ where $a \in \mathcal{B}$.

The following result summarizes some useful properties of the above functions.

Property 6.

– For any consistent subbase $\mathcal{S} \subseteq \Sigma$, $\mathcal{S} = \mathtt{Base}(\mathtt{Arg}(\mathcal{S}))$.
– The function Base is surjective but not injective.
– For any $\mathcal{E} \subseteq \mathtt{Arg}(\Sigma)$, $\mathcal{E} \subseteq \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$.
– The function Arg is injective but not surjective.

Another property that is important for the rest of the paper relates the notion of consistency of a set of formulas to that of conflict-freeness of a set of arguments.

Property 7. A set $\mathcal{S} \subseteq \Sigma$ is consistent *iff* $\mathtt{Arg}(\mathcal{S})$ is conflict-free.

The following example shows that the previous property does not hold for an arbitrary set of arguments.

Example 5. Let $\mathcal{E} = \{(\{x\}, x), (\{x \to y\}, x \to y), (\{\neg y\}, \neg y)\}$. It is obvious that $\mathcal{E}$ is conflict-free while $\mathtt{Base}(\mathcal{E})$ is not consistent.

In the rest of this paper, we assume that a knowledge base Σ contains only consistent formulas.

4.1 Recovering the preferred sub-theories

In this section, we will show that there is a full correspondence between the preferred sub-theories of a knowledge base Σ and the stable extensions of the PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$). Recall that the relation $\geq_{WLP}$ is based on the weakest link principle and privileges the arguments whose less important formulas are more important than the less important formulas of the other arguments. This relation is a total preorder and is defined over a knowledge base that is itself equipped with a total preorder. According to Property 5, the stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$) coincide with those of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}, \succeq_d$).

The first result shows that from a preferred sub-theory, it is possible to build a unique stable extension of the PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$).

Theorem 1. *Let Σ be a stratified knowledge base. For all preferred sub-theory $\mathcal{S}$ of Σ, it holds that:*

- $\texttt{Arg}(\mathcal{S})$ *is a stable extension of* ($\texttt{Arg}(\Sigma)$, *Undercut,* $\geq_{WLP}$)
- $\mathcal{S} = \texttt{Base}(\texttt{Arg}(\mathcal{S}))$

Similarly, we show that each stable extension of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$) is built from a unique preferred sub-theory of Σ.

Theorem 2. *Let Σ be a stratified knowledge base. For all stable extension $\mathcal{E}$ of* ($\texttt{Arg}(\Sigma)$, *Undercut,* $\geq_{WLP}$)*, it holds that:*

- $\texttt{Base}(\mathcal{E})$ *is a preferred sub-theory of* Σ
- $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$

The next theorem shows that there exists a one-to-one correspondence between preferred sub-theories of Σ and stable extensions of ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$).

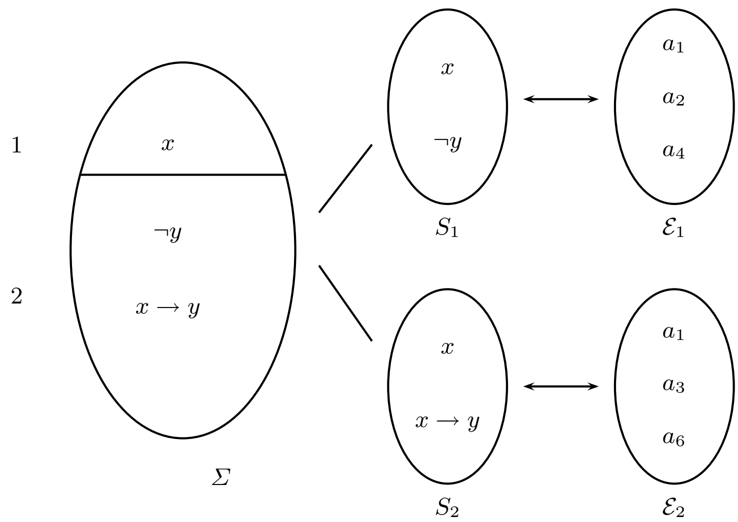
Theorem 3. *Let $\mathcal{T}$ = ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$) be a PAF over a stratified knowledge base Σ. The stable extensions of $\mathcal{T}$ are exactly the $\texttt{Arg}(\mathcal{S})$ where $\mathcal{S}$ ranges over the preferred sub-theories of Σ.*

From the above result, it follows that the PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$) has at least one stable extension unless the formulas of Σ are all inconsistent.

Corollary 1 *The PAF ($\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP}$) has at least one stable extension.*

Example 1 (Cont): Figure 1 shows the two preferred sub-theories of Σ as well as the two stable extensions of the corresponding PAF.

Fig. 1. Preferred sub-theories of Σ + Stable extensions of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{WLP})$



4.2 Recovering the democratic sub-theories

Recall that the democratic sub-theories of a knowledge base Σ generalize the preferred sub-theories when Σ is equipped with a partial preorder $\trianglerighteq$. Thus, in order to capture the democratic sub-theories, we will use the generalized version of the preference relation $\geq_{WLP} \subseteq \texttt{Arg}(\Sigma) \times \texttt{Arg}(\Sigma)$:

Definition 10 (Generalized weakest link principle [1]). *Let Σ be a knowledge base which is equipped with a partial preorder $\trianglerighteq$. For two arguments $(H, h), (H', h') \in \texttt{Arg}(\Sigma)$, $(H, h) \geq_{GWLP} (H', h')$ iff $\forall k \in H$, $\exists k' \in H'$ such that $k \triangleright k'$ (i.e. $k \trianglerighteq k'$ and not ($k' \trianglerighteq k$)).*

It can be shown that from each democratic sub-theory of a knowledge base Σ, a stable extension of $(\texttt{Arg}(\Sigma), \text{Undercut}, \geq_{GWLP})$ can be built.

Theorem 4. *Let Σ be a knowledge base which is equipped with a partial preorder $\trianglerighteq$. For all democratic sub-theory $\mathcal{S}$ of Σ, it holds that $\texttt{Arg}(S)$ is a stable extension of $(\texttt{Arg}(\Sigma), \textit{Undercut}, \geq_{GWLP})$.*

The following result shows that each stable extension of the PAF $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP})$ returns a maximal consistent subbase of Σ.

Theorem 5. *Let Σ be a knowledge base which is equipped with a partial preorder $\trianglerighteq$. For all stable extension $\mathcal{E}$ of $(\texttt{Arg}(\Sigma),$ Undercut $, \geq_{GWLP})$, it holds that:*

– $\texttt{Base}(\mathcal{E})$ *is a maximal (for set inclusion) consistent subbase of Σ.*

– $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$.

The following example shows that the stable extensions of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP})$ do not necessarily return democratic sub-theories.

Example 4 (Cont): Recall that $\Sigma = \{x, \neg x, y, \neg y\}$, $\neg x \trianglerighteq y$ and $\neg y \trianglerighteq x$. Let $\mathcal{S} = \{x, y\}$. It can be checked that the set $\texttt{Arg}(\mathcal{S})$ is a stable extension of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP})$. However, $\mathcal{S}$ is not a democratic sub-theory since $\{\neg x, \neg y\} \succeq_d \mathcal{S}$.

It can also be shown that the converse of the above theorem is not true. Indeed, a knowledge base may have a maximal consistent subbase $\mathcal{S}$ and $\texttt{Arg}(\mathcal{S})$ is not a stable extension of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP})$. Let us consider the following example.

Example 6. Let $\Sigma = \{x, \neg x\}$ and $x \rhd \neg x$. It is clear that $\{\neg x\}$ is a maximal consistent subbase of Σ while $\texttt{Arg}(\{\neg x\})$ is not a stable extension of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP})$.

The following result establishes a link between the 'best' maximal consistent subbases of Σ wrt the democratic relation $\succeq_d$ and the 'best' sets of arguments wrt the same relation $\succeq_d$.

Theorem 6. *Let $\mathcal{S}, \mathcal{S}' \subseteq \Sigma$ be maximal (for set inclusion) consistent subbases of Σ. It holds that $\mathcal{S} \succeq_d \mathcal{S}'$ iff $\texttt{Arg}(\mathcal{S}) \succeq_d \texttt{Arg}(\mathcal{S}')$.*

We also show that from each democratic sub-theory of Σ, one can build a stable extension of the corresponding rich PAF, and each stable extension of the rich PAF is built from a democratic sub-theory.

Theorem 7. *Let Σ be equipped with a partial preorder $\trianglerighteq$.*

– *For all democratic sub-theory $\mathcal{S}$ of Σ, $\texttt{Arg}(\mathcal{S})$ is a stable extension of the rich PAF $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP}, \succeq_d)$.*
– *For each stable extension $\mathcal{E}$ of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP}, \succeq_d)$, $\texttt{Base}(\mathcal{E})$ is a democratic sub-theory of Σ.*

Finally, we show that there is a one-to-one correspondence between the democratic sub-theories of a base Σ and the stable extensions of its corresponding rich PAF.

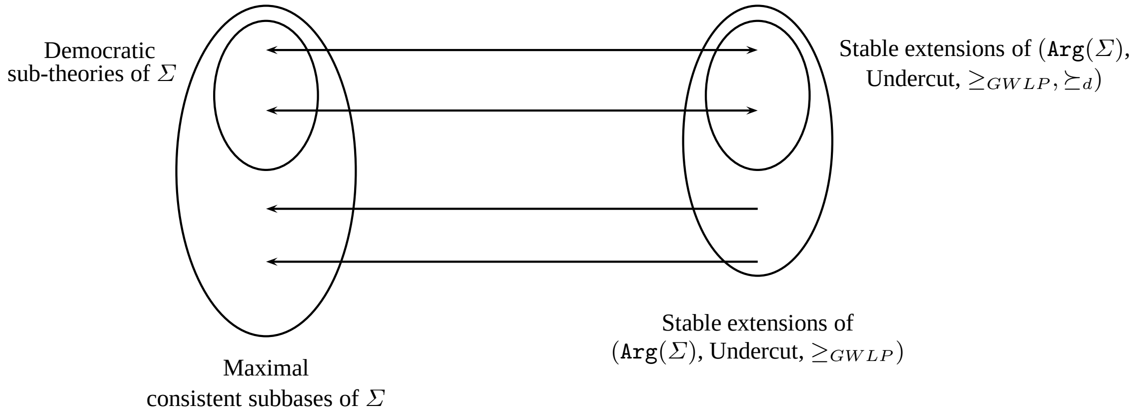
Theorem 8. *The stable extensions of $(\texttt{Arg}(\Sigma)$, Undercut, $\geq_{GWLP}, \succeq_d)$ are exactly the $\texttt{Arg}(\mathcal{S})$ where $\mathcal{S}$ ranges over the democratic subtheories of Σ.*

Figure 2 synthetizes the different links between the democratic sub-theories of a knowledge base Σ and the stable extensions of its corresponding PAF and rich PAF.

5 Conclusion

The paper has proposed a new approach for preference-based argumentation frameworks. This approach allows to encode two roles of preferences between arguments: handling critical attacks and refining the result of the evaluation. It is clearly argued in

Fig. 2. Summary



the paper that the two roles are completely independent and should be modeled in different ways and at different steps of the evaluation process. Then, we have shown that the approach is well-founded since it allows to recover very well known works on handling inconsistency in knowledge bases, namely the ones that restore the consistency of the knowledge base. Indeed, we have shown full correspondences between instances of the new PAF and respectively the preferred sub-theories defined by Brewka in [6] and the democratic sub-theories proposed by Cayrol, Royer and Saurel in [8].

References

1. L. Amgoud and C. Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *International Journal of Automated Reasoning*, 29 (2):125–169, 2002.
2. L. Amgoud and C. Cayrol. A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34:197–216, 2002.
3. L. Amgoud and S. Vesic. On the role of preferences in argumentation frameworks. In *Proceedings of the 22nd International Conference on Tools in AI, (ICTAI'10)*. IEEE, 2010.
4. T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.
5. S. Benferhat, D. Dubois, and H. Prade. Argumentative inference in uncertain and inconsistent knowledge bases. In *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, pages 411–419, 1993.
6. G. Brewka. Preferred subtheories: An extended logical framework for default reasoning. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 1043–1048, 1989.
7. C. Cayrol. On the relation between argumentation and non-monotonic coherence-based entailment. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence, Morgan Kaufmann*, pages 1443–1448, 1995.
8. C. Cayrol, V. Royer, and C. Saurel. Management of preferences in assumption-based reasoning. *Lecture Notes in Computer Science*, 682:13–22, 1993.
9. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence Journal*, 77:321–357, 1995.

10. N. Rescher and R. Manor. On inference from inconsistent premises. *Journal of Theory and decision*, 1:179–219, 1970.

Appendix

Proof of Property 1 Every set $\mathcal{E} \subseteq \mathcal{A}$ is conflict-free wrt $\mathcal{R}$ iff it is conflict-free wrt $\mathcal{R}_r$. Since extensions are conflict-free wrt $\mathcal{R}_r$, then they are conflict-free wrt $\mathcal{R}$.

Proof of Property 3 Let $\mathcal{S}$ be a democratic sub-theory. From Definition 9, $\mathcal{S}$ is consistent. Assume now that $\mathcal{S}$ is not a maximal (for set inclusion) consistent set. Thus, $\exists x \in \Sigma \setminus \mathcal{S}$ s.t. $\mathcal{S} \cup \{x\}$ is consistent. It is clear that $\mathcal{S} \cup \{x\} \succ_d \mathcal{S}$. This contradicts the fact that $\mathcal{S}$ is a democratic sub-theory.

Proof of Property 4 Let $\mathcal{E}, \mathcal{E}'$ be two stable extensions of $(\texttt{Arg}(\Sigma),\ \text{Undercut}\ , \geq)$, and let $\mathcal{E} \succeq_d \mathcal{E}'$ with $\mathcal{E} \neq \mathcal{E}'$. It is clear that $\neg(\mathcal{E} \subseteq \mathcal{E}')$ and $\neg(\mathcal{E}' \subseteq \mathcal{E})$. Let $a' \in \mathcal{E}' \setminus \mathcal{E}$ be such that $\forall a'' \in \mathcal{E}' \setminus \mathcal{E}$ it holds that $a' \geq a''$ (this is possible since $\geq$ is a total preorder). From $\mathcal{E} \succeq_d \mathcal{E}'$, we have that $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $a > a'$. This means that $\forall b' \in \mathcal{E}' \setminus \mathcal{E}$, $a > b'$. Since $\mathcal{E}'$ is a stable extension, then $\exists a'' \in \mathcal{E}'$ s.t. $a'' \mathcal{R}_r a$, i.e. $(a'' \mathcal{R} a$ and $\neg(a > a''))$ or $(a \mathcal{R} a''$ and $a'' > a)$. Sets $\mathcal{E}$ and $\mathcal{E}'$ are both conflict-free, so $a'' \in \mathcal{E}' \setminus \mathcal{E}$. Contradiction, since $\forall a'' \in \mathcal{E}' \setminus \mathcal{E}$ we have $a > a''$.

Proof of Property 6

- We show that $x \in \mathcal{S}$ iff $x \in \texttt{Base}(\texttt{Arg}(\mathcal{S}))$ where $\mathcal{S}$ is a consistent subbase of Σ. $(\Rightarrow)$ Let $x \in \mathcal{S}$. Since $\mathcal{S}$ is consistent, then the set $\{x\}$ is consistent as well. Thus, $(\{x\}, x) \in \texttt{Arg}(\mathcal{S})$. Consequently, $x \in \texttt{Base}(\texttt{Arg}(\mathcal{S}))$.
  $(\Leftarrow)$ Assume that $x \in \texttt{Base}(\texttt{Arg}(\mathcal{S}))$. Thus, $\exists a \in \texttt{Arg}(\mathcal{S})$ s.t. $x \in \texttt{Supp}(a)$. From the definition of an argument, $\texttt{Supp}(a) \subseteq \mathcal{S}$. Consequently, $x \in \mathcal{S}$.
- Let us show that the function Base is surjective. Let $\mathcal{S} \subseteq \Sigma$. From the first item of this property, the equality $\texttt{Base}(\texttt{Arg}(\mathcal{S})) = \mathcal{S}$ holds. It is clear that $\texttt{Arg}(\mathcal{S}) \in \mathcal{P}(\texttt{Arg}(\Sigma))$ ($\mathcal{P}(\texttt{Arg}(\Sigma))$ being the power set of $\texttt{Arg}(\Sigma)$).
  The following counter-example shows that the function Base is not injective: Let $\Sigma = \{x, x \to y\}$, $\mathcal{E} = \{(\{x\}, x), (\{x \to y\}, x \to y)\}$ and $\mathcal{E}' = \{(\{x\}, x), (\{x, x \to y\}, y)\}$. Since $\texttt{Base}(\mathcal{E}) = \texttt{Base}(\mathcal{E}') = \Sigma$, with $\mathcal{E} \neq \mathcal{E}'$ then Base is not injective.
- If $a \in \mathcal{E}$ where $\mathcal{E} \subseteq \texttt{Arg}(\Sigma)$, then $\texttt{Supp}(a) \subseteq \texttt{Base}(\mathcal{E})$. Consequently, $a \in \texttt{Arg}(\texttt{Base}(\mathcal{E}))$.
- Let us prove that Arg is injective. Let $\mathcal{S}, \mathcal{S}' \subseteq \Sigma$ with $\mathcal{S} \neq \mathcal{S}'$. Then, it must be that $\mathcal{S} \setminus \mathcal{S}' \neq \emptyset$ or $\mathcal{S}' \setminus \mathcal{S} \neq \emptyset$ (or both). Without loss of generality, let $\mathcal{S} \setminus \mathcal{S}' \neq \emptyset$ and let $x \in \mathcal{S} \setminus \mathcal{S}'$. If $\{x\}$ is consistent, then, $(\{x\}, x) \in \texttt{Arg}(\mathcal{S}) \setminus \texttt{Arg}(\mathcal{S}')$. Thus, $\texttt{Arg}(\mathcal{S}) \neq \texttt{Arg}(\mathcal{S}')$.
  We will now present an example that shows that this function is not surjective. Let $\Sigma = \{x, x \to y\}$ and $\mathcal{E} = \{(\{x\}, x), (\{x \to y\}, x \to y)\}$. It is clear that there exists no $\mathcal{S} \subseteq \Sigma$ s.t. $\mathcal{E} = \texttt{Arg}(\mathcal{S})$, since such a set $\mathcal{S}$ would contain Σ and, consequently, $\texttt{Arg}(\mathcal{S})$ would contain $(\{x, x \to y\}, y)$, an argument not belonging to $\mathcal{E}$.

Proof of Property 7 Let $\mathcal{S} \subseteq \Sigma$.

– Assume that $\mathcal{S}$ is consistent and $\mathtt{Arg}(\mathcal{S})$ is not conflict-free. This means that there exist $a, a' \in \mathtt{Arg}(\mathcal{S})$ s.t. a undercuts a'. From Definition 2 of undercut, it follows that $\mathtt{Supp}(a) \cup \mathtt{Supp}(a')$ is inconsistent. Besides, from the definition of an argument, $\mathtt{Supp}(a) \subseteq \mathcal{S}$ and $\mathtt{Supp}(a') \subseteq \mathcal{S}$. Thus, $\mathtt{Supp}(a) \cup \mathtt{Supp}(a') \subseteq \mathcal{S}$. Then, $\mathcal{S}$ is inconsistent. Contradiction.
– Assume now that $\mathcal{S}$ is inconsistent. This means that there exists a finite set $\mathcal{S}' = \{h_1, \ldots, h_k\}$ s.t.
  - $\mathcal{S}' \subseteq \mathcal{S}$
  - $\mathcal{S}' \vdash \perp$
  - $\mathcal{S}'$ is minimal (wrt. set inclusion) s.t. previous two items hold.

  Since $\mathcal{S}'$ is a minimal inconsistent set, then $\{h_1, \ldots, h_{k-1}\}$ and $\{h_k\}$ are consistent. Thus, $(\{h_1, \ldots, h_{k-1}\}, \neg h_k), (\{h_k\}, h_k) \in \mathtt{Arg}(\mathcal{S})$. Furthermore, those two arguments are conflicting (the former undercuts the latter). This means that $\mathtt{Arg}(\mathcal{S})$ is not conflict-free.

Proof of Theorem 1 Let $\mathcal{S}$ be a preferred sub-theory of a knowledge base Σ. Thus, $\mathcal{S}$ is consistent. From Property 7, it follows that $\mathtt{Arg}(\mathcal{S})$ is conflict-free. Assume that $\exists a \notin \mathtt{Arg}(\mathcal{S})$. Since $a \notin \mathtt{Arg}(\mathcal{S})$ and $\mathcal{S}$ is a maximal consistent subbase of Σ (according to Property 2), then $\exists h \in \mathtt{Supp}(a)$ s.t. $\mathcal{S} \cup \{h\} \vdash \perp$. Assume that $h \in \Sigma_j$. Thus, $\mathtt{Level}(\mathtt{Supp}(a)) \geq j$.

Since $\mathcal{S}$ is a preferred sub-theory of Σ, then $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j$ is a maximal (for set inclusion) consistent subbase of $\Sigma_1 \cup \ldots \cup \Sigma_j$. Thus, $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j \cup \{h\} \vdash \perp$. This means that there exists an argument $(\mathcal{S}', \neg h) \in \mathtt{Arg}(\mathcal{S})$ s.t. $\mathcal{S}' \subseteq \mathcal{S}_1 \cup \ldots \cup \mathcal{S}_j$. Thus, $\mathtt{Level}(\mathcal{S}') \leq j$. Consequently, $(\mathcal{S}', \neg h) \geq_{WLP} a$. Moreover, $(\mathcal{S}', \neg h)$ undercuts a. Thus, $(\mathcal{S}', \neg h)$ undercuts$_r a$.
The second part of the theorem follows directly from Property 6.

Proof of Theorem 2 Throughout the proof, we will use the notation $\mathcal{S}_i = \mathcal{S} \cap \Sigma_i$.

– We will first show that if $\mathcal{S} \subseteq \Sigma$, $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$ and $\mathcal{E}$ is a stable extension then $\mathcal{S} \in \mathtt{PST}$. We will suppose that $\mathcal{S} \notin \mathtt{PST}$ and we will prove that $\mathcal{E}$ is not a stable extension. If $\mathcal{S}$ is not consistent, then Property 7 implies that $\mathcal{E}$ is not conflict-free. Let us study the case when $\mathcal{S}$ is consistent but it is not a preferred subtheory. Thus, there exists $i \in \{1, \ldots, n\}$ such that $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i$ is not a maximal consistent set in $\Sigma_1, \ldots, \Sigma_i$. Let i be minimal s.t. $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i$ is not a maximal consistent set in $\Sigma_1, \ldots, \Sigma_i$. This means that there exists $x \notin \mathcal{S}$ s.t. $x \in \Sigma_i$ and $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i \cup \{x\}$ is consistent. Let $a' = (\{x\}, x)$. Since $\mathcal{E}$ is a stable extension, then $(\exists a \in \mathcal{E})$ s.t. $a \mathcal{R}_r a'$. Since $\mathcal{S}_1 \cup \ldots \cup \mathcal{S}_i \cup \{x\}$ is consistent then no argument in $\mathcal{E}$ having level at most i cannot be in conflict with a'. Thus, we have that $\nexists a \in \mathcal{E}$ s.t. $a \mathcal{R}_r a'$, which proves that $\mathcal{E}$ is not a stable extension.

– We will now prove that if $\mathcal{E} \subseteq \mathcal{A}$ is a stable extension of $(\mathcal{A}, \mathcal{R}, \geq)$ and $\mathcal{S} = \mathtt{Base}(\mathcal{E})$ then $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. Suppose the contrary. From Property 6, $\mathcal{E} \subseteq \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$, thus $\mathcal{E} \subsetneq \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$.

- Let us suppose that $\mathcal{S}$ is consistent. Since $\mathcal{S}$ is consistent, then Property 7 implies that $\texttt{Arg}(\mathcal{S})$ is conflict-free. Since we supposed that $\mathcal{E} \subsetneq \texttt{Arg}(\mathcal{S})$, then $\mathcal{E}$ is not maximal conflict-free, contradiction.
- Let us study the case when $\mathcal{S}$ is inconsistent. This means that there can be found a set $\mathcal{S}' = \{h'_1, \ldots, h'_k\}$ s.t.
  - $*$ $\mathcal{S}' \subseteq \mathcal{S}$
  - $*$ $\mathcal{S}' \vdash \perp$
  - $*$ $\mathcal{S}'$ is a minimal s.t. the previous two conditions are satisfied.

  Let us consider the set $\mathcal{E}'$ containing the following k arguments: $\mathcal{E}' = \{a'_1, \ldots, a'_k\}$, where $a'_i = (\mathcal{S}' \setminus h'_i, \neg h'_i)$. Since $(\forall h'_i \in \mathcal{S}')(\exists a \in \mathcal{E})$ s.t. $h'_i \in \texttt{Supp}(a)$ and since $\mathcal{E}$ is conflict-free then $(\nexists b \in \mathcal{E})$ s.t. $\texttt{Conc}(b) \in \{\neg h'_1, \ldots \neg h'_k\}$. Hence, $(\forall a'_i \in \mathcal{E}')$ we have that $a'_i \notin \mathcal{E}$. Formally, $\mathcal{E} \cap \mathcal{E}' = \emptyset$. This also means that, wrt. $\mathcal{R}$, no argument in $\mathcal{E}$ attacks any of arguments $a'_1, \ldots, a'_k$. Formally, $(\forall a' \in \mathcal{E}')(\nexists a \in \mathcal{E})$ s.t. $a \mathcal{R} a'$. Since $\mathcal{E}$ is a stable extension then arguments of $\mathcal{E}'$ must be attacked wrt. $\mathcal{R}_r$. We have just seen that they are not attacked wrt. $\mathcal{R}$. This means that:

  $$(\forall i \in \{1, \ldots, k\})(\exists a_i \in \mathcal{E})(a'_i \mathcal{R} a_i) \wedge (a_i > a'_i).$$

  For undercuts to exist, it is necessary that:

  $$(\forall i \in \{1, \ldots, k\})\ (h'_i \in \texttt{Supp}(a_i))\ \wedge\ (a_i > a'_i).$$

  From $(\forall i \in \{1, \ldots, k\}) a_i > a'_i$ we have $(\forall i \in \{1, \ldots, k\})$ $\texttt{Level}(\{h_i\}) \leq \texttt{Level}(\texttt{Supp}(a_i)) < \texttt{Level}(\texttt{Supp}(a'_i))$. This means that:

  $$(\forall i \in \{1, \ldots, k\})\ \texttt{Level}(\{h'_i\}) < max_{j \neq i} \texttt{Level}(\{h'_j\}).$$

  Let $l_i = \texttt{Level}(h'_i)$, for all $i \in \{1, \ldots, k\}$ and let $l_m \in \mathcal{S}'$ be s.t. $l_m = max\{l_1, \ldots, l_k\}$. Then, from the previous facts, we have:

  $$\begin{aligned} l_1 &< l_m \\ &\ldots \\ l_m &< max(\{l_1, \ldots, l_k\} \setminus \{l_m\}) \\ &\ldots \\ l_k &< l_m \end{aligned}$$

  The row m, i.e. $l_m < max(\{l_1, \ldots, l_k\} \setminus \{l_m\})$ is an obvious contradiction since we supposed that l_m is the maximal value in $\{l_1, \ldots, l_k\}$.

– Now, we have proved that:
  1. If $\mathcal{S} \subseteq \Sigma$, $\mathcal{E} = \texttt{Arg}(\mathcal{S})$ and $\mathcal{E}$ is a stable extension, then $\mathcal{S} \in \texttt{PST}$,
  2. If $\mathcal{E}$ is a stable extension then $\mathcal{E} = \texttt{Arg}(\texttt{Base}(\mathcal{E}))$.

  Let $\mathcal{E}$ be a stable extension and let $\mathcal{S} = \texttt{Base}(\mathcal{E})$. Then, from (2), $\mathcal{E} = \texttt{Arg}(\mathcal{S})$. From (1), $\mathcal{S} \in \texttt{PST}$.

Proof of Theorem 3

– Theorem 1 shows that $\texttt{Arg}(\texttt{PST}) \subseteq \texttt{Ext}$.

– Property 6 implies that $\mathtt{Arg}$ is injective.
– Let $\mathcal{E} \in \mathtt{Ext}$ and let $\mathcal{S} = \mathtt{Base}(\mathcal{E})$. From Theorem 2, we have $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. Theorem 2 yields also the conclusion that $\mathcal{S} \in \mathtt{PST}$. Thus, $\mathtt{Arg} : \mathtt{PST} \rightarrow \mathtt{Ext}$ is surjective.

Proof of Theorem 4 Let $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$ and let $x \rhd x'$ iff $x \trianglerighteq x'$ and not $x' \trianglerighteq x$. From Property 7, we see that $\mathcal{E}$ is conflict-free. We will prove that it attacks (wrt. $\mathcal{R}_r$) any argument in its exterior. Let $a' \in \mathcal{A} \setminus \mathcal{E}$ be an arbitrary argument. Since $a' \notin \mathcal{E}$ then $\exists h' \in \mathtt{Supp}(a')$ s.t.$h' \notin \mathcal{S}$. From $\mathcal{S} \in \mathtt{DMS}(\Sigma)$ we have that $\mathcal{S}$ is a maximal consistent set. It is clear that $\mathcal{S} \cup \{h'\} \vdash \perp$. Let us identify all its minimal conflicting subsets. Formally, let $C_1, \ldots, C_k$ be all sets which satisfy the following three conditions:

1. $C_i \subseteq \mathcal{S}$
2. $C_i \cup \{h'\} \vdash \perp$
3. C_i is minimal (wrt. set inclusion) s.t. the two previous conditions are satisfied.

Those sets allow to construct the following k arguments: $a_1 = (C_1, \neg h'), \ldots, a_k = (C_k, \neg h)$. It is obvious that all of them attack a' wrt. $\mathcal{R}$. If at least one of them attack a' wrt. $\mathcal{R}_r$, then the proof is over. Suppose the contrary. This would mean that $\forall i \in \{1, \ldots, k\}$, $a' > a_i$. Thus, $(\forall i \in \{1, \ldots, k\})$ $(\exists h_i \in C_i)$ s.t. $h' \rhd h_i$. In other words, for every argument a_i, there exists one formula $h_i \in \mathtt{Supp}(a_i)$, such that $h' \rhd h_i$. Let $H = \{h_1, \ldots, h_k\}$.

Now, we can define a set $\mathcal{S}'$ as follows: $\mathcal{S}' = \mathcal{S} \cup \{h'\} \setminus H$. We will show that $\mathcal{S}'$ is consistent. Suppose the contrary. Since $\mathcal{S}$ is consistent, then any inconsistent subset of $\mathcal{S}'$ must contain h'. Let $K_1, \ldots, K_j$ be all sets which satisfy the following conditions:

1. $K_i \subseteq \mathcal{S}' \setminus \{h'\}$
2. $K_i \cup \{h'\} \vdash \perp$
3. K_i is a minimal set s.t. the previous two conditions hold.

Let $K = \{K_1, \ldots, K_j\}$ and $C = \{C_1, \ldots, C_k\}$. It is easy to see that $K \subseteq C$ (this follows immediately from the fact that $\mathcal{S}' \setminus \{h'\} \subseteq \mathcal{S}$). Furthermore, since $(\forall C_i \in C)$ $(\exists h \in H)$ s.t. $h \in C_i$ then $(\forall K_i \in K)$ $(\exists h \in H)$ s.t. $h \in K_i$. Since for all K_i, we have that $K_i \cap H = \emptyset$ then it must be that $j = 0$, i.e. $K = \emptyset$. In other words, there are no inconsistent subsets of S', which means that S' is consistent.

We can notice that $\mathcal{S}' \setminus \mathcal{S} = \{h'\}$ and $\mathcal{S} \setminus \mathcal{S}' = \{h_1, \ldots, h_k\}$. Since $\mathcal{S}'$ is consistent, we see that $\mathcal{S}' \succ \mathcal{S}$. Contradiction with $\mathcal{S} \in \mathtt{DMS}(\Sigma)$.

Proof of Theorem 5 Let $\mathcal{S} = \mathtt{Base}(\mathcal{E})$.

– Let us suppose that $\mathcal{S}$ is consistent but that it is not a maximal consistent set. This means that $\exists h \in \Sigma \setminus \mathcal{S}$ s.t. $\mathcal{S} \cup \{h\}$ is consistent. From Property 7, $\mathcal{E}' = \mathtt{Arg}(\mathcal{S} \cup \{h\})$ is consistent. From Property 6, $\mathcal{E} \subseteq \mathcal{E}'$. The same result implies that $\mathcal{E} \neq \mathcal{E}'$. Thus, $\mathcal{E} \subsetneq \mathcal{E}'$, which means that $\mathcal{E}$ is not a maximal conflict-free set. Contradiction with the fact that $\mathcal{E}$ is a stable extension.
– Suppose now that $\mathcal{S}$ is inconsistent. This means that there can be found a set $S' = \{h'_1, \ldots, h'_k\}$ s.t.

- $\mathcal{S}' \subseteq \mathcal{S}$
- $\mathcal{S}' \vdash \bot$
- $\mathcal{S}'$ is a minimal s.t. the previous two conditions are satisfied.

Let us consider the set $\mathcal{E}'$ containing the following k arguments: $\mathcal{E}' = \{a'_1, \ldots, a'_k\}$, where $a'_i = (\mathcal{S}' \setminus h'_i, \neg h'_i)$. Since $(\forall h'_i \in \mathcal{S}')(\exists a \in \mathcal{E})$ s.t. $h'_i \in \mathtt{Supp}(a)$ and since $\mathcal{E}$ is conflict-free then $(\nexists b \in \mathcal{E})$ s.t. $\mathtt{Conc}(b) \in \{\neg h'_1, \ldots \neg h'_k\}$. Hence, $(\forall a'_i \in \mathcal{E}')$ we have that $a'_i \notin \mathcal{E}$. Formally, $\mathcal{E} \cap \mathcal{E}' = \emptyset$. This also means that, wrt. $\mathcal{R}$, no argument in $\mathcal{E}$ attacks any of arguments $a'_1, \ldots, a'_k$. Formally, $(\forall a' \in \mathcal{E}')(\nexists a \in \mathcal{E})$ s.t. $a \mathcal{R} a'$. Since $\mathcal{E}$ is a stable extension then arguments of $\mathcal{E}'$ must be attacked wrt. $\mathcal{R}_r$. We have just seen that they are not attacked wrt. $\mathcal{R}$. This means that:

$$(\forall i \in \{1, \ldots, k\})(\exists a_i \in \mathcal{E})(a'_i \mathcal{R} a_i) \wedge (a_i > a'_i).$$

For undercuts to exist, it is necessary that:

$$(\forall i \in \{1, \ldots, k\}) \ (h'_i \in \mathtt{Supp}(a_i)) \ \wedge \ (a_i > a'_i).$$

For $i = 1$, we have: $\exists i_1 \in \{1, \ldots, k\}$ s.t. $h'_1 \rhd h'_{i_1}$. For $i = i_1$, we have that $\exists i_2 \in \{1, \ldots, k\}$ s.t. $h'_{i_1} \rhd h'_{i_2}$, thus, $h'_1 \rhd h'_{i_1} \rhd h'_{i_2}$. After k consecutive applications of the same rule, we obtain: $h'_1 \rhd h'_{i_1} \rhd \ldots \rhd h'_{i_k}$. It is clearly a contradiction since on one hand, all the formulae in the chain are different because of the strict preference between them, and, on the other hand, set $\{h'_1, \ldots, h'_k\}$ contains k formulae, thus at least two of them in a chain of $k+1$ formulae must coincide.

This ends the first part of the proof. Let us now prove that $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. From Property 6, we have that $\mathcal{E} \subseteq \mathtt{Arg}(\mathcal{S})$. Suppose that $\mathcal{E} \subsetneq \mathtt{Arg}(\mathcal{S})$. In the first part of the proof, we have showed that $\mathcal{S}$ is a maximal consistent set. Thus, from Property 7, we have that $\mathtt{Arg}(\mathcal{S})$ is conflict-free. This simply means that $\mathcal{E}$ is not a maximal conflict-free set, contradiction.

Proof of Theorem 6 ($\Rightarrow$) Let $\mathcal{S} \succeq_d \mathcal{S}'$. Let $a' \in \mathcal{E}' \setminus \mathcal{E}$. Then $\exists h' \in \mathtt{Supp}(a')$ s.t. $h' \in \mathcal{S}' \setminus \mathcal{S}$. Since $\mathcal{S} \succeq_d \mathcal{S}'$ then $\exists h \in \mathcal{S} \setminus \mathcal{S}'$ s.t. $h \rhd h'$. Let $a = (\{h\}, h)$. It is clear that $a \in \mathcal{S} \setminus \mathcal{S}'$ and $a > a'$. Thus, $\mathcal{E} \succeq_d \mathcal{E}'$.
($\Leftarrow$) Let $\mathcal{E} \succeq_d \mathcal{E}'$. Let $h' \in \mathcal{S}' \setminus \mathcal{S}$. Then $a' = (\{h'\}, h') \in \mathcal{E}' \setminus \mathcal{E}$. Thus, $\exists a \in \mathcal{E} \setminus \mathcal{E}'$ s.t. $a > a'$. Since $a \in \mathcal{E} \setminus \mathcal{E}'$, then $\exists h \in \mathtt{Supp}(a)$ s.t. $h \in \mathcal{S} \setminus \mathcal{S}'$. It is clear that $h \rhd h'$.

Proof of Theorem 7

- From Theorem 4, we have that $\mathcal{E}$ is an extension of a basic PAF $(\mathcal{A}, \mathcal{R}, \geq)$. We will prove that it is also an extension of a rich PAF $(\mathcal{A}, \mathcal{R}, \geq, \succeq_d)$. Let us suppose the contrary, i.e. suppose that there exists $\mathcal{E}'$ s.t. $\mathcal{E}'$ is a stable extension and $\mathcal{E}' \succ_d \mathcal{E}$. Let $\mathcal{S}' = \mathtt{Base}(\mathcal{E}')$. From Theorem 5, $\mathcal{E}' = \mathtt{Arg}(\mathcal{S}')$. From the same theorem, we have that $\mathcal{S}'$ is maximal consistent set and from Theorem 6 that $\mathcal{S}' \succ_d \mathcal{S}$. Contradiction.
- Theorem 5 implies that $\mathcal{S}$ is a maximal conflict-free set and that $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. Suppose that $\mathcal{S} \notin \mathtt{DMS}(\Sigma)$. This means that $\exists \mathcal{S}' \subseteq \Sigma$ s.t. $\mathcal{S}' \in \mathtt{DMS}(\Sigma)$ and $\mathcal{S}' \succ_d \mathcal{S}$. From Theorem 4, $\mathcal{E}' = \mathtt{Arg}(\mathcal{S}')$ is a stable extension of a basic PAF. Theorem 6 implies that $\mathcal{E}' \succ_d \mathcal{E}$, contradiction.

Proof of Theorem 8 Let us denote $\mathtt{Ext}(\mathcal{T})$ the set of all extensions of a rich PAF $\mathcal{T}$. We will prove that $\mathtt{Arg} : \mathtt{DMS} \rightarrow \mathtt{Ext}(\mathcal{T})$ is a bijection.

- Theorem 7 shows that $\mathtt{Arg}(\mathtt{DMS}) \subseteq \mathtt{Ext}(\mathcal{T})$.
- Property 6 implies that $\mathtt{Arg}$ is injective.
- Let $\mathcal{E} \in \mathtt{Ext}(\mathcal{T})$ and let $\mathcal{S} = \mathtt{Base}(\mathcal{E})$. From Theorem 5, we have $\mathcal{E} = \mathtt{Arg}(\mathcal{S})$. Theorem 7 yields the conclusion that $\mathcal{S} \in \mathtt{DMS}$. Thus, $\mathtt{Arg} : \mathtt{DMS} \rightarrow \mathtt{Ext}$ is surjective.