Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

# Dealing with Generic Contrariness in Structured Argumentation

**Pietro Baroni, Massimiliano Giacomin**
DII, University of Brescia
Brescia, Italy
{pietro.baroni|massimiliano.giacomin}@unibs.it

**Beishui Liao**
Zhejiang University - Univ. of Luxembourg
Hangzhou, China - Luxembourg
baiseliao@zju.edu.cn

## Abstract

The adoption of a generic contrariness notion in *ASPIC*$^+$ substantially enhances its expressiveness with respect to other formalisms for structured argumentation. In particular, it opens the way to novel investigation directions, like the use of multivalued logics in the construction of arguments. This paper points out however that in the current version of *ASPIC*$^+$ a serious technical difficulty related with generic contrariness is present. With the aim of preserving the same level of generality, the paper provides a solution based on a novel notion of closure of the contrariness relation at the level of sets of formulas and an abstract representation of conflicts between sets of arguments. The proposed solution is shown to satisfy the same rationality postulates as *ASPIC*$^+$ and represents a starting point for further technical and conceptual developments in structured argumentation.

## 1 Introduction

Research on structured argumentation has witnessed an increasing interest in formalisms lying at an intermediate level of abstraction between specific argumentation systems and Dung's argumentation frameworks ($AF$s) [Dung, 1995]. Among these formalisms, a preminent role is played by *ASPIC*$^+$, presented in [Prakken, 2010] "as a general abstract model of argumentation with structured arguments" and developed in [Modgil and Prakken, 2013]. *ASPIC*$^+$ is an extension of *ASPIC* [Amgoud *et al.*, 2006], able to satisfy the same *rationality postulates* [Caminada and Amgoud, 2007] while offering a more general account of structured argumentation. Indeed, in [Modgil and Prakken, 2013] *ASPIC*$^+$ is shown to capture a variety of existing argumentation systems. A key aspect in this respect is the adoption of a generic, possibly asymmetric, relation of contrariness between formulas while *ASPIC* assumes standard negation, i.e. a binary and symmetric contrariness relation. While in [Prakken, 2010; Modgil and Prakken, 2013] it is mainly emphasized that generic contrariness is useful to encompass a larger variety of more specific literature proposals, it can be observed that this notion also opens the way to novel investigation directions, like the use of multivalued logics in the construction of arguments (see [Baroni *et al.*, 2015] for an initial analysis on this issue). Consider for instance a logic encompassing three truth values, namely **T**, **F**, **U**, standing for *true*, *false*, and *unknown* (or, better, *unknowable* like in [Baroni *et al.*, 2015]). In this context, an argument concluding that the truth value of a proposition $P$ is **F** is not just in conflict with any argument concluding that the truth value of $P$ is **T**, but also with any argument concluding that the truth value of $P$ is **U** and so on.

As to the evaluation of argument acceptance, both *ASPIC* and *ASPIC*$^+$ rely on $AF$s. The rationality postulates prescribe that the outcomes of this evaluation respect some closure and consistency properties when mapped back at the instantiated level. Both in *ASPIC* and in *ASPIC*$^+$ postulate satisfaction is ensured on the basis of some well-formedness assumptions, including some closure properties of the set of strict rules. However, we show that the combination of a generic contrariness relation and rule closure may give rise to counterintuitive results. To solve this problem while preserving the generality of *ASPIC*$^+$, we propose a modified version of the formalism based on two main standpoints:

- a notion of closure of the contrariness relation, involving sets of formulas rather than individual ones;
- an abstract representation for the evaluation of argument acceptance using sets of arguments and their conflicts.

On the technical side, the proposed formalism is shown to satisfy the rationality postulates of [Caminada and Amgoud, 2007], as a first and essential soundness requirement. On the conceptual side, being based on a non-binary notion of argument conflict and being oriented to making the actual roots of conflicts explicit, the formalism turns out to provide an alternative approach to the representation and management of articulated conflicts (like the *tandem* example in [Baroni *et al.*, 2011]) also with standard negation. Further, since it shares several intuitions with the formalism of *Abstract Argumentation Systems* ($AAS$s) [Vreeswijk, 1997], our proposal provides original formal tools enabling novel investigations about the relationships between $AAS$s, $AF$s, and *ASPIC*$^+$.

The paper is organized as follows. Section 2 provides the necessary background and Section 3 points out a problematic aspect of *ASPIC*$^+$. Section 4 introduces the proposed modifications to the formalism and Section 5 deals with the satisfaction of the rationality postulates. Section 6 provides an example and, finally, Section 7 concludes the paper.

## 2 Background

The following definition gives the minimal elements of Dung's theory of $AF$s used in this paper. The reader is referred to [Dung, 1995; Baroni *et al.*, 2011] for more details. We recall the definition of complete extension only since many literature semantics (grounded, stable, preferred, semi-stable, ideal) are complete-based (their extensions are a subset of the complete extensions) and the results proved in this paper directly carry over to any complete-based semantics.

**Definition 1** *An AF is a pair $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$ where $\mathcal{A}$ is a set of* arguments *and $\rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ describes their* attack relation*, so that $(\alpha, \beta) \in \rightarrow$ (denoted $\alpha \rightarrow \beta$) indicates that $\alpha$* attacks *$\beta$. For a set $S \subseteq \mathcal{A}$, the attackers of $S$ are defined as $S^- = \{\alpha \in \mathcal{A} \mid \exists \beta \in S : \alpha \rightarrow \beta\}$ and the attackees of $S$ are defined as $S^+ = \{\alpha \in \mathcal{A} \mid \exists \beta \in S : \beta \rightarrow \alpha\}$. A set $S \subseteq \mathcal{A}$ is* conflict-free *iff $S \cap S^+ = \emptyset$. $S$* defends *an argument $\alpha$ iff $\{\alpha\}^- \subseteq S^+$. The set of arguments defended by $S$ in $\mathcal{F}$ is denoted as $\mathcal{D}_\mathcal{F}(S)$. A set $S$ is a* complete *extension of $\mathcal{F}$, denoted $S \in \mathcal{E}_{\mathcal{CO}}(\mathcal{F})$, iff $S$ is conflict-free and $S = \mathcal{D}_\mathcal{F}(S)$.*

We recall in the sequel the essentials of the definition of the *ASPIC*$^+$ argumentation system, omitting some details not required by the present paper.

**Definition 2** *An* argumentation system *is a tuple $AS = (\mathcal{L}, ^-, \mathcal{R}, n)$ where:*

1. *$\mathcal{L}$ is a logical language*
2. *$^-$ is a* contrariness *function from $\mathcal{L}$ to $2^\mathcal{L}$ such that: (i) $\varphi$ is a* contrary *of $\psi$ if $\varphi \in \overline{\psi}$, $\psi \notin \overline{\varphi}$; (ii) $\varphi$ is a* contradictory *of $\psi$ (denoted by $\varphi = -\psi$) if $\varphi \in \overline{\psi}$, $\psi \in \overline{\varphi}$; (iii) each $\varphi \in \mathcal{L}$ has at least one contradictory*
3. *$\mathcal{R} = \mathcal{R}_S \cup \mathcal{R}_D$ is a set of strict ($\mathcal{R}_S$) and defeasible ($\mathcal{R}_D$) inference rules of the form $\varphi_1, \ldots, \varphi_n \rightarrow \varphi$ and $\varphi_1, \ldots, \varphi_n \Rightarrow \varphi$ respectively (where $\varphi_i, \varphi$ are meta-variables ranging over wff in $\mathcal{L}$), and $\mathcal{R}_S \cap \mathcal{R}_D = \emptyset$*
4. *$n : \mathcal{R}_D \rightarrow \mathcal{L}$ is a naming convention for $\mathcal{R}_D$.*

A knowledge base is a subset of $\mathcal{L}$ including certain (called *axioms*) and defeasible (called *ordinary*) premises. It gives rise to the notion of argumentation theory.

**Definition 3** *A* knowledge base *in an argumentation system $AS = (\mathcal{L}, ^-, \mathcal{R}, n)$ is a set $\mathcal{K} \subseteq \mathcal{L}$ consisting of two disjoint subsets $\mathcal{K}_n$ (the* axioms*) and $\mathcal{K}_p$ (the* ordinary premises*). The tuple $AT = (AS, \mathcal{K})$ is called an* argumentation theory.

Two notions of *consistency* are considered in *ASPIC*$^+$.

**Definition 4** *For any $S \subseteq \mathcal{L}$, let the* closure of $S$ under strict rules*, denoted $Cl_{\mathcal{R}_S}(S)$, be the smallest set containing $S$ and the consequent of any strict rule in $\mathcal{R}_S$, whose antecedents are in $Cl_{\mathcal{R}_S}(S)$. Then a set $S \subseteq \mathcal{L}$ is: (i)* directly consistent *iff $\nexists \varphi, \psi \in S$ such that $\varphi \in \overline{\psi}$; (ii)* indirectly consistent *iff $Cl_{\mathcal{R}_S}(S)$ is directly consistent.*

Arguments are built from a knowledge base using rules.

**Definition 5** *An argument $a$ on the basis of a knowledge base $\mathcal{K}$ in an argumentation system $(\mathcal{L}, ^-, \mathcal{R}, n)$ is:*

1. *$\varphi$ if $\varphi \in \mathcal{K}$ with: $Prem(a) = \{\varphi\}$; $Conc(a) = \varphi$; $Sub(a) = \{\varphi\}$; $Rules(a) = \emptyset$; $Top(a) =$ undefined.*
2. *$a_1, \ldots a_n \rightarrow (\Rightarrow) \psi$ if $a_1, \ldots a_n$ are arguments such that there exists a strict (defeasible) rule $Conc(a_1), \ldots Conc(a_n) \rightarrow (\Rightarrow) \psi$ in $\mathcal{R}_S$ ($\mathcal{R}_D$) with: $Prem(a) = Prem(a_1) \cup \ldots \cup Prem(a_n)$; $Conc(a) = \psi$; $Sub(a) = Sub(a_1) \cup \ldots \cup Sub(a_n) \cup \{a\}$; $Rules(a) = Rules(a_1) \cup \ldots \cup Rules(a_n) \cup \{Conc(a_1), \ldots, Conc(a_n) \rightarrow (\Rightarrow) \psi\}$; $Top(a) = Conc(a_1), \ldots, Conc(a_n) \rightarrow (\Rightarrow) \psi$; $DefRules(a) = \{r \mid r \in Rules(a) \cap \mathcal{R}_D\}$; $StRules(a) = \{r \mid r \in Rules(a) \cap \mathcal{R}_S\}$.*

*For any argument $a$, $Prem_n(a) = Prem(a) \cap \mathcal{K}_n$; $Prem_p(a) = Prem(a) \cap \mathcal{K}_p$. $a$ is:* strict *if $DefRules(a) = \emptyset$,* defeasible *if $DefRules(a) \neq \emptyset$;* firm *if $Prem(a) \subseteq \mathcal{K}_n$;* plausible *if $Prem(a) \nsubseteq \mathcal{K}_n$;* fallible *if $a$ is plausible or defeasible;* finite *if $Rules(a)$ is finite.*

We assume, as in [Modgil and Prakken, 2013], that the set $Prem(a)$ of premises of an argument $a$ is always finite. An argument may include both fallible (ordinary premises and defeasible rules) and infallible (axioms and strict rules) elements. The following definition is based on this distinction.

**Definition 6** *For any set of arguments $\{a_1, \ldots, a_n\}$ the argument $a$ is a strict continuation of $\{a_1, \ldots, a_n\}$ iff $Prem_p(a) = \bigcup_{i=1}^n Prem_p(a_i)$; $DefRules(a) = \bigcup_{i=1}^n DefRules(a_i)$; $StRules(a) \supseteq \bigcup_{i=1}^n StRules(a_i)$; $Prem_n(a) \supseteq \bigcup_{i=1}^n Prem_n(a_i)$.*

Some further notations are worth introducing.

**Notation 1**
1. *Given $S \subseteq \mathcal{L}$, $S \vdash \varphi$ denotes that there exists a strict argument $a$ such that $Conc(a) = \varphi$, with $Prem(a) \subseteq S$.*
2. *$S \vdash_{min} \varphi$ denotes that $S \vdash \varphi$ and $\nexists T \subsetneq S : T \vdash \varphi$.*
3. *Given a set of arguments $\Sigma$, $Prem(\Sigma) \triangleq \bigcup_{a \in \Sigma} Prem(a)$, and similarly for $Conc(\Sigma)$, $Sub(\Sigma)$, $Rules(\Sigma)$, $Top(\Sigma)$, $DefRules(\Sigma)$, $StRules(\Sigma)$.*

The notion of *c-consistency* is based on contradiction.

**Definition 7** *A set $S \subseteq \mathcal{L}$ is* c-consistent *if for no $\varphi$ it holds that $S \vdash \varphi, -\varphi$. Otherwise $S$ is* c-inconsistent*. $S$ is minimally c-inconsistent iff $S$ is c-inconsistent and $\forall S' \subsetneq S$, $S'$ is c-consistent. An argument $a$ on the basis of a knowledge-base $\mathcal{K}$ in an argumentation system $(\mathcal{L}, ^-, \mathcal{R}, n)$ is c-consistent iff $Prem(a)$ is c-consistent.*

Accordingly, several properties for an argumentation theory can be formulated.

**Definition 8** *Let $AT = (AS, \mathcal{K})$ be an argumentation theory, where $AS = (\mathcal{L}, ^-, \mathcal{R}, n)$. $AT$ is:*

1. closed under contraposition *iff for all $S \subseteq \mathcal{L}$, $\psi \in S$, $\varphi \in \mathcal{L}$, if $S \vdash \varphi$ then $S \setminus \{\psi\} \cup \{-\varphi\} \vdash -\psi$;*
2. closed under transposition *iff if $\varphi_1, \ldots, \varphi_n \rightarrow \psi \in \mathcal{R}_S$ then, for $i = 1 \ldots n$, $\varphi_1, \ldots, \varphi_{i-1}, -\psi, \varphi_{i+1}, \ldots, \varphi_n \rightarrow -\varphi_i \in \mathcal{R}_S$;*
3. axiom consistent *iff $Cl_{\mathcal{R}_S}(\mathcal{K}_n)$ is directly consistent;*
4. c-classical *iff for any minimal c-inconsistent $S \subseteq \mathcal{L}$ and for any $\varphi \in S$, it holds that $S \setminus \{\varphi\} \vdash -\varphi$;*

5. well formed *if whenever* $\varphi$ *is a contrary of* $\psi$ *then* $\psi \notin \mathcal{K}_n$ *and* $\psi$ *is not the consequent of a strict rule.*

Three kinds of attack among arguments are considered.

**Definition 9** *An argument* $a$ attacks *an argument* $b$ *iff* $a$ undercuts*,* rebuts*, or* undermines $b$ *where:*

- $a$ undercuts $b$ *(on* $b'$*) iff* $Conc(a) \in \overline{n(r)}$ *for some* $b' \in Sub(b)$ *such that* $r = Top(b')$ *is defeasible.*
- $a$ rebuts $b$ *(on* $b'$*) iff* $Conc(a) \in \overline{\varphi}$ *for some* $b' \in Sub(b)$ *of the form* $b''_1, \ldots, b''_n \Rightarrow \varphi$*. In such a case* $a$ contrary-rebuts $b$ *iff* $Conc(a)$ *is a contrary of* $\varphi$*.*
- $a$ undermines $b$ *(on* $b'$*) iff* $Conc(a) \in \overline{\varphi}$ *for some* $b' = \varphi$*,* $\varphi \in Prem_p(b)$*. In such a case* $a$ contrary-undermines $b$ *iff* $Conc(a)$ *is a contrary of* $\varphi$*.*

Attack effectiveness, in some cases, depends on an ordering $\preceq$ among arguments, where $a \preceq b$ means that argument $b$ is at least as preferred as $a$ and, following [Prakken, 2010], we assume that $\preceq$ is a preorder. As usual $a \prec b$ iff $a \preceq b$ and $b \not\preceq a$; $a \approx b$ iff $a \preceq b$ and $b \preceq a$.

**Definition 10** *An argument ordering* $\preceq$ *is reasonable iff:*

1. $\forall a, b$ *if* $a$ *is strict and firm and* $b$ *is plausible or defeasible, then* $b \prec a$*;*
2. $\forall a, b$ *if* $a$ *is strict and firm then* $a \not\prec b$*;*
3. $\forall a, a', b$ *with* $a'$ *a strict continuation of* $\{a\}$*, if* $a \not\prec b$ *then* $a' \not\prec b$ *and if* $b \not\prec a$ *then* $b \not\prec a'$
4. *let* $\Theta = \{a_1, \ldots, a_n\}$ *be a finite set of arguments and for* $i = 1 \ldots n$ *let* $a^{+\backslash i}$ *be some strict continuation of* $\Theta \setminus \{a_i\}$*. Then, it is not the case that* $\forall i\ a^{+\backslash i} \prec a_i$*.*

Effective attacks give rise to defeat.

**Definition 11** *Let* $a$ *attack* $b$ *on* $b'$*. If* $a$ *undercuts, contrary-rebuts, or contrary-undermines* $b$ *on* $b'$*, then* $a$ preference-independent *attacks* $b$ *on* $b'$*, otherwise* $a$ preference-dependent *attacks* $b$ *on* $b'$*. Then,* $a$ defeats $b$ *iff for some* $b'$ *either* $a$ preference-independent *attacks* $b$ *on* $b'$ *or* $a$ preference-dependent *attacks* $b$ *on* $b'$ *and* $a \not\prec b'$*.* $a$ strictly defeats $b$ *iff* $a$ *defeats* $b$ *and* $b$ *does not defeat* $a$*.*

Two kinds of *structured argumentation frameworks* are defined from an argumentation theory, using the attack relation.

**Definition 12** *Let* $AT = (AS, \mathcal{K})$ *be an argumentation theory. A* (c-)structured argumentation framework *((c)*$-SAF$*) defined by* $AT$ *is a triple* $(\Sigma, \mathcal{C}, \preceq)$ *where* $\Sigma$ *is the set of all (c-consistent) finite arguments constructed from* $\mathcal{K}$ *in* $AS$ *(henceforth called the set of arguments on the basis of* $AT$*),* $\preceq$ *is an ordering on* $\Sigma$ *and* $(a, b) \in \mathcal{C}$ *iff* $a$ *attacks* $b$*.*

Well-definedness of (c-)structured argumentation frameworks depends on the properties of the underlying theory.

**Definition 13** *A* $SAF$ *defined by an* $AT$ *is* well-defined *iff* $AT$ *is c-classical, axiom consistent, well formed and closed under contraposition or closed under transposition; A* $c{-}SAF$ *defined by an* $AT$ *is* well-defined *iff* $AT$ *is axiom consistent, well formed and closed under contraposition or closed under transposition.*

A traditional Dung's framework is derived from a $(c){-}SAF$ using the defeat relation.

**Definition 14** *Let* $\Delta = (\Sigma, \mathcal{C}, \preceq)$ *be a* $(c){-}SAF$*, and* $\mathcal{D} \subseteq \Sigma \times \Sigma$ *be the defeat relation according to Def. 11. The* $AF$ *corresponding to* $\Delta$ *is defined as* $\mathcal{F}_\Delta = (\Sigma, \mathcal{D})$*.*

Compliance with rationality postulates is then proved in [Modgil and Prakken, 2013] under the assumptions above.

**Theorem 1** *Let* $\Delta = (\Sigma, \mathcal{C}, \preceq)$ *be a well-defined* $(c){-}SAF$ *with reasonable* $\preceq$ *and* $E$ *a* complete *extension of* $\mathcal{F}_\Delta$*. Then*

- $\forall a \in E$ *if* $a' \in Sub(a)$ *then* $a' \in E$*;*
- $\{Conc(a) \mid a \in E\} = Cl_{\mathcal{R}_S}(\{Conc(a) \mid a \in E\})$*;*
- $\{Conc(a) \mid a \in E\}$ *is consistent;*
- $Cl_{\mathcal{R}_S}(\{Conc(a) \mid a \in E\})$ *is consistent.*

## 3 A problem in *ASPIC*$^+$ without negation

Def. 2 is liberal about the notion of contrariness and contradiction. In particular, it requires every $\varphi \in \mathcal{L}$ to have a contradictory (to ensure that the items of Def. 8 using $-$ are well-founded) but leaves open the possibility of having more than one contradictory. However this may give rise to difficulties with the closure properties stated in Def. 8. To see this, consider a simple language $\mathcal{L}^6 = \{\mathtt{B}, \mathtt{M}, \mathtt{D}, \mathtt{T}, \mathtt{A}, \mathtt{O}\}$ meant to intuitively correspond to the properties of being *Bachelor*, *Married*, *Divorced*, *Teen-ager*, *Adult*, and *Old* respectively, where the former three properties are mutually exclusive and so are the three latter ones, hence $\overline{\mathtt{B}} = \{\mathtt{M}, \mathtt{D}\}$, $\overline{\mathtt{M}} = \{\mathtt{B}, \mathtt{D}\}$, $\overline{\mathtt{D}} = \{\mathtt{B}, \mathtt{M}\}$, and $\overline{\mathtt{T}} = \{\mathtt{A}, \mathtt{O}\}$, $\overline{\mathtt{A}} = \{\mathtt{T}, \mathtt{O}\}$, $\overline{\mathtt{O}} = \{\mathtt{T}, \mathtt{A}\}$. Assume then that as a strict rule (e.g. imposed by law) a teenager is a bachelor, i.e. $\mathtt{T} \rightarrow \mathtt{B} \in \mathcal{R}_S$.

Suppose now that closure under transposition is applied to this rule. We get $-\mathtt{B} \rightarrow -\mathtt{T}$, however both $-\mathtt{B}$ and $-\mathtt{T}$ admit two replacements, namely $\mathtt{M}$ or $\mathtt{D}$ and $\mathtt{A}$ or $\mathtt{O}$ respectively. Hence a plain application of closure under transposition generates four strict rules: $\mathtt{M} \rightarrow \mathtt{A}$, $\mathtt{M} \rightarrow \mathtt{O}$, $\mathtt{D} \rightarrow \mathtt{A}$, $\mathtt{D} \rightarrow \mathtt{O}$. It follows that both $\{\mathtt{M}\}$ and $\{\mathtt{D}\}$ are c-inconsistent. Then a c-classical $AT$ (required for a well-defined $SAF$) would impose $\emptyset \vdash -\mathtt{M}$ and $\emptyset \vdash -\mathtt{D}$, leading to $\emptyset \vdash \mathtt{B}, \mathtt{D}, \mathtt{M}$, i.e. contradictory conclusions should be strictly derived from the empty set. Given that $AT$ must be axiom consistent, it follows that $\mathcal{L}^6$ with $\mathtt{T} \rightarrow \mathtt{B} \in \mathcal{R}_S$ is forbidden in a well-defined $SAF$.

A well-defined $c{-}SAF$ does not require a c-classical $AT$, but arguments must be c-consistent, which, in this case, implies in particular that neither $\mathtt{M}$ nor $\mathtt{D}$ can be a premise. To put it in other words, the fact that a teen-ager must be bachelor prevents you to have as a premise that a person is married or divorced. Note that these observations do not affect the correctness of the results in [Modgil and Prakken, 2013], rather they show that multiple contradictories may give rise to a collapse of the expressiveness of the system: consistency is preserved at the price of very counterintuitive limitations.

One could fix this problem by requiring in Def. 8 every element of the language to have exactly one contradictory. Indeed the main example presented in [Modgil and Prakken, 2013] is coherent with this restriction. It may be noted however that this amounts to actually reintroduce negation within the language, though in disguise. While using negation (and possibly also disjunction) would be appropriate to this toy example, proposing it as the general solution to these difficulties would compromise the aim of *ASPIC*$^+$ of being as language-independent as possible. Quoting [Prakken, 2010], "formulas like *bachelor* and *married* can, if desired, be declared contradictory without having to reason with an axiom $\neg$(bachelor $\wedge$ married)". For similar reasons, we do not consider other possible very specific technical fixes, like for instance forbidding that two contraries of a formula are in turn contrary each other. Thus, the goal of this paper is to devise an alternative solution to achieve the same desirable results obtained in [Modgil and Prakken, 2013] while keeping the underlying language as generic as possible, namely equipped only with a basic contrariness function.

## 4 Revising *ASPIC*$^+$

We assume the same basic elements of *ASPIC*$^+$ just removing any further specification and constraint on the contrariness function. Thus, in Def. 2 items 1, 3, and 4, are unmodified while in item 2 we drop (iii). Defs. 3, 4, 5, 6, as well as Notation 1 are also left unmodified.

We then leave apart c-consistency (Def. 7) and consider the properties of Def. 8, used to introduce well-definedness. To start, we keep only the properties of axiom consistency and well-formedness and focus on the closure requirements. Closure under transposition imposes a sort of explicit completeness of the set of strict rules, based on the $-$ relation, while closure under contraposition imposes this completeness implicitly through the notion of strict derivation $\vdash$.

The use of $-$ being problematic, (explicit or implicit) closure of strict rules cannot be defined. One can, however, pursue the complementary idea of establishing a closure of the contrariness relation on the basis of the strict rules, according to the fact that the contrariness relation captures some form of (directional) incompatibility. The relevant intuition is as follows. Let $\varphi \in \overline{\psi}$ and $\chi \rightarrow \varphi \in \mathcal{R}_S$. Since $\varphi$ is incompatible with $\psi$ and $\varphi$ strictly follows from $\chi$, one may require that $\chi$ too is regarded as (directionally) incompatible with $\psi$. Further, let $\omega \rightarrow \psi \in \mathcal{R}_S$. Since $\varphi$ is incompatible with $\psi$ and when $\omega$ holds also $\psi$ necessarily holds, one may require $\varphi$ to be incompatible with $\omega$ too. This needs two complements.

First, a strict rule may have multiple premises, e.g. in $\mathcal{R}_S$ we may have $\eta, \theta \rightarrow \varphi$. Then, it is the set $\{\eta, \theta\}$ to be incompatible with $\psi$. Similarly, if $\kappa, \lambda \rightarrow \psi$ we get that $\varphi$ is incompatible with $\{\kappa, \lambda\}$. These examples show that an extended contrariness relation involving sets of formulas is needed. Intuitively, this can be justified by the fact that negation is a collective notion (it includes whatever is incompatible with the negated entity): if it is removed, the underlying collections of elements have to be dealt with explicitly.

Second, strict derivation may involve several rules. Hence the previous considerations need to be generalized to the cases where we have $\vdash$ instead of $\rightarrow$ in the formulas above. More precisely, in order to ensure minimality in the identification of incompatibilities, $\vdash_{min}$ should be considered.

Accordingly, we introduce an extended contrariness relation between subsets of $\mathcal{L}$ obtained as a "completion" of the basic contrariness relation $^-$. A preliminary precisation is required. Taking into account Def. 5 and Notation 1, in an argumentation theory strict derivation must start from the knowledge base, i.e. $S \vdash_{min} \varphi$ implies $S \subseteq \mathcal{K}$. However, we aim at defining a notion of closure that does not depend on the set of premises but only on the contrariness relation and the set of strict rules. To this purpose we need to define a notion of strict derivability, rather than of strict derivation, namely the possibility to derive a conclusion from some elements of the language using strict rules only, independently of the presence of these elements in the knowledge base.

**Definition 15** *Given an argumentation system* $AS = (\mathcal{L}, ^-, \mathcal{R}, n)$ *the strict knowledge base* $\mathcal{K}^*_{AS}$ *for AS is given by* $\mathcal{K}_n = \mathcal{L}$, $\mathcal{K}_p = \emptyset$ *and the corresponding argumentation theory is defined as* $AT^*_{AS} = (AS, \mathcal{K}^*_{AS})$. $S \vdash^* \varphi$ *and* $S \vdash^*_{min} \varphi$ *denote respectively that* $S \vdash \varphi$ *and* $S \vdash_{min} \varphi$ *in* $AT^*_{AS}$.

Note that from the finiteness of the set of premises of any argument, it follows that whenever $S \vdash^*_{min} \varphi$, $S$ is finite.

**Definition 16** *Given an argumentation system* $AS = (\mathcal{L}, ^-, \mathcal{R}, n)$, *let* $\mathcal{EC}^1(AS)$, $\mathcal{EC}^2(AS)$,$\mathcal{EC}^3(AS)$ *be the following subsets of* $2^{\mathcal{L}} \times 2^{\mathcal{L}}$

- $\mathcal{EC}^1(AS) = \{(\{\varphi\}, \{\psi\}) \mid \varphi \in \overline{\psi}\}$;
- $\mathcal{EC}^2(AS) = \{(S, \{\psi\}) \mid S \vdash^*_{min} \varphi$ *and* $\varphi \in \overline{\psi}\}$;
- $\mathcal{EC}^3(AS) = \{(S, T) \mid T \vdash^*_{min} \psi$ *and* $(S, \{\psi\}) \in \mathcal{EC}^1(AS) \cup \mathcal{EC}^2(AS)\}$.

*The extended contrariness relation is defined as* $\mathcal{EC}(AS) = \mathcal{EC}^1(AS) \cup \mathcal{EC}^2(AS) \cup \mathcal{EC}^3(AS) \subseteq 2^{\mathcal{L}} \times 2^{\mathcal{L}}$.

A set is $\mathcal{EC}$-incompatible if it corresponds to the union of two contrary subsets according to $\mathcal{EC}(AS)$.

**Definition 17** *Given an* $AS = (\mathcal{L}, ^-, \mathcal{R}, n)$, *a set* $S \subseteq \mathcal{L}$ *is* $\mathcal{EC}$-incompatible, *denoted as* $S \in \mathcal{ECI}(AS)$, *if there is some* $T, U$ *such that* $S = T \cup U$ *and* $(T, U) \in \mathcal{EC}(AS)$.

We can now revise accordingly the various forms of attack.

**Definition 18** *Given an argumentation theory* $AT = (AS, \mathcal{K})$, *a set of arguments* $\Sigma$ attacks *an argument* $b$ *iff* $\Sigma$ undercuts, rebuts, *or* undermines $b$ *where:*

- $\Sigma$ undercuts $b$ *(on* $b'$*) iff* $Conc(\Sigma) \cup \{n(r)\} \in \mathcal{ECI}(AS)$ *for some* $b' \in Sub(b)$ *such that* $r = Top(b') \in \mathcal{R}_D$.
- $\Sigma$ rebuts $b$ *(on* $b'$*) iff for some* $b' \in Sub(b)$ *of the form* $b''_1, \ldots, b''_n \Rightarrow \varphi$ *the following condition holds:* $\exists T, U$ *such that* $T \cup U = Conc(\Sigma) \cup \{\varphi\}$, $(T, U) \in \mathcal{EC}(AS)$ *and* $\varphi \in U$. *In this case* $\Sigma$ contrary rebuts $b$ *if there is no* $V, W$ *such that* $V \cup W = Conc(\Sigma) \cup \{\varphi\}$, $(V, W) \in \mathcal{EC}(AS)$ *and* $\varphi \in V$.
- $\Sigma$ undermines $b$ *(on* $b'$*) iff for some* $b' = \varphi$, $\varphi \in Prem_p(b)$ *the following condition holds:* $\exists T, U$ *such that* $T \cup U = Conc(\Sigma) \cup \{\varphi\}$, $(T, U) \in \mathcal{EC}(AS)$ *and* $\varphi \in U$. *In this case* $\Sigma$ contrary undermines $b$ *if there is no* $V, W$ *such that* $V \cup W = Conc(\Sigma) \cup \{\varphi\}$, $(V, W) \in \mathcal{EC}(AS)$ *and* $\varphi \in V$.

As the effectiveness of some attacks depends on the preference relation, we need to generalize the notion of preference ordering to our extended context.

**Definition 19** *Given an argument ordering $\preceq$ on a set of arguments $\Sigma$, we extend $\preceq$ to $2^\Sigma \times \Sigma$ as follows. An argument $a$ is at least as preferred as a set of arguments $\Theta$, denoted as $\Theta \preceq a$ iff $\exists b \in \Theta$ such that $b \preceq a$. $a$ is strictly preferred to $\Theta$, denoted $\Theta \prec a$ iff $\exists b \in \Theta$ such that $b \prec a$, not strictly preferred to $\Theta$, denoted $\Theta \not\prec a$ iff $\nexists b \in \Theta$ such that $b \prec a$.*

Accordingly, we can define an extended notion of defeat.

**Definition 20** *Let the set of arguments $\Theta$ attack an argument $b$ on $b'$ according to Def. 18. If $\Theta$ undercuts, contrary-rebuts, or contrary-undermines $b$ on $b'$, then $\Theta$* preference-independent *attacks $b$ on $b'$, otherwise $\Theta$* preference-dependent *attacks $b$ on $b'$. Then, $\Theta$* defeats *$b$, denoted as $\Theta \rightsquigarrow b$, iff either $\Theta$* preference-independent *attacks $b$ on $b'$ or $\Theta$* preference-dependent *attacks $b$ on $b'$ and $\Theta \not\prec b'$.*

Lemmata 1 and 2 are used later (the proofs are omitted).

**Lemma 1** *If $\Theta$ is finite then $\exists a \in \Theta : (\Theta \setminus \{a\}) \not\prec a$.*

**Lemma 2** *Given an argumentation theory $AT = (AS, \mathcal{K})$, if $\Theta \rightsquigarrow a$ then $\Theta \rightsquigarrow b$ for every $b$ such that $a \in Sub(b)$.*

We can now define an $AF$ based on the above relation of defeat in order to evaluate the justification status of arguments. The idea is that the nodes of the framework represent relevant sets of arguments. In particular we need a node for each singleton corresponding to a produced argument, and a node for each set that defeats some produced argument.

**Definition 21** *Given an argumentation theory $AT = (AS, \mathcal{K})$ with ordering $\preceq$, let $\mathcal{ARGS}(AT)$ be the set of the arguments produced in $AS$ on the basis of $\mathcal{K}$. The set of relevant sets of arguments of $AT$, denoted as $\mathcal{RS}(AT)$, is defined as $\mathcal{RS}(AT) = \{\{a\} \mid a \in \mathcal{ARGS}(AT)\} \cup \{\Theta \mid \Theta \subseteq \mathcal{ARGS}(AT)$ and $\exists b \in \mathcal{ARGS}(AT) : \Theta \rightsquigarrow b\}$.*

In the $AF$ a relevant set of arguments attacks another one (denoted as D-attacks) simply if it defeats one of its members.

**Definition 22** *Let $\Sigma, \Theta \in \mathcal{RS}(AT)$ for an argumentation theory $AT$. $\Sigma$ D-attacks $\Theta$, denoted as $\Sigma \twoheadrightarrow \Theta$, iff $\exists a \in \Theta : \Sigma \rightsquigarrow a$.*

The relevant-set-based $AF$ is defined accordingly.

**Definition 23** *Given an argumentation theory $AT = (AS, \mathcal{K})$, the $\mathcal{RS}$-based argumentation framework induced by $AT$ is defined as $\mathcal{RS}{-}AF(AT) = (\mathcal{RS}(AT), \twoheadrightarrow)$.*

**Notation 2** *Each node of the induced $\mathcal{RS}$-based argumentation framework $\mathcal{RS}{-}AF(AT)$ corresponds to a set $\Sigma$ of arguments of $AT$. To make this more explicit, we write $\|\Sigma\|$ to denote the node of $\mathcal{RS}{-}AF(AT)$ corresponding to $\Sigma$. The set of (instantiated) arguments supported by an extension $E$ of $\mathcal{RS}{-}AF(AT)$ is defined as $\mathcal{SA}(E) = \bigcup_{\|\Sigma\| \in E} \Sigma$.*

## 5 Rationality properties

In this section we show that the revised version of *ASPIC*$^+$ is well-behaved and satisfies the rationality postulates (some proofs are omitted due to space limitations). The following lemmata show that the acceptance of sets of arguments and of singletons in $\mathcal{RS}{-}AF(AT)$ are in accordance.

**Lemma 3** *Given an argumentation theory $AT = (AS, \mathcal{K})$, let $\mathcal{C} \in \mathcal{E_{CO}}(\mathcal{RS}{-}AF(AT))$. If $\|\Sigma\| \in \mathcal{C}$ then $\forall a \in \Sigma$ $\|\{a\}\| \in \mathcal{C}$. If $a \in \mathcal{SA}(\mathcal{C})$ then $\|\{a\}\| \in \mathcal{C}$.*

**Proof:** Note that the second statement is a direct consequence of the first one, proved below. From Def. 21, if $\Sigma \in \mathcal{RS}(AT)$ then $\forall a \in \Sigma$, $\{a\} \in \mathcal{RS}(AT)$. Let $\|\Sigma\| \in \mathcal{C}$ and $a \in \Sigma$. $\mathcal{C} \cup \{\|\{a\}\|\}$ is conflict-free. In fact, from Def. 22, for any $\Theta \in \mathcal{RS}(AT)$, if $\Theta \twoheadrightarrow \{a\}$, then $\Theta \twoheadrightarrow \Sigma$, and $\|\Theta\| \in \mathcal{C}$ would contradict the conflict-freeness of $\mathcal{C}$. The case $\{a\} \twoheadrightarrow \Theta$ leads to $\Theta \twoheadrightarrow \{a\}$ since $\mathcal{C}$ defends itself and hence to the same contradiction. Analogously, we get $\|\{a\}\|^- \subseteq \|\Sigma\|^-$ and since $\|\Sigma\|^- \subseteq \mathcal{C}^+$, also $\|\{a\}\|^- \subseteq \mathcal{C}^+$. Hence $\mathcal{C}$ defends $\|\{a\}\|$ and, from the properties of complete extensions, it must be the case that $\|\{a\}\| \in \mathcal{C}$. □

**Lemma 4** *Let $AT = (AS, \mathcal{K})$, $\mathcal{C} \in \mathcal{E_{CO}}(\mathcal{RS}{-}AF(AT))$ and $\Sigma \in \mathcal{RS}(AT)$. If $\forall a \in \Sigma$ $\|\{a\}\| \in \mathcal{C}$ then $\|\Sigma\| \in \mathcal{C}$.*

The next lemma follows directly from the previous ones.

**Lemma 5** *Let $AT = (AS, \mathcal{K})$, $\mathcal{C} \in \mathcal{E_{CO}}(\mathcal{RS}{-}AF(AT))$ and $\Sigma, \Theta, (\Sigma \cup \Theta) \in \mathcal{RS}(AT)$. If $\|\{\Sigma\}\|, \|\{\Theta\}\| \in \mathcal{C}$ then $\|\Sigma \cup \Theta\| \in \mathcal{C}$.*

We now prove the properties established in Theorems 12-15 of [Modgil and Prakken, 2013] and corresponding to the rationality postulates of [Caminada and Amgoud, 2007].

**Theorem 2** *Let $AT = (AS, \mathcal{K})$, $\mathcal{C} \in \mathcal{E_{CO}}(\mathcal{RS}{-}AF(AT))$. If $a \in \mathcal{SA}(\mathcal{C})$ then $\forall b \in Sub(a)$ $b \in \mathcal{SA}(\mathcal{C})$.*

**Proof:** From Lemma 3 $\|\{a\}\| \in \mathcal{C}$. We show that $\forall b \in Sub(a)$ $\|\{b\}\| \in \mathcal{C}$, implying $b \in \mathcal{SA}(\mathcal{C})$. From Lemma 2 and Defs. 21-22, if $\Theta \twoheadrightarrow \{b\}$ then $\Theta \twoheadrightarrow \{a\}$. It follows that $\|\{b\}\|^- \subseteq \|\{a\}\|^-$. Since $\|\{a\}\| \in \mathcal{C}$ we get $\|\{b\}\|^- \subseteq \mathcal{C}^+$ and $\|\{b\}\|^- \cap \mathcal{C} = \emptyset$. Further, it cannot be that $\|\{b\}\|^+ \cap \mathcal{C} \neq \emptyset$ since, by the properties of complete extensions, this would contradict $\|\{b\}\|^- \cap \mathcal{C} = \emptyset$. Using again the properties of complete extensions we get $\|\{b\}\| \in \mathcal{C}$. □

**Theorem 3** *Let $AT = (AS, \mathcal{K})$, $\mathcal{C} \in \mathcal{E_{CO}}(\mathcal{RS}{-}AF(AT))$. Then $\{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\} = Cl_{\mathcal{R}_S}(\{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\})$.*

**Proof:** Suppose by contradiction that $\exists \varphi$ such that $\varphi \in Cl_{\mathcal{R}_S}(\{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\})$ and $\varphi \notin \{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\}$. Since $\varphi \in Cl_{\mathcal{R}_S}(\{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\})$ there is an argument $b$ with $Conc(b) = \varphi$ such that $b$ is a strict continuation of a set of arguments $\Sigma \subseteq \mathcal{SA}(\mathcal{C})$. Moreover from $\varphi \notin \{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\}$ we get $b \notin \mathcal{SA}(\mathcal{C})$ and $\|\{b\}\| \notin \mathcal{C}$. From the properties of complete extensions, it follows $\|\{b\}\|^- \neq \emptyset$. Given that $b$ is a strict continuation of a set of arguments $\Sigma$, by Def. 18 it can only receive attacks through some of its subarguments belonging to $Sub(\Sigma)$. Hence $\|\{b\}\|^- = \bigcup_{c \in Sub(\Sigma)} \|\{c\}\|^-$. From $\Sigma \subseteq \mathcal{SA}(\mathcal{C})$ and Theorem 2 we get $Sub(\Sigma) \subseteq \mathcal{SA}(\mathcal{C})$ and from Lemma 3 we get $\forall c \in Sub(\Sigma)$ $\|\{c\}\| \in \mathcal{C}$. From the properties of complete extensions and the equality $\|\{b\}\|^- = \bigcup_{c \in Sub(\Sigma)} \|\{c\}\|^-$ it follows $\|\{b\}\|^- \subseteq \mathcal{C}^+$ and $\|\{b\}\|^- \cap \mathcal{C} = \emptyset$. Further, it cannot be that $\|\{b\}\|^+ \cap \mathcal{C} \neq \emptyset$ since, by the properties of complete extensions, this would contradict $\|\{b\}\|^- \cap \mathcal{C} = \emptyset$. It follows that it must be the case that $\|\{b\}\| \in \mathcal{C}$. Contradiction. □

**Theorem 4** *Given $AT = (AS, \mathcal{K})$ with argument ordering $\preceq$ satisfying item 1 of Def. 10, let $\mathcal{C} \in \mathcal{E}_{\mathcal{CO}}(\mathcal{RS}{-}AF(AT))$. If $AT$ is axiom consistent and well-formed then $\{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\}$ is directly consistent.*

**Theorem 5** *Given $AT = (AS, \mathcal{K})$ with argument ordering $\preceq$ satisfying item 1 of Def. 10, let $\mathcal{C} \in \mathcal{E}_{\mathcal{CO}}(\mathcal{RS}{-}AF(AT))$. If $AT$ is axiom consistent and well-formed then $Cl_{\mathcal{R}_S}(\{Conc(a) \mid a \in \mathcal{SA}(\mathcal{C})\})$ is directly consistent.*

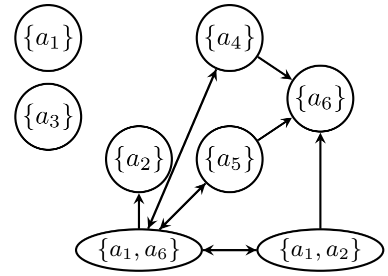


Figure 1: The "Married or 16?" example.

## 6 An example

Consider language $\mathcal{L}^{10} = \mathcal{L}^6 \cup \{\texttt{HC}, \texttt{GC}, \texttt{FC}, \texttt{WR}\}$, where intuitively HC means that the person we are talking about has a birth certificate dated 16 years ago, GC represents the default assumption that the certificate is genuine, FC means that the certificate is fake, and WR means that the person wears a wedding ring. The contrariness relation is as in Sect. 3 with the only addition of $\overline{\texttt{GC}} = \{\texttt{FC}\}$, i.e. FC is a contrary of GC. Let $\mathcal{R}_S = \{\texttt{T} \rightarrow \texttt{B}; \texttt{HC}, \texttt{GC} \rightarrow \texttt{T}\}$ and $\mathcal{R}_D = \{\texttt{WR} \Rightarrow \texttt{M}\}$. From $\mathcal{R}_S$ we obtain $\{\texttt{T}\} \vdash^*_{min} \texttt{B}$, $\{\texttt{HC}, \texttt{GC}\} \vdash^*_{min} \texttt{T}$, $\{\texttt{HC}, \texttt{GC}\} \vdash^*_{min} \texttt{B}$. Applying Def. 16 and omitting $\mathcal{EC}^1(AS)$, we get $\mathcal{EC}^2(AS) = \{(\{\texttt{T}\}, \{\texttt{M}\}), (\{\texttt{T}\}, \{\texttt{D}\}), (\{\texttt{HC}, \texttt{GC}\}, \{\texttt{A}\}), (\{\texttt{HC}, \texttt{GC}\}, \{\texttt{O}\}), (\{\texttt{HC}, \texttt{GC}\}, \{\texttt{M}\}), (\{\texttt{HC}, \texttt{GC}\}, \{\texttt{D}\})\}$ and then $\mathcal{EC}^3(AS) = \{(\{\texttt{M}\}, \{\texttt{T}\}), (\{\texttt{D}\}, \{\texttt{T}\}), (\{\texttt{A}\}, \{\texttt{HC}, \texttt{GC}\}), (\{\texttt{O}\}, \{\texttt{HC}, \texttt{GC}\}), (\{\texttt{M}\}, \{\texttt{HC}, \texttt{GC}\}), (\{\texttt{D}\}, \{\texttt{HC}, \texttt{GC}\})\}$.

Suppose now $\mathcal{K}_n = \{\texttt{HC}, \texttt{WR}\}$ (it is certain that the person has a birth certificate and wears a wedding ring) and $\mathcal{K}_p = \{\texttt{GC}\}$ (it is defeasibly assumed that the certificate is genuine). We get the following arguments: $a_1 = \texttt{HC}$, $a_2 = \texttt{GC}$, $a_3 = \texttt{WR}$, $a_4 = [a_1, a_2 \rightarrow \texttt{T}]$, $a_5 = [a_4 \rightarrow \texttt{B}]$, $a_6 = [a_3 \Rightarrow \texttt{M}]$.

From Def. 18 we have that both $\{a_4\}$ and $\{a_5\}$ rebut $a_6$, $\{a_1, a_2\}$ rebuts $a_6$, and $\{a_1, a_6\}$ undermines (and also rebuts) $a_2$ and its superarguments $a_4$ and $a_5$. All these attacks are preference dependent. To keep the presentation simple we assume that all fallible arguments (i.e. $a_2, a_4, a_5, a_6$) are equally preferred and are less preferred than the firm arguments $a_1$ and $a_3$. Under this assumption, all the above attacks are successful and give rise to a defeat according to Def. 20, i.e., summing up, $\{a_4\} \rightsquigarrow a_6$, $\{a_5\} \rightsquigarrow a_6$, $\{a_1, a_2\} \rightsquigarrow a_6$, $\{a_1, a_6\} \rightsquigarrow a_2$, $\{a_1, a_6\} \rightsquigarrow a_4$, $\{a_1, a_6\} \rightsquigarrow a_5$. Then, by Def. 21, $\mathcal{RS}(AT) = \{\{a_1\}, \{a_2\}, \{a_3\}, \{a_4\}, \{a_5\}, \{a_6\}, \{a_1, a_2\}, \{a_1, a_6\}\}$ and the D-attack relation $\twoheadrightarrow$ consists of the replication of the defeat relation described above (with the attacked arguments replaced by the corresponding singletons) with the addition of a mutual D-attack between $\{a_1, a_2\}$ and $\{a_1, a_6\}$ and of attacks from $\{a_4\}$ and $\{a_5\}$ to $\{a_1, a_6\}$.

Figure 1 shows the resulting $AF$. Its complete extensions are $\mathcal{C}_1 = \{\|\{a_1\}\|, \|\{a_3\}\|\}$, $\mathcal{C}_2 = \{\|\{a_1\}\|, \|\{a_3\}\|, \|\{a_6\}\|, \|\{a_1, a_6\}\|\}$, $\mathcal{C}_3 = \{\|\{a_1\}\|, \|\{a_2\}\|, \|\{a_3\}\|, \|\{a_4\}\|, \|\{a_5\}\|, \|\{a_1, a_2\}\|\}$. $\mathcal{C}_1$ is the grounded extension [Dung, 1995] supporting only the conclusions HC and WR. $\mathcal{C}_2$ and $\mathcal{C}_3$ correspond to the two alternative choices arising from the incompatibility between the fallible arguments $a_2$ (the certificate is genuine) and $a_6$ (the person is married) given that it is certain that s/he has a certicate asserting that s/he is a teenager and that there is no preference between the alternatives. Note that the structure of the $AF$ evidences this mutual attack as the root of all the inconsistencies. In fact, the acceptance of all other arguments (apart the unattacked ones) basically depends on the choice between $\|\{a_1, a_2\}\|$ and $\|\{a_1, a_6\}\|$. It may also be noted that the attacks from $\|\{a_4\}\|$ and $\|\{a_5\}\|$ against $\|\{a_6\}\|$ and $\|\{a_1, a_6\}\|$ could be suppressed without affecting the final outcome. This suggests that, at least in some cases, the same results can be achieved with a more compact representation. This issue is an important direction of future work.

As a variation of the example, assume that we add an ordinary premise FC since there is some evidence suggesting the falsity of the certificate. Then we get an additional argument $a_7 = \texttt{FC}$ which preference-independent attacks $a_2$ and its superarguments $a_4$ and $a_5$. Skipping some obvious steps we obtain a modified framework where, in particular, $\|\{a_7\}\| \twoheadrightarrow \|\{a_1, a_2\}\|$ and, as expected, there is exactly one complete extension $\mathcal{C}'_1 = \{\|\{a_1\}\|, \|\{a_3\}\|, \|\{a_6\}\|, \|\{a_7\}\|, \|\{a_1, a_6\}\|\}$ corresponding to the unique surviving (but still defeasible) alternative.

Due to space limitations, other application examples can not be presented in detail. We remark however that the proposed formalism provides a natural treatment of cases where conflicts between arguments arise from ternary (or, generally, more than binary) incompabilities. These cases have attracted attention in the literature since they are somehow challenging for modelling approaches (like $AF$s) based on binary incompatibility/attack relations. For instance the *tandem* example [Baroni *et al.*, 2011] concerns three persons (John, Mary, and Suzy) who all want to go cycling on a tandem (represented by three axioms $jw$, $mw$, $sw$) and one assumes defeasibly that if one wants to go then s/he will go (represented by three rules $jw \Rightarrow jt$, $mw \Rightarrow mt$, $sw \Rightarrow st$). Clearly only two of them (no matter who) can actually go. Using *ASPIC* this is modeled with a constraint $\neg(jt \wedge mt \wedge st)$. Altogether this representation gives rise in *ASPIC* to the construction of ten arguments, three of which, concluding respectively $\neg jt$, $\neg mt$, and $\neg st$, are in mutual conflict each other. In the $\mathcal{RS}{-}AF(AT)$ representation the mutual conflict involves three nodes corresponding to three couples of arguments with conclusions $\{jt, mt\}$, $\{jt, st\}$, $\{mt, st\}$ respectively, thus making more explicit that the root of the conflicts is the incompatibility among the three possible pairs of tandem riders.

## 7 Conclusions

The ability to properly deal with generic contrariness is an important feature for structured argumentation, as it ensures a very high expressiveness and appears to enable novel developments like the use of multivalued logics for argument construction. In this paper we have focused on generic contrariness in *ASPIC*$^{+}$, an expressively rich and technically complex formalism whose properties are an important and evolving investigation topic, currently witnessing a great deal of interest. As an example, recently Dung [Dung, 2014] has proposed a new definition of attack relation[1] in *ASPIC*$^{+}$ in order to ensure the satisfaction of some axioms not respected by the attack relations considered in [Modgil and Prakken, 2013].

This paper contributes to the advancement of *ASPIC*$^{+}$ on a different side, by pointing out that the use of a generic contrariness relation without explicit negation can become problematic in some cases and by providing a solution based on two main pillars: a notion of closure of the contrariness relation at a set level and, accordingly, a notion of conflict and an $AF$ representation involving sets of arguments. The revised version of *ASPIC*$^{+}$ satisfies the same rationality postulates, while posing fewer requirements on the properties of argument ordering than in [Modgil and Prakken, 2013].

Starting from the results provided in this paper, many further technical developments can be envisaged, including, for instance, the characterization of the cases where our proposal and the one in [Modgil and Prakken, 2013] are equivalent, the identification of necessary/sufficient conditions for the problems pointed out in Sect. 3 to occur, and the study of the complexity properties of computing the closure of the contrariness relation and the induced $AF$ in specific instances of the *ASPIC*$^{+}$ framework. Further, it will be interesting to explore the relationships and, possibly, the intertranslatability between the abstract representation based on Dung's $AF$s in Def. 23 and alternative representations based on frameworks where attacks originating from sets of arguments are encompassed [Bochman, 2003; Nielsen and Parsons, 2006; Oren and Norman, 2008; Gabbay, 2009].

We remark however that the contribution of this paper goes beyond the technical advancement of *ASPIC*$^{+}$. Indeed, we believe that the notions proposed in this paper are useful as general conceptual tools in the study of rule-based structured argumentation and can enable fruitful theoretical developments, independently of the specific framework adopted. To exemplify, we plan to investigate the relationship between argumentation and the four-valued Belnap-Dunn logic [Belnap, 1977] and to exploit the set-based abstract representation to provide a unifying view of different rule-based argumentation systems available in the literature, including in particular the one proposed in [Vreeswijk, 1997] where a collective notion of conflict is encompassed.

## Acknowledgments

This paper was partially supported by the National Natural Science Foundation of China (No.61175058) and Zhejiang Provincial Natural Science Foundation of China (No. LY14F030014).

[1] While we used the original definition of *ASPIC*$^{+}$, our considerations and results also apply to the version of [Dung, 2014].

## References

[Amgoud *et al.*, 2006] L. Amgoud, L. Bodenstaff, M. Caminada, P. McBurney, S. Parsons, H. Prakken, J. van Veenen, and G. Vreeswijk. Final review and report on formal argumentation system. Technical Report Deliverable D2.6, ASPIC project IST-FP6-002307, 2006.

[Baroni *et al.*, 2011] P. Baroni, M. Caminada, and M. Giacomin. An introduction to argumentation semantics. *Knowledge Engineering Review*, 26(4):365–410, 2011.

[Baroni *et al.*, 2015] P. Baroni, M. Giacomin, and B. Liao. I don't care, I don't know ... I know too much! On incompleteness and undecidedness in abstract argumentation. In *Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation - Essays Dedicated to Gerhard Brewka on the Occasion of His 60th Birthday*, volume 9060 of *LNCS*, pages 265–280. Springer, 2015.

[Belnap, 1977] N. D. Belnap. How a computer should think. In G. Ryle, editor, *Contemporary aspects of philosophy*, pages 30–56. Oriel Press, 1977.

[Bochman, 2003] A. Bochman. Collective argumentation and disjunctive logic programming. *J. of Logic and Computation*, 13(3):405–428, 2003.

[Caminada and Amgoud, 2007] M. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artif. Intell.*, 171:286 – 310, 2007.

[Dung, 1995] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artif. Intell.*, 77(2):321–357, 1995.

[Dung, 2014] P. M. Dung. An axiomatic analysis of structured argumentation for prioritized default reasoning. In *Proc. of the 21st European Conf. on Artificial Intelligence (ECAI 2014)*, pages 267–272, 2014.

[Gabbay, 2009] D. M. Gabbay. Fibring argumentation frames. *Studia Logica*, 93(2-3):231–295, 2009.

[Modgil and Prakken, 2013] S. Modgil and H. Prakken. A general account of argumentation with preferences. *Artif. Intell.*, 195:361 – 397, 2013.

[Nielsen and Parsons, 2006] S. H. Nielsen and S. Parsons. A generalization of Dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In *Third Int. Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2006), Revised Selected and Invited Papers*, volume 4766 of *LNCS*, pages 54–73. Springer, 2006.

[Oren and Norman, 2008] N. Oren and T. J. Norman. Semantics for evidence-based argumentation. In *Proc. of the 2nd Int. Conf. on Computational Models of Argument (COMMA 2008)*, pages 276–284, 2008.

[Prakken, 2010] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument & Computation*, 1(2):93–124, 2010.

[Vreeswijk, 1997] G. Vreeswijk. Abstract argumentation systems. *Artif. Intell.*, 90(1-2):225–279, 1997.