Maxi-Consistent Operators in Argumentation

Srdjan Vesic[1,2]

Abstract. This paper studies an instantiation of Dung-style argumentation system with classical propositional logic. Our goal is to explore the link between the result obtained by using argumentation to deal with an inconsistent knowledge base and the result obtained by using maximal consistent subsets of the same knowledge base. Namely, for a given attack relation and semantics, we study the question: does every extension of the argumentation system correspond to exactly one maximal consistent subset of the knowledge base? We study the class of attack relations which satisfy that condition. We show that such a relation must be conflict-dependent, must not be valid, must not be conflict-complete, must not be symmetric etc. Then, we show that some attack relations serve as lower or upper bounds with respect to the condition we study (e.g. we show that if an attack relation contains "canonical undercut" then it does not satisfy this condition). By using our results, we show for each attack relation and each semantics whether or not they satisfy the aforementioned condition. Finally, we interpret our results and discuss more general questions, like does (and when) this link is a desirable property. This work will help us obtain our long-term goal, which is to better understand the role of argumentation and, more particularly, the expressivity of logic-based instantiations of Dung-style argumentation frameworks.

1 INTRODUCTION

Argumentation is a reasoning model based on the generation and evaluation of arguments. There are a number of proposals for defining a computational model of argument [6, 16]. Nowadays, most of the work in argumentation is based on the abstract argumentation theory proposed by Dung [11]. This paper studies the case when Dung's framework is instantiated by building arguments in classical propositional logic (e.g. [4, 12]). The advantages of such an approach are that it can benefit from all the results proved in the case of abstract argumentation, and in addition, it allows for using some important and useful notions from the underlying logic like consistency, inference, logical equivalence, etc.

One of the most important features of argumentation is that it can be used as a tool for *inconsistency handling*. Indeed, in many real-world scenarios, a knowledge base from which arguments are constructed is inconsistent. If we are given an inconsistent set of classical propositional formulae Σ, classical inference relation is useless, since from an inconsistent set of formulae, we can draw any conclusion. Thus, a natural and well-known way to deal with this problem is to identify *maximal (for set inclusion) consistent subsets* of Σ and to consider them as possible (and mutually exclusive) consistent points of view. Those maximal consistent sets represent a consistent output given an inconsistent knowledge base at the input.

Given a set of formulae Σ, let us consider the corresponding argumentation framework $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$, where for a set $S \subseteq \Sigma$, we denote by $\mathtt{Arg}(S)$ the set of all arguments that can be built from S, and by $\mathcal{R}$ the relation used for identifying attacks between arguments. The most important output of an argumentation system is the set of its *extensions*, which are the sets of mutually acceptable arguments. From an argumentation system containing conflicts, representing some kind of inconsistency in the underlying logic, we obtain extensions, that are conflict-free sets for all well-known argumentation semantics.

There are works [1, 2, 5, 9, 10, 12] which study the link between the result obtained by using an argumentation system and the notions of the underlying logic, like consistency, inference, etc. This paper continues that line of research, since its goal is also to study the link between a knowledge base and the corresponding argumentation system.

We have seen that the most common way to deal with inconsistency in a knowledge base is to identify its maximal consistent subsets, whereas for an argumentation system, this is done by calculating its extensions. We formalise the main research task of this paper as follows: what is the link between maximal consistent subsets of Σ and extensions of $\mathcal{F}$? More particularly: **When is the function Arg a bijection between maximal consistent subsets of Σ and extensions of $\mathcal{F}$?**

If an attack relation satisfies the previous condition, then we will say that it satisfies ($\mathtt{MC} \leftrightarrow \mathtt{Ext}$).

In this paper, we study the class of attack relations which satisfy this condition. First, we study the properties of those relations and identify several conditions they must satisfy. For example, we show that if an attack relation $\mathcal{R}$ satisfies the condition we study, then: the base of any extension must be consistent; $\mathcal{R}$ must be conflict-dependent; $\mathcal{R}$ must not be valid; $\mathcal{R}$ must not be symmetric, ...

After studying general properties such a relation must (not) satisfy, we examine virtually all attack relations used in logic-based argumentation. We show that some of them serve as lower or upper bounds w.r.t. the condition we study, e.g. we show that if an attack relation is more specific than "canonical undercut" (i.e. if it contains "canonical undercut") then it does not satisfy ($\mathtt{MC} \leftrightarrow \mathtt{Ext}$). Then, we analyse all those attack relations and for each of them prove whether or not it satisfies the aforementioned condition, by using the results from the first part of the paper. At the end, we interpret the results and discuss more general questions, like is satisfying (or not satisfying) the previous condition a good or a bad thing for an attack relation (and a semantics) and what does it say about a propositional logic instantiation of Dung's argumentation framework?

[1] The author was funded by the National Research Fund, Luxembourg. His work on this paper was carried out during the tenure of an ERCIM "Alain Bensoussan" Fellowship Programme. This Programme is supported by the Marie Curie Co-funding of Regional, National and International Programmes (COFUND) of the European Commission.

[2] Computer Science and Communication, University of Luxembourg, Luxembourg, email: srdjan.vesic@uni.lu

2 BASICS OF ARGUMENTATION

We will use classical propositional logic for instantiating Dung's argumentation system. $\mathcal{L}$ denotes the set of well-formed formulae, $\vdash$ stands for classical entailment, and $\equiv$ for logical equivalence. We denote by Σ a finite set of classical propositional formulae from which arguments are constructed. We will use the notation $\texttt{MC}(\Sigma)$ for the set of all maximal consistent subsets of Σ. A logical argument is most commonly defined as a pair $(support, conclusion)$.

Definition 1 (Argument) *An **argument** is a pair (Φ, α) such that $\Phi \subseteq \Sigma$ is a minimal (for set inclusion) consistent set of formulae such that $\Phi \vdash \alpha$.*

For an argument $a = (\Phi, \alpha)$, we will use the function $\texttt{Supp}(a) = \Phi$ to denote its support and $\texttt{Conc}(a) = \alpha$ to denote its conclusion. For a given set of formulae S, we denote by $\texttt{Arg}(S)$ the set of arguments constructed from S. Formally, $\texttt{Arg}(S) = \{a \mid a \text{ is an argument and } \texttt{Supp}(a) \subseteq S\}$. Let $\texttt{Arg}(\mathcal{L})$ denote the set of all arguments that can be made from propositional logic. For a given set of arguments $\mathcal{E}$, we denote $\texttt{Base}(\mathcal{E}) = \bigcup_{a \in \mathcal{E}} \texttt{Supp}(a)$. We suppose that function $\texttt{Arg}$ is defined on $\mathcal{L}$ and that function $\texttt{Base}$ is defined on $\texttt{Arg}(\mathcal{L})$; by slightly abusing the notation, we will sometimes write $\texttt{Arg}$ (respectively $\texttt{Base}$) for the restriction of these functions on any set of formulae (respectively arguments).

Definition 2 (Argumentation system) *An argumentation system (AS) is a pair $(\mathcal{A}, \mathcal{R})$ where $\mathcal{A} \subseteq \texttt{Arg}(\mathcal{L})$ is a set of arguments and $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ a binary relation. For each pair $(a, b) \in \mathcal{R}$, we say that a attacks b. We will also sometimes use notation $a\mathcal{R}b$ instead of $(a, b) \in \mathcal{R}$.*

Since arguments are built from formulae, we suppose that an attack relation is to be defined by specifying a condition s.t. for any two arguments a and b, we have that a attacks b if and only if the condition from the definition of attack relation is satisfied. For example, such a condition can be that the conclusion of a is logically equivalent to the negation of the conclusion of b. We suppose that all attack relations are defined on the set $\texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$, and that for any set $\mathcal{A} \subseteq \texttt{Arg}(\mathcal{L})$, we use the restriction of the attack relation on the set $\mathcal{A} \times \mathcal{A}$. That is why, in order to simplify notation, we will simply write $\mathcal{R}$ for an attack relation defined on the set $\texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$ as well as for the restriction of that attack relation on any set $\mathcal{A} \times \mathcal{A}$, with $\mathcal{A} \subseteq \texttt{Arg}(\mathcal{L})$.

In order to determine mutually acceptable sets of arguments, different semantics have been introduced in argumentation. We first introduce the basic notions of conflict-freeness and defence.

Definition 3 (Conflict-free, defence) *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AS, $\mathcal{E} \subseteq \mathcal{A}$ and $a \in \mathcal{A}$.*

- *$\mathcal{E}$ is* conflict-free *iff $\nexists a, b \in \mathcal{E}$ s.t. $a \mathcal{R} b$*
- *$\mathcal{E}$ defends a iff for every $b \in \mathcal{A}$ we have that if $b \mathcal{R} a$ then there exists $c \in \mathcal{E}$ s.t. $c \mathcal{R} b$.*

Definition 4 (Acceptability semantics) *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AS and $\mathcal{B} \subseteq \mathcal{A}$. We say that a set $\mathcal{B}$ is* admissible *iff it is conflict-free and defends all its elements.*

- *$\mathcal{B}$ is a* complete *extension iff $\mathcal{B}$ defends all its arguments and contains all the arguments it defends.*
- *$\mathcal{B}$ is a* preferred *extension iff it is a maximal (w.r.t. set inclusion) admissible set.*
- *$\mathcal{B}$ is a* stable *extension iff $\mathcal{B}$ is conflict-free and for all $a \in \mathcal{A} \setminus \mathcal{B}$, there exists $b \in \mathcal{B}$ such that $b \mathcal{R} a$.*
- *$\mathcal{B}$ is a* semi-stable *extension iff $\mathcal{B}$ is a complete extension and the union of the set $\mathcal{B}$ and the set of all arguments attacked by $\mathcal{B}$ is maximal (for set inclusion).*
- *$\mathcal{B}$ is a* grounded extension *iff $\mathcal{B}$ is a minimal (for set inclusion) complete extension.*
- *$\mathcal{B}$ is an* ideal *extension iff $\mathcal{B}$ is a maximal (for set inclusion) admissible set contained in every preferred extension.*

For an argumentation system $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ we will denote by $\texttt{Ext}_x(\mathcal{F})$; or, by a slight abuse of notation, by $\texttt{Ext}_x(\mathcal{A}, \mathcal{R})$ the set of its extensions w.r.t. semantics x. We will use abbreviations c, p, s, ss, g and i for respectively complete, preferred, stable, semi-stable, grounded and ideal semantics. For example, $\texttt{Ext}_p(\mathcal{F})$ denotes the set of preferred extensions argumentation system $\mathcal{F}$.

3 DEFINING THE PROBLEM

Until now, we have specified how to construct an argumentation system from a knowledge base Σ. Namely, given Σ, construct the corresponding argumentation system $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$, and then, using a chosen semantics, calculate extensions. Note that the only component which we did not specify until now is the attack relation. That is because in logical argumentation, different attack relations are used.

As already mentioned, the goal of this paper is to study the link between the result obtained from the set of formulae Σ and the result obtained from an argumentation system whose arguments are built using the formulae from Σ. Since all the other components of our system except semantics and attack relation are fixed, it is obvious that this link depends exclusively on those two components. Let us now formally define a criterion we will study throughout the article.

Definition 5 (MC $\leftrightarrow$ Ext) *Let x be any argumentation semantics. We say that attack relation $\mathcal{R}$ satisfies $(\texttt{MC} \leftrightarrow \texttt{Ext}_x)$ if and only if for every finite set of propositional formulae Σ we have that*

$$\texttt{Arg} \textit{ is a bijection between } \texttt{MC}(\Sigma) \textit{ and } \texttt{Ext}_x(\texttt{Arg}(\Sigma), \mathcal{R})$$

This means that for every $S \in \texttt{MC}(\Sigma)$, $\texttt{Arg}(S) \in \texttt{Ext}(\texttt{Arg}(\Sigma), \mathcal{R})$ and for every $\mathcal{E} \in \texttt{Ext}(\texttt{Arg}(\Sigma), \mathcal{R})$, there exists $S \in \texttt{MC}(\Sigma)$ such that $\mathcal{E} = \texttt{Arg}(S)$. In other words, we ask for the function $\texttt{Arg}$ to be a bijection between the set of maximal conflict-free subsets of Σ and extensions of the corresponding argumentation system w.r.t. the given attack relation and semantics. For example, we will say that an attack relation $\mathcal{R}$ satisfies $(\texttt{MC} \leftrightarrow \texttt{Ext}_c)$ if and only if for any finite Σ, we have that $\texttt{Arg}$ is a bijection between $\texttt{MC}(\Sigma)$ and $\texttt{Ext}_c(\texttt{Arg}(\Sigma), \mathcal{R})$. Sometimes, when it is clear from the context which semantics we refer to or when a semantics is not important, we will use the simplified notation $(\texttt{MC} \leftrightarrow \texttt{Ext})$. We say that an attack relation $\mathcal{R}$ falsifies $(\texttt{MC} \leftrightarrow \texttt{Ext}_x)$ if and only if $\mathcal{R}$ does not satisfy $(\texttt{MC} \leftrightarrow \texttt{Ext}_x)$.

There are two ways to study the link between a knowledge base and the corresponding argumentation system. In the *first scenario*, one starts with the knowledge base Σ and constructs all the arguments. Then, the link between Σ and $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$ is studied. In this case, we call a corresponding argumentation system a *complete system*. Complete systems are infinite, but for every complete system, there exists an equivalent finite system [3]. In the *second scenario*, we start with an argumentation system $\mathcal{F} = (\mathcal{A}, \mathcal{R})$, we *define* $\Sigma \overset{\text{def}}{=} \texttt{Base}(\mathcal{A})$, and then study the link between Σ and $\mathcal{F}$. In this case, we deal with an *incomplete system*. There is an important difference between those two scenarios. Namely, in the first case, all the arguments that can be built from Σ are considered when calculating $\mathtt{Ext}_x(\mathcal{F})$. In the second case, Σ contains all the formulae from $\mathcal{A}$, but in $\mathcal{F}$, not all formulae are equally represented. It is clear that in the second scenario, one cannot expect Arg to be a bijection between $\mathtt{MC}(\Sigma)$ and $\mathtt{Ext}(\mathcal{A}, \mathcal{R})$.

Due to the space restrictions, we do not further discuss the differences between complete and incomplete systems; that is not the goal of this paper. However, we find it necessary to point out that they exist, in order to make the context of our research question clear. In the second scenario, it is not reasonable to expect any correlation between the result obtained directly from Σ and from $\mathcal{F}$. That is why, in the rest of the paper we will suppose the first scenario, and study the properties of attack relations (and semantics) which satisfy $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$.

4 PROPERTIES OF RELATIONS SATISFYING (MC↔EXT)

In this section, we analyse properties of attack relations satisfying $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$. We first show that if this condition is satisfied, then the function $\mathtt{Base} : \mathtt{Ext}(\mathcal{F}) \to \mathtt{MC}(\Sigma)$ is the inverse function of the function $\mathtt{Arg} : \mathtt{MC}(\Sigma) \to \mathtt{Ext}(\mathcal{F})$.

Proposition 1 *Let $\mathcal{R}$ be an attack relation and x an acceptability semantics. If $\mathcal{R}$ satisfies $(\mathtt{MC} \leftrightarrow \mathtt{Ext}_x)$ then:*

- *for every $S \in \mathtt{MC}(\Sigma)$, we have that $S = \mathtt{Base}(\mathtt{Arg}(S))$,*
- *for every $\mathcal{E} \in \mathtt{Ext}_x(\mathcal{F})$, we have that $\mathcal{E} = \mathtt{Arg}(\mathtt{Base}(\mathcal{E}))$.*

The previous result allows to easily show that if an attack relation satisfies $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$, then every extension has a consistent base and the union of its arguments' conclusions is consistent.

Corollary 1 *Let $\mathcal{R}$ be an attack relation and x a semantics. Let $\mathcal{R}$ satisfy $(\mathtt{MC} \leftrightarrow \mathtt{Ext}_x)$ and let Σ be a finite set of formulae. Denote $\mathcal{F} = (\mathtt{Arg}(\Sigma), \mathcal{R})$. Then, for every $\mathcal{E} \in \mathtt{Ext}_x(\mathcal{F})$, we have:*

- $\mathtt{Base}(\mathcal{E})$ *is consistent*
- $\bigcup_{a \in \mathcal{E}} \mathtt{Conc}(a)$ *is consistent*

Note that we can use the previous result to show that an attack relation does *not* satisfy $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$. Namely, if an attack relation returns extensions having inconsistent bases, then it violates $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$.

Corollary 2 *Let $\mathcal{R}$ be an attack relation, and x an acceptability semantics. If there exists a finite knowledge base Σ s.t. there exists an extension $\mathcal{E} \in \mathtt{Ext}_x(\mathtt{Arg}(\Sigma), \mathcal{R})$ s.t. $\mathtt{Base}(\mathcal{E})$ is inconsistent, then $\mathcal{R}$ does not satisfy $(\mathtt{MC} \leftrightarrow \mathtt{Ext}_x)$.*

4.1 On conflict-dependence and validity

In this subsection, we study the link between satisfying $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$ and being conflict-dependent or valid. An attack relation is *conflict-dependent* if whenever an argument attacks another one, the union of their supports is inconsistent [1].

Definition 6 (Conflict-dependent) *Let $\mathcal{R} \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ be an attack relation. We say that $\mathcal{R}$ is conflict-dependent iff for every $a, b \in \mathtt{Arg}(\mathcal{L})$, if $(a, b) \in \mathcal{R}$ then $\mathtt{Supp}(a) \cup \mathtt{Supp}(b) \vdash \bot$.*

We will now prove that conflict-dependence is a necessary condition for satisfying $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$. To be completely precise, we here specify that we will say that a *semantics x returns conflict-free sets* iff for any argumentation system $(\mathcal{A}, \mathcal{R})$, for any $\mathcal{E} \in \mathtt{Ext}_x(\mathcal{A}, \mathcal{R})$, it holds that $\mathcal{E}$ is conflict-free w.r.t. $\mathcal{R}$.

Proposition 2 *Let $\mathcal{R}$ be an attack relation and x any semantics which returns conflict-free sets. If $\mathcal{R}$ satisfies $(\mathtt{MC} \leftrightarrow \mathtt{Ext}_x)$, then $\mathcal{R}$ is conflict-dependent.*

Having proved this, we know that a relation satisfying $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$ enjoys all the properties of conflict-dependent relations. For example, it was shown that if an attack relation is conflict-dependent, then there are no self-attacking arguments [1].

Corollary 3 *Let $\mathcal{R}$ be an attack relation and x any semantics which returns conflict-free sets. If $\mathcal{R}$ satisfies $(\mathtt{MC} \leftrightarrow \mathtt{Ext}_x)$ then for any argument $a \in \mathtt{Arg}(\mathcal{L})$, we have that s.t. $(a, a) \notin \mathcal{R}$.*

This means that we have another way to identify (some of the) attack relations not satisfying $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$: namely, if for an attack relation there exists a self-attacking argument, then the given attack relation falsifies $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$ for all semantics returning conflict-free sets. Let us now study the notion of validity [2].

Definition 7 (Valid) *Let $\mathcal{R} \subseteq \mathtt{Arg}(\mathcal{L}) \times \mathtt{Arg}(\mathcal{L})$ be an attack relation. We say that $\mathcal{R}$ is valid iff for every $B \subseteq \mathtt{Arg}(\mathcal{L})$ it holds that if B is conflict-free, then $\mathtt{Base}(B)$ is consistent.*

We show that this property is incompatible with conflict-dependence.

Proposition 3 *There exists no attack relation which is both conflict-dependent and valid.*

This means that if an attack relation $\mathcal{R}$ satisfies $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$ then there must exist a set $\mathcal{E}$ which is conflict-free w.r.t. $\mathcal{R}$ but whose base is inconsistent.

Corollary 4 *Let $\mathcal{R}$ be an attack relation and x an acceptability semantics returning conflict-free sets and let $\mathcal{R}$ satisfy $(\mathtt{MC} \leftrightarrow \mathtt{Ext}_x)$. Then, $\mathcal{R}$ is not valid.*

The previous result is useful since if an attack relation is valid, we can immediately conclude that it violates $(\mathtt{MC} \leftrightarrow \mathtt{Ext}_x)$ for all well-known semantics. On the more general level, note that we see that asking for every conflict-free set to have a consistent base is very demanding. Roughly speaking, this is due the fact that attacks are binary whereas minimal conflicts may be ternary (or of a greater cardinality). An alternative condition to be studied in the context where validity was introduced would be to ask for every *admissible* set (or for every extension under a given semantics) to have a consistent base.

4.2 Satisfying (MC↔Ext) and different acceptability semantics

In this subsection we study the properties related to particular semantics. We show that if an attack relation satisfies $(\mathtt{MC} \leftrightarrow \mathtt{Ext})$ for stable semantics, then it also satisfies it for semi-stable semantics. Then we identify conditions under which an attack relation satisfies

($\texttt{MC} \leftrightarrow \texttt{Ext}$) for stable semantics. We provide a similar result for preferred semantics. We also identify a sufficient condition so that an attack relation falsifies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) under complete semantics. Then, we discuss the case of single-extension semantics.

First, suppose that $\mathcal{R}$ satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$). This means that for any finite set of formulae Σ, function $\texttt{Arg}$ is a bijection between $\texttt{MC}(\Sigma)$ and $\texttt{Ext}_s(\texttt{Arg}(\Sigma), \mathcal{R})$. Since every finite set of formulae has at least one maximal consistent subset, then for every Σ, it must be that $(\texttt{Arg}(\Sigma), \mathcal{R})$ has at least one stable extension. Since there are stable extensions, then stable and semi-stable semantics coincide [8]. Thus, we obtain the following proposition.

Proposition 4 *Let $\mathcal{R}$ be an attack relation. If $\mathcal{R}$ satisfies* ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$) *then:*

- *for every finite set of formulae Σ and $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$, we have that $\texttt{Ext}_s(\mathcal{F}) = \texttt{Ext}_{ss}(\mathcal{F})$*
- *$\mathcal{R}$ satisfies* ($\texttt{MC} \leftrightarrow \texttt{Ext}_{ss}$).

Let us now prove that in the case of stable semantics, if the image w.r.t. $\texttt{Arg}$ of every maximal consistent set is an extension and if the base of every extension is consistent, then the attack relation in question satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$).

Proposition 5 *Let $\mathcal{R}$ be an attack relation. If for every set of formulae Σ and $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$, we have:*

- *for all $S \in \texttt{MC}(\Sigma)$, $\texttt{Arg}(S) \in \texttt{Ext}_s(\mathcal{F})$, and*
- *for all $\mathcal{E} \in \texttt{Ext}_s(\mathcal{F})$, $\texttt{Base}(\mathcal{E})$ is consistent*

then $\mathcal{R}$ satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$).

We will prove that similar two conditions are sufficient to guarantee that $\mathcal{R}$ satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) under *preferred* semantics.

Proposition 6 *Let $\mathcal{R}$ be an attack relation. If for every set of formulae Σ and $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$, we have:*

- *for all $S \in \texttt{MC}(\Sigma)$, $\texttt{Arg}(S) \in \texttt{Ext}_p(\mathcal{F})$, and*
- *for all $\mathcal{E} \in \texttt{Ext}_p(\mathcal{F})$, $\texttt{Base}(\mathcal{E})$ is consistent*

then $\mathcal{R}$ satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}_p$).

As a consequence of the two previous results, we can identify a sufficient condition so that $\mathcal{R}$ satisfies both ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$) and ($\texttt{MC} \leftrightarrow \texttt{Ext}_p$).

Corollary 5 *Let $\mathcal{R}$ be an attack relation. If for every set of formulae Σ and $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$, we have:*

- *for all $S \in \texttt{MC}(\Sigma)$, $\texttt{Arg}(S) \in \texttt{Ext}_s(\mathcal{F})$, and*
- *for all $\mathcal{E} \in \texttt{Ext}_p(\mathcal{F})$, $\texttt{Base}(\mathcal{E})$ is consistent*

then $\mathcal{R}$ satisfies both ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$) *and* ($\texttt{MC} \leftrightarrow \texttt{Ext}_p$).

Let us now show that if an attack relation returns a stable extension having an inconsistent base, then it violates ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for stable, semi-stable, preferred and complete semantics.

Proposition 7 *Let $\mathcal{R}$ be an attack relation. If there exists a finite set of formulae Σ s.t. $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$ has a stable extension $\mathcal{E}$ s.t. $\texttt{Base}(\mathcal{E})$ is inconsistent, then $\mathcal{R}$ falsifies* ($\texttt{MC} \leftrightarrow \texttt{Ext}_x$) *for $x \in \{s, ss, p, c\}$.*

Let us now study the case of complete semantics. We will show that it is not possible for an attack relation to satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}_c$). The only condition we use in our result is that for any argument a, if a has a formula φ in its support, and $\neg\varphi \in \Sigma$, then there exists an argument $b \in \texttt{Arg}(\Sigma)$ such that b attacks a.

Proposition 8 *Let $\mathcal{R}$ be an attack relation s.t. for every finite set of formulae Σ, for every $a \in \texttt{Arg}(\Sigma)$, for every $\varphi \in \texttt{Supp}(a)$, if there exists $\psi \in \Sigma$ s.t. $\psi \equiv \neg\varphi$ then there exists $b \in \texttt{Arg}(\Sigma)$ s.t. $(b, a) \in \mathcal{R}$. Then, $\mathcal{R}$ does not satisfy* ($\texttt{MC} \leftrightarrow \texttt{Ext}_c$).

What about the semantics which always return a unique extension, like grounded and ideal semantics? In such a case, it is not reasonable to expect that there is a bijection between $\texttt{MC}(\Sigma)$ and the set of extensions, since there can be several maximal consistent subsets of Σ. Let us formally state this fact.

Proposition 9 *If x is a semantics s.t. for any argumentation system $\mathcal{F}$ we have $|\texttt{Ext}_x(\mathcal{F})| = 1$ then there is no attack relation $\mathcal{R}$ which satisfies* ($\texttt{MC} \leftrightarrow \texttt{Ext}_x$).

The previous simple result is not surprising. The idea between those semantics is to have one extension that contains all the arguments that should be accepted according to any point of view. Thus, we can expect a link between the set of formulae not belonging to any minimal inconsistent set and those extensions. Note that the sufficient conditions for $\mathcal{R}$ were identified [12] so that for any finite set Σ and $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$ we have that the grounded and the ideal semantics coincide and that the extension is exactly $\texttt{Arg}(\Sigma \setminus (\Phi_1 \cup \ldots \cup \Phi_k))$ where $\{\Phi_1, \ldots, \Phi_k\}$ is the set of all minimal (for set inclusion) inconsistent subsets of Σ.

5 IDENTIFYING CLASSES OF ATTACK RELATIONS (NOT-)SATISFYING (MC↔EXT)

Until now, we have identified several properties that any attack relation satisfying ($\texttt{MC} \leftrightarrow \texttt{Ext}$) must satisfy. We also provided several results closely related to the choice of a specific acceptability semantics. In this section, we will identify classes of attack relations which satisfy, does not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$), or serve as lower (upper) bounds (w.r.t. set inclusion) for (non-)satisfying ($\texttt{MC} \leftrightarrow \texttt{Ext}$).

A particular class of attack relations, namely symmetric relations, have already been criticised for violating some important properties [1]. From that perspective, the next result is not surprising, as we show that any symmetric attack relation violates ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for all the semantics from Definition 4.

Proposition 10 *If $\mathcal{R}$ is a symmetric attack relation, then for every $x \in \{s, ss, p, c, g, i\}$, $\mathcal{R}$ falsifies* ($\texttt{MC} \leftrightarrow \texttt{Ext}_x$).

We will now identify another class of attack relations that do not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$). Namely, we will show that any (possible) attack which generate "too many attacks" falsifies ($\texttt{MC} \leftrightarrow \texttt{Ext}$). First, we need to formally define what we mean by "too many attacks". We do this by introducing the notion of conflict-completeness.

Definition 8 (Conflict-complete) *Let $\mathcal{R} \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$ be an attack relation. We say that $\mathcal{R}$ is conflict-complete iff for every minimal conflict $C \subseteq \mathcal{L}$ (i.e. for every inconsistent set whose every proper subset is consistent), for every $C', C'' \subseteq C$ s.t. $C' \neq \emptyset$, $C'' \neq \emptyset$, $C' \cup C'' = C$, for every argument a s.t. $\texttt{Supp}(a) = C'$, there exists an argument a'' s.t. $\texttt{Supp}(a'') = C''$ and $(a'', a') \in \mathcal{R}$.*

Now, we can show that if an attack relation is conflict-complete, then it falsifies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for stable, semi-stable, preferred and complete semantics.

Proposition 11 *Let $\mathcal{R}$ be an attack relation. If $\mathcal{R}$ is conflict-complete then $\mathcal{R}$ does not satisfy* ($\texttt{MC} \leftrightarrow \texttt{Ext}_x$) *for $x \in \{s, ss, p, c\}$.*

We will now define all (to the best of our knowledge) attack relations used in logic-based argumentation. Then, we prove that some of them are lower (upper) bounds (w.r.t. set inclusion) for non-satisfying ($\texttt{MC} \leftrightarrow \texttt{Ext}$). If $\Phi = \{\varphi_1, \ldots, \varphi_k\}$ is a set of formulae, notation $\bigwedge \Phi$ stands for $\varphi_1 \wedge \ldots \wedge \varphi_k$.

Definition 9 (Attack relations) *Let $a, b \in \texttt{Arg}(\mathcal{L})$. We define the following attack relations:*

- defeat*: $a\mathcal{R}_d b$ iff $\texttt{Conc}(a) \vdash \bigwedge \neg\texttt{Supp}(b)$*
- direct defeat*: $a\mathcal{R}_{dd} b$ iff $\exists \varphi \in \texttt{Supp}(b)$ s.t. $\texttt{Conc}(a) \vdash \neg\varphi$*
- undercut*: $a\mathcal{R}_u b$ iff $\exists \Phi \subseteq \texttt{Supp}(b)$ s.t. $\texttt{Conc}(a) \equiv \neg \bigwedge \Phi$*
- direct undercut*: $a\mathcal{R}_{du} b$ iff $\exists \varphi \in \texttt{Supp}(b)$ s.t. $\texttt{Conc}(a) \equiv \neg\varphi$*
- canonical undercut*: $a\mathcal{R}_{cu} b$ iff $\texttt{Conc}(a) \equiv \neg \bigwedge \texttt{Supp}(b)$*
- rebut*: $a\mathcal{R}_r b$ iff $\texttt{Conc}(a) \equiv \neg\texttt{Conc}(b)$*
- defeating rebut*: $a\mathcal{R}_{dr} b$ iff $\texttt{Conc}(a) \vdash \neg\texttt{Conc}(b)$*
- rebut + direct undercut*: $a\mathcal{R}_{rdu} b$ iff $a\mathcal{R}_r b$ or $a\mathcal{R}_{du} b$*
- conflicting attack*: $a\mathcal{R}_c b$ iff $\texttt{Supp}(a) \cup \texttt{Supp}(b) \vdash \perp$*

The reader can easily check that $\mathcal{R}_{cu}$ is conflict-complete, which leads to the conclusion that any attack relation which contains canonical undercut (in the set-theoretic sense) is also conflict-complete.

Proposition 12 *Let $\mathcal{R} \subseteq \texttt{Arg}(\mathcal{L}) \times \texttt{Arg}(\mathcal{L})$ be an attack relation. If $\mathcal{R}_{cu} \subseteq \mathcal{R}$ then $\mathcal{R}$ is conflict-complete.*

Thus, from Proposition 11, we conclude that every attack relation which contain "canonical undercut" falsifies ($\texttt{MC} \leftrightarrow \texttt{Ext}$).

Corollary 6 *Let $\mathcal{R}$ be an attack relation. If $\mathcal{R}_{cu} \subseteq \mathcal{R}$, (i.e. for every $a, b \in \texttt{Arg}(\mathcal{L})$, we have that $(a, b) \in \mathcal{R}_{cu}$ implies $(a, b) \in \mathcal{R}$), then $\mathcal{R}$ does not satisfy* ($\texttt{MC} \leftrightarrow \texttt{Ext}_x$) *for $x \in \{s, ss, p, c\}$.*

As a consequence, $\mathcal{R}_u, \mathcal{R}_d$ and $\mathcal{R}_c$ also falsify ($\texttt{MC} \leftrightarrow \texttt{Ext}$). Furthermore, we identified another class of attack relations, this time based on rebutting, which do not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$). Namely, any attack relation contained in defeating rebut falsifies ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$).

Proposition 13 *Let $\mathcal{R}$ be an attack relation. If $\mathcal{R} \subseteq \mathcal{R}_{dr}$ then $\mathcal{R}$ does not satisfy* ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$).

One of the consequences of the previous result is that $\mathcal{R}_r$ falsify ($\texttt{MC} \leftrightarrow \texttt{Ext}_s$).

6 PARTICULAR ATTACK RELATIONS AND (MC↔EXT)

In the previous section, we identified classes of relations which do not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$). In this section, we examine all the attack relations from Definition 9. By using the results we presented until now, we will prove that *direct undercut* and *direct defeat* satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for stable, semi-stable and preferred semantics, and falsify it for other semantics, while other attack relations fail to satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for any semantics. Note that it has been proved [10] that direct undercut satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) in the case of stable semantics. From Proposition 4, we conclude that $\mathcal{R}_{du}$ satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for semi-stable semantics. So, we only need to prove that $\mathcal{R}_{du}$ satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) in the case of preferred semantics.

Proposition 14 *Attack relation $\mathcal{R}_{du}$ satisfies* ($\texttt{MC} \leftrightarrow \texttt{Ext}_x$) *for $x \in \{s, ss, p\}$.*

Let us now show that $\mathcal{R}_{dd}$ also satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for stable, semi-stable and preferred semantics.

Proposition 15 *Attack relation $\mathcal{R}_{dd}$ satisfies* ($\texttt{MC} \leftrightarrow \texttt{Ext}_x$) *for $x \in \{s, ss, p\}$.*

We have already seen that no relation satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for the grounded or ideal semantics. By using Proposition 8, it is easy to conclude that neither $\mathcal{R}_{du}$ nor $\mathcal{R}_{dd}$ satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}_c$). Let us now prove that the remaining attack relations from Definition 9 do not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for neither of the semantics from Definition 4.

Proposition 16 *Attack relations $\mathcal{R}_d$, $\mathcal{R}_u$, $\mathcal{R}_{cu}$, $\mathcal{R}_r$, $\mathcal{R}_{dr}$, $\mathcal{R}_{rdu}$, $\mathcal{R}_c$ falsify* ($\texttt{MC} \leftrightarrow \texttt{Ext}$) *for stable, semi-stable, preferred, complete, grounded and ideal semantics.*

7 DISCUSSION AND RELATED WORK

In this paper, we studied the following question: for which attack relations and semantics, the function Arg is a bijection between the set of maximal consistent subsets of a knowledge base and extensions of the argumentation system obtained by constructing all the arguments from that knowledge base? Our main motivation is to identify the similarities and differences in the results obtained by using an argumentation-based approach and a non argumentation-based approach. Practical benefits of our work are: (i) "validate" argumentation-based approaches by showing in which cases they return a result comparable with that of some of well-known non-argumentation based approaches (ii) reduce computational complexity by using the simpler approach in the cases when the result obtained by an argumentation-based and a non argumentation-based approach is the same (iii) if differences are found, try to understand why they arise (what is the supplementary information induced by an attack relation) in order to use them (if we find that an argumentation-based approach models some situations better than a non argumentation-based approach) or to avoid them (if they represent problems, e.g. returning extensions having inconsistent bases). Our long-term goal is to better understand the role of argumentation and, more particularly, the expressivity of logic-based instantiations of Dung-style argumentation frameworks.

Let us now review the case when an attack relation satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for a given semantics. A positive aspect of this situation is that the result obtained from the argumentation system is comparable to that obtained by using maximal consistent sets of the knowledge base for reasoning. A possible criticism of such a relation is that it is useless, since we can obtain the same result without using argumentation at all. But, this is far from being true; namely, argumentation can be used for explanatory purposes. For example, if one wants to know why a certain conclusion is accepted, an argument having that conclusion can be presented. That argument can be attacked by other arguments and so on. It might be possible to *construct only a part of the argumentation graph* related to the argument in question, thus having a better knowledge representation (i.e. ignoring the parts of the knowledge base unrelated to the argument one wants to concentrate on).

Let us now consider the case when an attack relation does not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$). Since we supposed no preferences between the formulae of Σ, then the attack relation is introducing some hidden bias which results in some formulae being privileged. Whether we want or not this type of bias can be discussed, but we think that when doing this, we must at least be aware of that fact. An important question in such a case is: what type of formulae becomes privileged / unprivileged? To try to answer that question, let us take another look at the attack relations not satisfying ($\texttt{MC} \leftrightarrow \texttt{Ext}$) that we studied. Note that all the attack relations from the literature which do not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$) may return extensions having inconsistent bases. So, they cannot be used (at least if the goal of the argumentation process is to deal with inconsistency and somehow resolve it). Another situation would be the one where there are no extensions having inconsistent bases, but $\mathcal{R}$ does not satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$). This seems less in contradiction with the idea of using argumentation for inconsistency handling, since in that case, some maximal consistent sets do not yield extensions, but at least no extension has an inconsistent base. We did not encounter this situation for any of the attack relations used in the literature, and identifying and studying them will be a part of future work.

We now review the related work. The paper by Cayrol [10] is one of the early works relating the results obtained directly from a knowledge base and by using an argumentative approach. In that paper, it was shown that "direct undercut" satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$) for stable semantics, but no results for other semantics or attack relations were provided. We not only studied other attack relations and other semantics, but also provided a general study of properties an attack relation satisfying ($\texttt{MC} \leftrightarrow \texttt{Ext}$) must also satisfy.

Amgoud and Vesic [5] generalised the result by Cayrol [10] for the case of prioritised knowledge base, by showing that $\texttt{Arg}$ is a bijection between preferred sub-theories [7] (which generalise maximal consistent sets in case of prioritised knowledge base) and stable extensions of an instantiation of the preference-based argumentation system they developed.

Amgoud and Besnard [1, 2] also studied the link between a knowledge base and the corresponding argumentation system. Those papers introduced some important notions like conflict-dependence and validity of an attack relation and proved many results related to consistency in the underlying logic. However, note that the criterion ($\texttt{MC} \leftrightarrow \texttt{Ext}$) was neither defined nor studied in those papers; they provided [2, Corollary 1] a link between $\texttt{MC}(\Sigma)$ and maximal conflict-free sets of $\mathcal{F} = (\texttt{Arg}(\Sigma), \mathcal{R})$. Furthermore, this result is proved under hypotheses which are impossible to satisfy: the attack relation should be both valid and conflict-dependent, which is impossible (as proved in Proposition 3).

A recent paper by Gorogiannis and Hunter [12] studied the properties of attack relations in the case when a Dung-style argumentation system is instantiated with classical propositional logic. Our work is related to those ideas, however, the focus of our paper is different: we proved the properties that must be satisfied by any attack relation satisfying ($\texttt{MC} \leftrightarrow \texttt{Ext}$) and we identified the attack relations that serve as lower and upper bounds of classes of relations non-satisfying ($\texttt{MC} \leftrightarrow \texttt{Ext}$).

Considering the difference between complete and incomplete systems, note that many works in the 1990s [13, 15, 17] yield conceptual and philosophical arguments supporting partial computation (i.e. incomplete systems). The goal of this paper is not to argue that complete systems are in any sense "better" than incomplete ones (or vice versa), but only to study the possibilities and limits of (propositional) logic-based instantiations of Dung-style argumentation systems.

One of the challenges left for future work is to find a set of conditions such that an attack relation satisfies those conditions if and only if it satisfies ($\texttt{MC} \leftrightarrow \texttt{Ext}$). Recall that direct undercut and direct defeat are the only attack relations from the literature that satisfy ($\texttt{MC} \leftrightarrow \texttt{Ext}$). We believe that this fact deserves more attention and in the future work we will try to identify the reasons behind this fact. Our intuition is that the notion of conflict-completeness will play an important role in explaining this phenomenon.

A long-term research agenda includes defining and studying such links for other argumentation formalisms (i.e. not only for Dung-style argumentation frameworks instantiated with classical logic). For example, considering the framework proposed by Modgil and Prakken [14], we would like to investigate the questions like: is there a link between the set of conclusions obtained by using a non-argumentative system based on the same strict and defeasible rules and the result obtained by the corresponding argumentation system? If yes, what are advantages / drawbacks of an argumentative approach and in which case to use it? If no, why this happens and what is the intrinsic added value of argumentation?

ACKNOWLEDGEMENTS

The author would like to thank Leendert van der Torre for his comments which helped improve this paper considerably.

REFERENCES

[1] L. Amgoud and Ph. Besnard, 'Bridging the gap between abstract argumentation systems and logic', in *International Conference on Scalable Uncertainty Management (SUM'09)*, pp. 12–27, (2009).

[2] L. Amgoud and Ph. Besnard, 'A formal analysis of logic-based argumentation systems', in *International Conference on Scalable Uncertainty Management (SUM'10)*, pp. 42–55, (2010).

[3] L. Amgoud, Ph. Besnard, and S. Vesic, 'Identifying the core of logic-based argumentation systems', in *23rd International Conference on Tools with Artificial Intelligence (ICTAI'11)*, pp. 633–636, (2011).

[4] 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).

[5] L. Amgoud and S. Vesic, 'Handling inconsistency with preference-based argumentation', in *Proceedings of SUM'10*, pp. 56–69, (2010).

[6] Ph. Besnard and A. Hunter, *Elements of Argumentation*, MIT Press, 2008.

[7] G. Brewka, 'Preferred subtheories: An extended logical framework for default reasoning', in *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'89)*, pp. 1043–1048, (1989).

[8] M. Caminada, 'Semi-stable semantics', in *Proceedings of the 1st International Conference on Computational Models of Argument (COMMA'06)*, pp. 121–130. IOS Press, (2006).

[9] M. Caminada and L. Amgoud, 'On the evaluation of argumentation formalisms', *Artificial Intelligence Journal*, **171 (5-6)**, 286–310, (2007).

[10] C. Cayrol, 'On the relation between argumentation and non-monotonic coherence-based entailment', in *Proceedings of IJCAI'95*, pp. 1443–1448, (1995).

[11] 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).

[12] N. Gorogiannis and A. Hunter, 'Instantiating abstract argumentation with classical logic arguments: Postulates and properties', *Artificial Intelligence Journal*, **175**, 1479–1497, (2011).

[13] R.P. Loui, 'Process and policy: Resource-bounded nondemonstrative reasoning', *Computational Intelligence*, **14**(1), 1–38, (1998).

[14] S. Modgil and H. Prakken, 'Revisiting preferences and argumentation', in *Proceedings of IJCAI'11*, pp. 1021–1026, (2011).

[15] J. Pollock, 'How to reason defeasibly', *Artificial Intelligence Journal*, **57**, 1–42, (1992).

[16] *Argumentation in Artificial Intelligence*, eds., I. Rahwan and G. Simari, Springer, 2009.

[17] G. Vreeswijk, 'Abstract argumentation systems', *Artificial Intelligence Journal*, **90**, 225–279, (1997).