Journal of Artificial Intelligence Research 56 (2016) 197-245 Submitted 07/15; published 06/16

Two Aspects of Relevance in Structured Argumentation: Minimality and Paraconsistency

Diana Grooters DIANAGROOTERS@GMAIL.COM
ORTEC Finance Rotterdam, The Netherlands

Henry Prakken H.PRAKKEN@UU.NL
Department of Information and Computing Sciences, Utrecht University
Faculty of Law, University of Groningen
The Netherlands

Abstract

This paper studies two issues concerning relevance in structured argumentation in the context of the $ASPIC^+$ framework, arising from the combined use of strict and defeasible inference rules. One issue arises if the strict inference rules correspond to classical logic. A longstanding problem is how the trivialising effect of the classical Ex Falso principle can be avoided while satisfying consistency and closure postulates. In this paper, this problem is solved by disallowing chaining of strict rules, resulting in a variant of the $ASPIC^+$ framework called $ASPIC^\star$, and then disallowing the application of strict rules to inconsistent sets of formulas. Thus in effect Rescher & Manor's paraconsistent notion of weak consequence is embedded in $ASPIC^\star$.

Another issue is minimality of arguments. If arguments can apply defeasible inference rules, then they cannot be required to have subset-minimal premises, since defeasible rules based on more information may well make an argument stronger. In this paper instead minimality is required of applications of strict rules throughout an argument. It is shown that under some plausible assumptions this does not affect the set of conclusions. In addition, circular arguments are in the new $ASPIC^\star$ framework excluded in a way that satisfies closure and consistency postulates and that generates finitary argumentation frameworks if the knowledge base and set of defeasible rules are finite. For the latter result the exclusion of chaining of strict rules is essential.

Finally, the combined results of this paper are shown to be a proper extension of classical-logic argumentation with preferences and defeasible rules.

1. Introduction

One tradition in the logical study of argumentation is to allow for arguments that combine strict and defeasible inference rules. This approach was introduced in AI by Pollock (1987, 1990, 1992, 1994, 1995), was studied in the past also by e.g. Lin and Shoham (1989), Simari and Loui (1992), Vreeswijk (1997), Prakken and Sartor (1997) and Garcia and Simari (2004) and is currently studied by e.g. Dung and Thang (2014), Dung (2014, 2016) and in work on the $ASPIC^+$ framework (Prakken, 2010; Modgil & Prakken, 2013, 2014). Strict inference rules are intended to capture deductively valid inferences, where the truth of the premises guarantee the truth of the conclusion. Defeasible inference rules are meant to capture presumptive inferences, where the premises create a presumption in favour of the conclusion, which can be refuted by evidence to the contrary. Much research in this tradition has shown that the idea of defeasible inference rules makes sense. For example, Pollock applied it to

formalize his theory of defeasible epistemic reasons, which includes reasons concerning perception, memory, enumerative induction, the statistical syllogism and temporal persistence. Moreover, several publications on $ASPIC^+$ use defeasible inference rules to formalise Walton (1996)-style presumptive argumentation schemes (Prakken, 2010; Modgil & Prakken, 2014) and to apply these to legal reasoning (Prakken, Wyner, Bench-Capon, & Atkinson, 2015) or policy debate (Bench-Capon, Prakken, & Visser, 2011). Finally, Garcia and Simari's (2004) Defeasible Logic Programming approach has been applied in many domains.

In this tradition, two issues arise concerning relevance, namely, minimality of arguments and paraconsistency in strict-rule application. We will study both issues in the context of the $ASPIC^+$ framework. The choice of $ASPIC^+$ for these purposes is justified by its framework nature, which allows the study of various classes of instantiations. Moreover, it has been shown that various other approaches can be reconstructed as instantiations of the $ASPIC^+$ framework. Prakken (2010) showed this for assumption-based argumentation as reconstructed by Dung, Mancarella, and Toni (2007) as an instance of abstract argumentation (Dung, 1995), and this result carries over to the original formulation of assumption-based argumentation (Bondarenko, Dung, Kowalski, & Toni, 1997) for all known semantics except semi-stable and eager semantics (cf. Caminada, Sa, Alcantara, & Dvorak, 2015). Furthermore, Modgil and Prakken (2013) reconstructed two forms of classical argumentation with premise attack as studied by Gorogiannis and Hunter (2011) and several uses of Tarskian abstract logics as studied by Amgoud and Besnard (2013) as instances of $ASPIC^+$. For these reasons, results in terms of $ASPIC^+$ are representative for large classes of argumentation systems.

$ASPIC^+$ is sometimes criticised on the fact that it allows for instantiations with bad properties but such criticism is besides the point, since it ignores the framework nature of $ASPIC^+$(Prakken & Modgil, 2012). Being a framework instead of a concrete system, $ASPIC^+$ is intended to allow the study of properties of various instantiations, such as whether they satisfy the rationality postulates of Caminada and Amgoud (2007). It is the very idea of the framework to allow for 'bad' instantiations so that they can be identified. Therefore, the framework cannot be criticised on the existence of such bad instantiations. Moreover, there is a growing body of results on 'good' instantiations of $ASPIC^+$ (Caminada & Amgoud, 2007; Prakken, 2010; Modgil & Prakken, 2013; Dung, 2014, 2016; Caminada, Modgil, & Oren, 2014; Grooters & Prakken, 2014; Wu & Podlaszewski, 2015) and this paper aims at identifying another class of good instantiations.

One relevance issue discussed in this paper is minimality of arguments. In deductive approaches to argumentation (e.g., Besnard & Hunter, 2008; Gorogiannis & Hunter, 2011; Amgoud & Besnard, 2013) arguments are required to have a subset-minimal set of premises. However, if arguments can apply defeasible inference rules, then this requirement is undesirable, since defeasible rules that are based on more information may well make an argument stronger. For example, *Observations done in ideal circumstances are usually correct* is stronger than *Observations are usually correct*. Note that this remark does not apply to strict inference rules, so it still makes sense to improve efficiency by requiring that strict inference rules are only applied to a subset-minimal set of formulas. So far, no system in the defeasible-rule tradition enforces this requirement. One contribution of this paper is to do so for the $ASPIC^+$ approach and to show that under some plausible conditions on its argument ordering, this does not affect the set of conclusions.

Another aspect of minimality is circularity. So far presentations of $ASPIC^+$ have not prevented arguments from repeating conclusions of their subarguments. Yet in argumentation theory circular arguments are generally regarded as fallacious, so it makes sense to exclude them. In this paper we will do so and prove that the results on the rationality postulates are not affected by it. Moreover, we will prove that excluding non-circular arguments has some computational benefits.

Another relevance issue arises if the strict inference rules are chosen to correspond to classical logic. Then a longstanding unsolved problem originally identified by Pollock (1994, 1995) is how the trivialising effect of the classical Ex Falso principle can be avoided when two arguments that use defeasible rules have contradictory conclusions. The problem is especially hard since any solution should arguably preserve satisfaction of the rationality postulates of consistency and strict closure (Caminada & Amgoud, 2007).

In a nutshell, the problem is as follows. Suppose two arguments have contradictory conclusions φ and $\neg\varphi$. If the strict inference rules include the Ex Falso principle that an inconsistent set implies any formula, then these two arguments can be combined into an argument for $\neg\psi$ for any formula ψ. This combined argument can potentially defeat any argument for ψ by applying the Ex Falso inference rule to their joint conclusions. So when there are arguments for contradictory conclusions, any other argument is potentially under threat, which is clearly undesirable, since the conflict about φ is in general unrelated to ψ.

Pollock (1994, 1995) thought that he had avoided such 'trivialising' arguments by allowing for multiple labellings, but Caminada (2005) showed that Pollock's solution does not fully avoid them. The problem is a genuine one, since there arguably is a real need for argumentation systems that allow for combinations of strict and defeasible inferences and that, moreover, allow for the full reasoning power of a deductive logic. Although for many cases less expressiveness may suffice, a full theory of the logic of argumentation cannot exclude the general case.

To solve the problem, two approaches are possible. One is to change the definitions of the argumentation framework, while the other is to derive the strict inference rules from a weaker logic than classical logic. The first approach is taken by Wu (2012) and Wu and Podlaszewski (2015), who for the $ASPIC^+$ framework require that for each argument the set of conclusions of all its subarguments are classically consistent. They show that this solution works for a restricted version of $ASPIC^+$ without preferences, but they give counterexamples to the consistency postulates for the case with preferences.

A second approach to solve the problem is to replace classical logic as the source for strict rules with a weaker, monotonic paraconsistent logic, in order to invalidate the Ex Falso principle as a valid strict inference rule. This paper explores this possibility. We first show that two well-known paraconsistent logics, the system C_ω of Da Costa (1974) and the Logic of Paradox of Priest (1979, 1989), cannot be used for these purposes, since they induce violation of the postulate of indirect consistency. We then show that using Rescher and Manor's (1970) paraconsistent consequence notion satisfies all closure and consistency postulates and also avoids trivialisation. While thus initially taking the second approach, we will have to combine it with the first approach (changing the definitions) since it will turn out that chaining strict rules in arguments has to be disallowed. This change in turn motivates a new interpretation of Caminada and Amgoud's (2007) strict-closure postulate

and the introduction of a new rationality postulate of 'logical closure'. This contribution of our paper is based on and extends results of Grooters and Prakken (2014).

After making these contributions, we will argue that in combination they shed light on the relation between the adapted version of $ASPIC^+$ and classical argumentation as studied by Besnard and Hunter (2008) and Gorogiannis and Hunter (2011), in which arguments are essentially classical proofs from consistent and subset-minimal subsets of a classical knowledge base. For two of their versions of classical argumentation with premise attack the adapted version of $ASPIC^+$ is shown to a be proper extension with defeasible rules and preferences. This observation justifies a combined treatment of both issues (minimality of arguments and paraconsistency) in the same paper.

Caminada, Carnielli, and Dunne (2012) formulated a new set of rationality postulates in addition to those of Caminada and Amgoud (2007), to characterise cases under which the trivialisation problem is avoided (called the postulates of non-interference and crash-resistance). Wu (2012) and Wu and Podlaszewski (2015) prove for their adaptation of $ASPIC^+$ with consistent arguments that these new postulates are satisfied for complete semantics. However, we will not attempt to prove Caminada et al.'s (2012) postulates, for two reasons. First, we want to obtain results for other semantics as well and, second, we will argue in Section 10 that Caminada et al.'s postulates in fact capture a stronger intuitive notion than the one we study in this paper.

The remainder of this article is organised as follows. First in Section 2 the $ASPIC^+$ framework is summarised and in Section 3 the rationality postulates of Caminada and Amgoud (2007) are presented. Then in Section 4 the trivialisation problem is illustrated in more detail, after which in Section 5 instantiations of $ASPIC^+$ with the paraconsistent logics LP and C_ω are studied as an attempt to avoid trivialisation in the face of inconsistency. It will be shown that these instantiations violate the rationality postulate on indirect consistency. Then in Section 6 Rescher and Manor's (1970) paraconsistent consequence notion is introduced as another attempt to avoid trivialisation. It turns out that its embedding in $ASPIC^+$ requires an adaptation of the $ASPIC^+$ framework into a framework called $ASPIC^\star$, which disallows chaining of strict rules, which in turn motivates new notions of strict closure and indirect consistency. Then in Section 7 the first main contribution of this paper is proved: satisfaction of the closure and consistency postulates by the instantiation of the $ASPIC^\star$ framework with Rescher and Manor's consequence notion. In Section 8 the second and third main contribution is presented: an equivalence result between versions of $ASPIC^\star$ with and without minimality constraints on strict inferences, and proofs that show that a version of $ASPIC^\star$ that excludes circular arguments is well-behaved. In Section 9 we present our fourth main result, namely, that $ASPIC^\star$ with minimal arguments properly generalises two versions of classical argumentation. Finally, in Section 10 we discuss our results and put them in the context of related work.

2. The $ASPIC^+$ Framework

In this section, the $ASPIC^+$ framework is reviewed. Since it makes use of Dung's (1995) theory of abstract argumentation, that theory is first briefly summarised. An *abstract argumentation framework* (AF) is a pair $(\mathcal{A}, \mathcal{D})$, where $\mathcal{A}$ is a set of *arguments* and $\mathcal{D} \subseteq \mathcal{A} \times \mathcal{A}$ is a binary relation of *defeat*. An argument A defeats argument B if $(A, B) \in \mathcal{D}$. A

set S of arguments defeats an argument B if there is an argument $A \in S$ such that A defeats B. A set S defeats a set S' if there is an argument $A \in S'$ such that S defeats A. A set of arguments is said to be *conflict-free* if it does not attack itself; otherwise it is *conflicting*. A set $S \subseteq \mathcal{A}$ defends an argument $A \in \mathcal{A}$ iff for each $Bin\mathcal{A}$ that defeats A there exists a $C \in S$ that defeats B. A set is *admissible* if it is conflict-free and defends itself by attacking each argument attacking S. Each argumentation framework has zero or more extensions, which intuitively are maximal sets of arguments that can be accepted together since they are conflict-free and defend all their members against attacks. Formally, *extensions* are admissible sets with some additional properties. They can be defined according to Dung's 'characteristic function'.

Definition 2.1. [**Dung's characteristic function** F] $F_{AF} : 2^{\mathcal{A}} \rightarrow 2^{\mathcal{A}}$ such that $F_{AF}(S) = \{A \in \mathcal{A} | A \text{ is defended by } S\}$.

Henceforth the subscript AF will be omitted if there is no danger of confusion.

Definition 2.2. [**Extensions of abstract argumentation frameworks**] For any $AF = (\mathcal{A}, \mathcal{D})$ and any $E \subseteq \mathcal{A}$:

- E is *conflict-free* iff there are no A, B in E such that $(A, B) \in \mathcal{D}$.
- E is *admissible* iff E is conflict-free and E defends each $A \in E$.
- E is a *complete extension* of AF iff E is conflict-free and $F_{AF}(E) = E$.
- E is a *preferred extension* of AF iff E is a set-inclusion-maximal complete extension of AF.
- E is a *stable extension* of AF iff E is conflict-free and for all $A \notin E$ there exists a $B \in E$ such that B defeats A.
- E is the *grounded extension* of AF iff E is the set-inclusion-minimal complete extension of AF.

Finally, for $T \in \{\text{complete, preferred, grounded, stable}\}$, X is *sceptically* or *credulously* T-justified if X belongs to all, respectively at least one, T extension. Other notions of extensions have been proposed in the literature but in this paper we confine ourselves to these four notions.

The *ASPIC*$^+$ framework (Prakken, 2010; Modgil & Prakken, 2013) gives structure to Dung's arguments and defeat relation. As in the work of Vreeswijk (1997) it defines arguments as directed acyclic inference graphs formed by applying strict or defeasible inference rules to premises formulated in some logical language. Intuitively, strict rules *guarantee* the truth of their consequent if their antecedents are true, while defeasible rules only create a *presumption* in favour of the truth of their consequent if their antecedents are true. Arguments can be attacked on their (ordinary) premises and on their applications of defeasible inference rules. Some attacks succeed as *defeats*, which is partly determined by preferences. The acceptability status of arguments is then defined by applying any of Dung's (1995) semantics for abstract argumentation frameworks to the resulting set of arguments with its defeat relation.

Below the special case with symmetric negation of the version of $ASPIC^+$ defined by Modgil and Prakken (2013) is presented, but with some minor improvements. Nontrivial improvements will be indicated below when they are made.

$ASPIC^+$ is not a system but a framework for specifying systems. As said above, the framework is intended to allow the study of properties of instantiations, such as whether they satisfy the rationality postulates of Caminada and Amgoud (2007). To this end it defines the notion of an abstract *argumentation system* as a structure consisting of a logical language $\mathcal{L}$ with a unary negation symbol $\neg$, a set $\mathcal{R}$ consisting of two disjoint subsets $\mathcal{R}_s$ and $\mathcal{R}_d$ of strict and defeasible inference rules, and a naming convention n in $\mathcal{L}$ for defeasible rules in order to talk in $\mathcal{L}$ about the applicability of defeasible rules. All these elements are left undefined in general and have to be specified for each specific instantiation.

Definition 2.3. [**Argumentation systems**] An *argumentation system* is a triple $AS = (\mathcal{L}, \mathcal{R}, n)$ where:

- $\mathcal{L}$ is a nonempty logical language with a unary negation symbol $\neg$.
- $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$ 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 φ_i, φ are meta-variables ranging over wff in $\mathcal{L}$), and $\mathcal{R}_s \cap \mathcal{R}_d = \emptyset$.
- $n : \mathcal{R}_d \longrightarrow \mathcal{L}$ is a naming convention for defeasible rules.

Informally, $n(r)$ is a wff in $\mathcal{L}$, which says that rule $r \in \mathcal{R}$ is applicable. We write $\psi = -\varphi$ just in case $\psi = \neg\varphi$ or $\varphi = \neg\psi$. Note that $-$ is not part of the logical language $\mathcal{L}$ but a metalinguistic function symbol to obtain more concise definitions. Furthermore, if there is no danger for confusion, we will sometimes write the sequence of antecedents of a strict or defeasible rule as a set.

Example 2.1. An example argumentation system is with

$\mathcal{L} = \{p, \neg p, q, \neg q, r, \neg r, s, \neg s, t, \neg t, r_1, r_2, \neg r_1, \neg r_2\}$,
$\mathcal{R}_s = \{p, r \rightarrow s; \rightarrow \neg r_1\}$, $\mathcal{R}_d = \{q \Rightarrow r; t \Rightarrow \neg s\}$,
$n(q \Rightarrow r) = r_1$ and $n(t \Rightarrow \neg s) = r_2$.

$ASPIC^+$ as a framework abstracts from the origins of the strict and defeasible rules. Several ways to identify rules are possible. One way, quite usual in AI, is to let the rules express domain-specific knowledge. For example, the strict rules could contain terminological knowledge such as 'bachelors are not married', and the defeasible rules could contain defeasible generalisations such as 'Birds fly' or defeasible norms such as 'Thou shalt not lie'. Another way is to base the rules on general accounts of deductive and defeasible reasoning. For example, the strict rules might be chosen to correspond to a monotonic logic and the defeasible rules might be instantiated with argument schemes (Walton, 1996). These two ways to identify inference rules are only pragmatically different; formally, the $ASPIC^+$ framework treats rules as inference rule regardless of their origin. In this paper we abstract from the origin of the defeasible rules and we focus on the choice of strict rules. We are in particular concerned with instantiations of $ASPIC^+$ in which the strict rules are chosen to correspond to a monotonic logic (although several results will apply more generally). In

such instantiations $\mathcal{R}_s$ is defined as follows, given a monotonic consequence notion $\vdash_L$ for a logic L:

- $\mathcal{R}_s = \{S \to \varphi \mid S \vdash_L \varphi \text{ and } S \text{ is finite}\}$

If $\mathcal{R}_s$ is defined in this way over the logical language of L, then we say that $\mathcal{R}_s$ *corresponds to* logic L.

Definition 2.4. [**Knowledge bases**] A *knowledge base* in an $AS = (\mathcal{L}, \mathcal{R}, n)$ is a set $\mathcal{K} \subseteq \mathcal{L}$ consisting of two disjoint subsets $\mathcal{K}_n$ and $\mathcal{K}_p$ (the *necessary* and *ordinary premises*).

Intuitively, the necessary premises are certain knowledge and thus cannot be attacked, whereas the ordinary premises are uncertain and thus can be attacked.

Definition 2.5. [**Consistency and strict closure**] For any $X \subseteq \mathcal{L}$, let the closure of X under strict rules, denoted $Cl_{\mathcal{R}_s}(X)$, be the smallest set containing X and the consequent of any strict rule in $\mathcal{R}_s$ whose antecedents are in $Cl_{\mathcal{R}_s}(X)$. Then a set $X \subseteq \mathcal{L}$ is

- *directly consistent* iff $\nexists\ \psi$, $\varphi \in X$ such that $\psi = -\varphi$;
- *indirectly consistent* iff $Cl_{\mathcal{R}_s}(X)$ is directly consistent.

Example 2.2. In our example argumentation system, an example of a directly inconsistent set is $\{p, \neg p\}$ and an example of a directly consistent but indirectly inconsistent set is $\{p, r, \neg s\}$. Finally, an example of a closure under strict rules is $Cl_{\mathcal{R}_s}(\{p, r\}) = \{p, r, s, \neg r_1\}$.

Arguments can be constructed step-by-step from knowledge bases by chaining inference rules into directed acyclic graphs (or trees if no formula is used more than once). In what follows, for a given argument the function `Prem` returns all its premises, `Conc` returns its conclusion and `Sub` returns all its sub-arguments, while `TopRule` returns the last rule used in the argument.

Definition 2.6. [**Argument**] An *argument* A on the basis of a knowledge base $\mathcal{K}$ in an argumentation system $(\mathcal{L}, \mathcal{R}, n)$ is:

1. φ if $\varphi \in \mathcal{K}$ with:
   $\texttt{Prem}(A) = \{\varphi\}$;
   $\texttt{Conc}(A) = \varphi$;
   $\texttt{Sub}(A) = \{\varphi\}$;
   $\texttt{TopRule}(A) =$ undefined.

2. $A_1, \ldots A_n \to/\Rightarrow \psi$ if $A_1, \ldots, A_n$ are arguments such that R_s/R_d contains the strict/defeasible rule $\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \to/\Rightarrow \psi$, with:
   $\texttt{Prem}(A) = \texttt{Prem}(A_1) \cup \ldots \cup \texttt{Prem}(A_n)$,
   $\texttt{Conc}(A) = \psi$,
   $\texttt{Sub}(A) = \texttt{Sub}(A_1) \cup \ldots \cup \texttt{Sub}(A_n) \cup \{A\}$;
   $\texttt{TopRule}(A) = \texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \to/\Rightarrow \psi$.

Each of these functions $\texttt{Func}$ are also defined on sets of arguments $S = \{A_1, \ldots, A_n\}$ as follows: $\texttt{Func}(S) = \texttt{Func}(A_1) \cup \ldots \cup \texttt{Func}(A_n)$. If an argument only uses strict rules, the argument is said to be *strict*, otherwise it is *defeasible*. If an argument only has necessary premises, then the argument is *firm*, otherwise it is *plausible*. For any argument A we define $\texttt{Prem}_n(A) = \texttt{Prem}(A) \cap \mathcal{K}_n$ and $\texttt{Prem}_p(A) = \texttt{Prem}(A) \cap \mathcal{K}_p$. The set of all arguments that just consist of a necessary premise are denoted with $NP(S)$.

Example 2.3. If our example argumentation system is combined with a knowledge base with $\mathcal{K}_n = \{p\}$ and $\mathcal{K}_p = \{q, t\}$, then the following arguments can be constructed:

$$\begin{array}{lll} A_1 = & p & \qquad A_5 = \quad A_3 \Rightarrow \neg s \\ A_2 = & q & \qquad A_6 = \quad A_1, A_4 \rightarrow s \\ A_3 = & t & \qquad A_7 = \quad A_5 \rightarrow \neg r_1 \\ A_4 = & A_2 \Rightarrow r & \end{array}$$

Argument A_1 is strict and firm, A_2 and A_3 are strict and plausible, and the remaining arguments are all defeasible and plausible.

In Figure 1 these arguments are visualised. The type of a premise is indicated with a superscript and defeasible inferences are displayed with dotted lines. The dotted boxes and thick arrows will be explained below in Example 2.4.

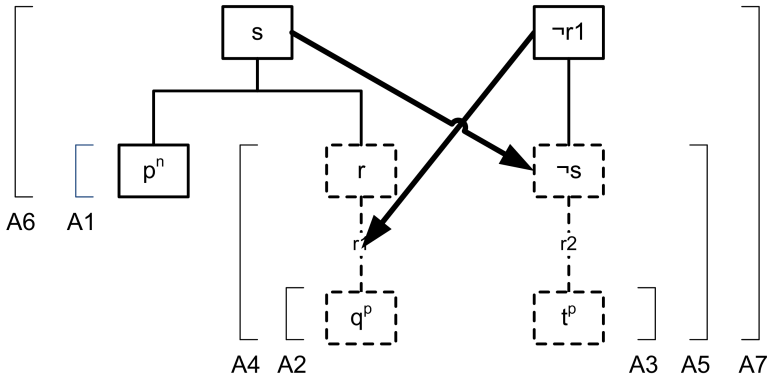

Figure 1: Arguments and attacks of Example 2.1. The premises are at the bottom and the conclusion at the top of the tree. Thin vertical links between boxes are inferences while the thick diagonal links are attacks. The type of a premise is indicated with a superscript and defeasible inferences, underminable premises and rebuttable conclusions are displayed with dotted lines.

Arguments can be attacked in three ways: on their ordinary premises (undermining attack), on a defeasible inference (undercutting attack) or on the conclusion of a defeasible inference.

Definition 2.7. [**Attack**] An argument A *attacks* an argument B iff A *undercuts*, *rebuts* or *undermines* B, where:

- *A undercuts* argument B (on B') iff $\texttt{Conc}(A) = -n(r)$ and $B' \in \texttt{Sub}(B)$ such that B''s top rule r is defeasible.
- *A rebuts* argument B (on B') iff $\texttt{Conc}(A) = -\varphi$ for some $B' \in \texttt{Sub}(B)$ of the form $B''_1, \ldots, B''_n \Rightarrow \varphi$.
- *A undermines* argument B (on φ) iff $\texttt{Conc}(A) = -\varphi$ for some $\varphi \in \texttt{Prem}_p(B)$.

Example 2.4. In our running example A_6 rebuts A_5 and therefore also A_6 rebuts A_7 on A_5 (since A_5 is a subargument of A_7). Note that A_5 does not rebut A_6 since A_6 has a strict top rule. Furthermore, A_7 undercuts A_4 and A_6 on A_4. In Figure 1 rebuttable conclusions are visualised with dotted boxes and direct defeat relations are displayed with thick arrows. Note that the indirect attacks of A_6 on A_7 and A_7 on A_6 are not explicitly visualised.

An argument A is a *basic fallible argument*[1] iff $A \in K_p$ or $\texttt{TopRule}(A) \in \mathcal{R}_d$. A basic fallible argument is thus an argument that has a defeasible top rule or equates an ordinary premise and so can be attacked on its final conclusion or inference. The set of all basic fallible arguments for a set of arguments S is denoted by $FA(S)$.

Argumentation systems plus knowledge bases form argumentation theories. These are in turn combined with a preference ordering on the set of all arguments constructible in the theory, to induce structured argumentation frameworks. Like the elements of argumentation systems, the nature of the $ASPIC^+$ argument ordering is undefined in general and has to be specified for each specific instantiation.

Definition 2.8. [**Structured Argumentation Frameworks**] Let AT be an *argumentation theory* $(AS, \mathcal{K})$. A *structured argumentation framework (*SAF*)* defined by AT, is a triple $\langle \mathcal{A}, \mathcal{C}, \preceq \rangle$ where $\mathcal{A}$ is the set of all finite arguments constructed from $\mathcal{K}$ in AS, $\preceq$ is a binary relation on $\mathcal{A}$, and $(X, Y) \in \mathcal{C}$ iff X attacks Y.

Unlike Modgil and Prakken (2013) we will not consider versions of $ASPIC^+$ that require arguments to have consistent premises (except briefly in Section 9 for purposes of comparison). In our approach strict arguments with inconsistent premises will be handled by our choice to let $\mathcal{R}_s$ correspond to a paraconsistent logic, which will prevent trivialisation, i.e., it will prevent systems that generate arguments for any random conclusion from contradictions. Furthermore, we will leave it to the users whether they want to ensure that there are no defeasible rules with inconsistent antecedents. This can be left to the user since allowing defeasible rules with inconsistent antecedents does not cause trivialisation. Furthermore, all results proved in this paper still hold whether defeasible rules with inconsistent antecedents are excluded or not.

The notion of *defeat* can then be defined as follows. Undercutting attacks succeed as *defeats* independently of preferences over arguments, since they are meant to express exceptions to defeasible inference rules. Rebutting and undermining attacks succeed only if the attacked argument is not stronger than the attacking argument ($A \prec B$ is defined as usual as $A \preceq B$ and $B \not\preceq A$).

Definition 2.9. [**Defeat**] *A defeats B* iff:

1. This is a renaming of Dung and Thang's (2014) notion of a *basic defeasible* argument.

1. A undercuts B; or

2. A rebuts/undermines B on B' and $A \not\prec B'$.

Example 2.5. In our running example A_6 defeats A_5 and A_7 unless $A_6 \prec A_5$. Furthermore, with any preference relation between A_4 and A_7, even with $A_7 \prec A_4$, we have that A_7 defeats A_4 (and thus A_7 also defeats A_6).

SAFs generate abstract argumentation frameworks in the sense of Dung (1995), which can then be used to evaluate arguments and their conclusions:

Definition 2.10. [**Argumentation frameworks**]
An *abstract argumentation framework (AF) corresponding to a* $SAF = \langle \mathcal{A}, \mathcal{C}, \preceq \rangle$ is a pair $(\mathcal{A}, D)$ such that D is the defeat relation on $\mathcal{A}$ determined by SAF.

Let $T \in$ {complete, preferred, grounded, stable} and let $\mathcal{L}$ be from the AT defining SAF. A wff $\varphi \in \mathcal{L}$ is *sceptically T-justified* in SAF if φ is the conclusion of a sceptically T-justified argument, and *credulously T-justified* in SAF if φ is not sceptically T-justified and is the conclusion of a credulously T-justified argument.

Example 2.6. Suppose in our running example that $\mathcal{D}$ is such that A_6 defeats A_5 and A_7 while A_7 defeats A_4 and A_6. The resulting AF is visualised twice in Figure 2. The grounded extension is $\{A_1, A_2, A_3\}$ while there are two preferred extensions $E_1 = \{A_1, A_2, A_3, A_4, A_6\}$ and $E_2 = \{A_1, A_2, A_3, A_5, A_7\}$. Both preferred extensions are also stable. The two preferred extensions are visualised in Figure 2: members of an extension are coloured white.

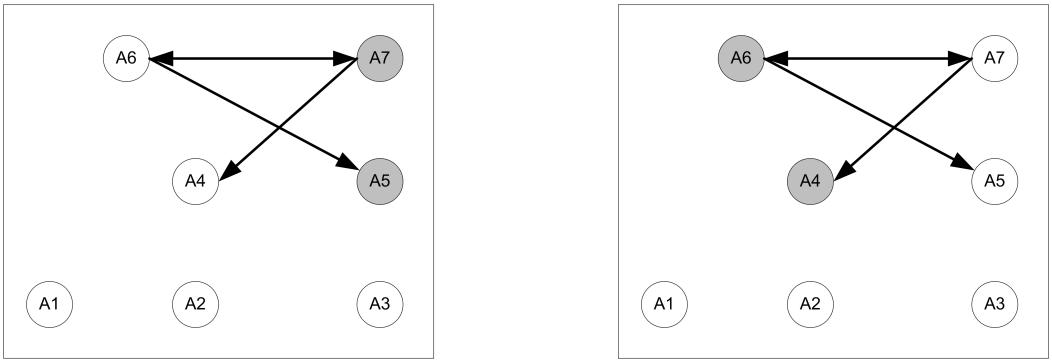

Figure 2: Two preferred extensions of the Dung AF of Example 2.1.

We finally need the notion of a strict continuation of a set arguments, which we define in a slightly different way than Modgil and Prakken (2013). The new definition is arguably simpler but does not affect the proofs of Modgil and Prakken. It identifies arguments that are formed by extending a set of arguments with only strict inferences into a new argument, so that the new argument can only be attacked on the arguments that it extends.

Definition 2.11. [**Strict continuations**] The set of *strict continuations* of a set of arguments from $\mathcal{A}$ is the smallest set satisfying the following conditions:

1. Any argument A is a strict continuation of $\{A\}$.
2. If $A_1, \ldots, A_n$ and $S_1, \ldots, S_n$ are such that for each $i \in \{1, \ldots, n\}$, A_i is a strict continuation of S_i and $B_1, \ldots, B_n$ are all strict-and-firm arguments, and $\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n), \texttt{Conc}(B_1), \ldots, \texttt{Conc}(B_m) \rightarrow \varphi$ is a strict rule in R_s, then $A_1, \ldots, A_n, B_1, \ldots, B_n \rightarrow \varphi$ is a strict continuation of $S_1 \cup \ldots \cup S_n$.

If argument A is a strict continuation of arguments $A_1, \ldots, A_n$, then A is a *strict argument over* $\{\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n)\}$.

Example 2.7. In our running example all arguments are strict continuations of themselves while A_6 is a strict continuation of $\{A_1, A_4\}$ and A_7 is a strict continuation of A_5. Suppose we temporarily add a strict rule $p, s \rightarrow \neg r_2$ to $\mathcal{R}_s$. Then $A_8 = A_1, A_6 \rightarrow \neg r_2$ is a strict continuation of $\{A_6\}$.

3. The Rationality Postulates

Extensions of abstract argumentation frameworks are intuitively maximal sets of arguments that can be rationally accepted together given such frameworks. Dung's (1995) various semantics, yielding different types of extensions, can be seen as various alternative ways to formalize rationality constraints on acceptable sets of arguments. When arguments have structure, additional rationality constraints can defined for extensions above those of abstract argumentation semantics. Caminada and Amgoud (2007) proposed the following rationality postulates for structured argumentation.

> **Subargument closure:** For every extension E, if an argument A is in E then all subarguments of A are in E.
>
> **Closure under strict rules:** For every extension E, the set $\texttt{Conc}(E)$ is closed under application of strict rules.
>
> **Direct consistency:** For every extension E, the set $\texttt{Conc}(E)$ is directly consistent.
>
> **Indirect consistency:** For every extension E, the set $\texttt{Conc}(E)$ is indirectly consistent.

Note that closure under strict rules and direct consistency together imply indirect consistency.

Modgil and Prakken (2013) identify a set of conditions under which *ASPIC*$^+$ satisfies all four postulates. The first condition is that the set of strict rules is either closed under transposition or closed under contraposition.

Definition 3.1. [**Closure under transposition, (Modgil & Prakken, 2013)**] A set of strict rules $\mathcal{R}_s$ is said to be *closed under transposition* if for each rule $\varphi_1, \ldots, \varphi_n \rightarrow \psi$ in $\mathcal{R}_s$ all the rules of the form $\varphi_1, \ldots, \varphi_{i-1}, -\psi, \varphi_{i+1}, \ldots, \varphi_n \rightarrow -\varphi_i$ for any $-\varphi_i$ and $-\psi$[2] also belong to $\mathcal{R}_s$. An argumentation theory $(AS, \mathcal{K})$ is *closed under transposition* if the strict rules $\mathcal{R}_s$ of AS are closed under transposition.

2. Note that a wff φ can have more than one ψ such that $\varphi = -\psi$. For example, $\neg p = -p$ and $\neg p = -\neg\neg p$.

Definition 3.2. [**Closure under contraposition, (Modgil & Prakken, 2013)**] An argumentation system is said to be *closed under contraposition* if for all $X \subseteq \mathcal{L}$, all $s \in X$ and all φ it holds that if $\varphi \in Cl_{\mathcal{R}_s}(X)$ then $-\psi \in Cl_{\mathcal{R}_s}(X \backslash \{\psi\} \cup \{-\varphi\})$ for any $-\varphi$ and $-\psi$. An argumentation theory $(AS, \mathcal{K})$ is *closed under contraposition* if the argumentation system AS is closed under contraposition.

The second condition states that the argument ordering should have the following properties:

Definition 3.3. [**Reasonable argument ordering, (Modgil & Prakken, 2013)**] $\preceq$ is a *reasonable argument ordering* if and only if:

- $\forall A, B$, if A is strict and firm and B is plausible or defeasible, then $B \prec A$;
- $\forall A, B$, if B is strict and firm, then $B \not\prec A$;
- $\forall A, A', B, C$, such that $C \prec A$, $A \prec B$ and A' is a strict continuation of $\{A\}$, then $C \prec A'$, $A' \prec B$;
- Let $\{C_1, \ldots, C_n\}$ be a finite subset of $\mathcal{A}$ and for $i = 1, \ldots, n$ let $C^{+/i}$ be some strict continuation of $\{C_1, \ldots, C_{i-1}, C_{i+1}, \ldots, C_n\}$. Then it is not the case that $\forall i,\ C^{+/i} \prec C_i$.

Modgil and Prakken (2013) identify several types of argument orderings that are reasonable.

The third condition is *axiom consistency*.

Definition 3.4. [**Axiom consistent, (Modgil & Prakken, 2013)**] An argumentation theory is *axiom consistent* if and only if $Cl_{\mathcal{R}_s}(\mathcal{K}_n)$ is consistent.

Modgil and Prakken (2013) prove that any argumentation theory that satisfies all three conditions only induces extensions that satisfy all four rationality postulates.

4. The Trivialisation Problem

In this section we illustrate the trivialisation problem in more detail. The following abstract example illustrates the problems that can arise if the strict rules of an argumentation system correspond to classical logic, i.e. $X \rightarrow \varphi \in \mathcal{R}_s$ if and only if $X \vdash \varphi$ and X is finite (where $\vdash$ denotes classical consequence).

Example 4.1. Let $\mathcal{R}_d = \{p \Rightarrow q;\ r \Rightarrow \neg q;\ t \Rightarrow s\}$, $\mathcal{K}_p = \emptyset$ and $\mathcal{K}_n = \{p, r, t\}$, while $\mathcal{R}_s$ corresponds to classical logic. Then the corresponding AF includes the following arguments:

A_1: p $\quad$ A_2: $A_1 \Rightarrow q$
B_1: r $\quad$ B_2: $B_1 \Rightarrow \neg q$ $\quad$ C: $A_2, B_2 \rightarrow \neg s$
D_1: t $\quad$ D_2: $D_1 \Rightarrow s$

Figure 3 displays these arguments and their attack relations. Dotted lines indicate defeasible inferences and dotted boxes indicate rebuttable conclusions. Argument C attacks D_2.

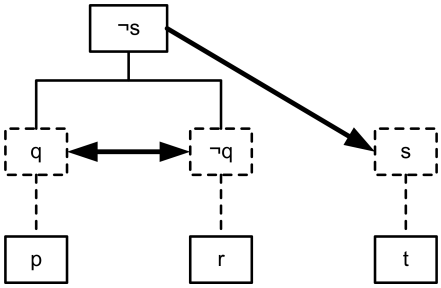

Figure 3: Illustrating trivialisation

Whether C defeats D_2 depends on the argument ordering but plausible argument orderings are possible in which $C \not\prec D_2$ and so C defeats D_2. This is problematic, since s can be any formula, so any defeasible argument unrelated to A_2 or B_2, such as D_2, can, depending on the argument ordering, be defeated by C. Clearly, this is extremely harmful, since the existence of just a single case of mutual rebutting attack, which is very common, could trivialise the system. It should be noted that simply disallowing application of strict rules to inconsistent sets of formulas does not help, since then an argument for $\neg s$ can still be constructed as follows:

A_3: $A_2 \rightarrow q \vee \neg s$
C': $A_3, B_2 \rightarrow \neg s$

Note that argument C' does not apply any strict inference rule to an inconsistent set of formulas.

This example suggests the following formalisation of the property of trivialisation.

Definition 4.1 (Trivialising argumentation systems)**.** An argumentation system AS is *trivialising* iff for all $\varphi, \psi \in \mathcal{L}$ and all knowledge bases $\mathcal{K}$ such that $\{\varphi, \neg\varphi\} \subseteq \mathcal{K}$ a strict argument on the basis of $\mathcal{K}$ can be constructed in AS with conclusion ψ.

We are then interested in defining classes of non-trivialising argumentation systems. The argumentation system in our example is clearly trivialising since $\mathcal{R}_s$ contains strict rules $\varphi, \neg\varphi \rightarrow \psi$ for all $\varphi, \psi \in \mathcal{L}$.

Example 4.1 does not cause any problems for preferred or stable semantics, since A_2 and B_2 attack each other and at least one of these attacks will (with non-circular argument orderings) succeed as defeat. Therefore, all preferred or stable extensions contain either A_2 or B_2 but not both. Since both A_2 and B_2 attack C (by directly attacking one of its subarguments), C is for each preferred or stable extension defeated by at least one argument in the extension, so C is not in any of these extensions, so D_2 is in all these extensions. This is intuitively correct since there is no connection between D_2 and the arguments A_2 and B_2.

In fact, the only semantics defined by Dung (1995) that has problems with Example 4.1 is grounded semantics. Since both A_2 and B_2 defeat each other, neither of them is in the grounded extension. So that extension does not defend D_2 against C and therefore does not contain D_2.

Pollock (1994, 1995) thought that the just-given line of reasoning for preferred semantics suffices to show that his recursive-labelling approach (which was later in Jakobovits & Vermeir, 1999 proved to be equivalent to preferred semantics) adequately deals with this problem. However, Caminada (2005) showed that the example can be extended in ways that also cause problems for preferred and stable semantics. Essentially, he replaced the facts p and r with defeasible arguments for p and r and let both these arguments be defeated by a self-defeating argument. On the one hand, such self-defeating arguments cannot be in any extension, since extensions are conflict free. However, if a self-defeating argument is not defeated by other arguments, it prevents any argument that it defeats from being acceptable with respect to an extension. In our example, if both A_2 and B_2 are defeated by a self-defeating argument that is otherwise undefeated, then neither A_2 not B_2 is in any extension, so no argument in an extension defends D_2 against C.

A critic of $ASPIC^+$ or Pollock's approach might argue that the problem is caused by the combination of strict (i.e., deductive) and defeasible inference rules. Indeed, in classical argumentation (Besnard & Hunter, 2008; Gorogiannis & Hunter, 2011) the problem can be easily avoided by requiring that the premises of an argument are consistent. However, there are reasons to believe that classical logic is too strong to be able to model all forms of defeasible reasoning; see, for instance, the discussions by Brewka (1991), Ginsberg (1994) and Prakken (2012). Furthermore, in assumption-based argumentation (ABA) as reconstructed by Dung et al. (2007), which only has strict inference rules but does not require them to be classical, and which does not require that the premises of arguments are consistent, the problem may or may not arise depending on how it is instantiated. When reconstructed in $ASPIC^+$ as by Prakken (2010), ABA arguments are built from ordinary premises $\mathcal{K}_p$ and strict inference rules $\mathcal{R}_s$. The following example (in the notation of the $ASPIC^+$ reconstruction of ABA) shows that the trivialisation problem can also arise in ABA.

Example 4.2. Take $\mathcal{K}_p = \{p, \neg p, s\}$ and let $\mathcal{R}_s$ correspond to classical logic, i.e, $S \rightarrow \varphi \in \mathcal{R}_s$ iff S is a finite set of wff that classically implies φ. Then the following arguments can be constructed.

$A : p$
$B : \neg p$
$C : A, B \rightarrow \neg s$
$D : s$

The 'trivialising' argument C prevents argument D from being in any extension.

The problem now is to instantiate and/or redefine $ASPIC^+$ in a way that avoids the trivialising effects of strict inferences from an inconsistent set, while still satisfying the rationality postulates of Caminada and Amgoud (2007).

5. Instantiating $ASPIC^+$ with Two Well-Known Paraconsistent Togics

As said in the introduction, one way to avoid trivialisation is to derive the strict rules of $ASPIC^+$ from a paraconsistent logic. A logical consequence relation $\vdash_L$ is said to be paraconsistent if it is not 'explosive', i.e. when it does not hold for all A and B that

$\{A, \neg A\} \vdash_L B$. In this section we investigate this strategy for two well-known paraconsistent logics, the Logic of Paradox of Priest (1979, 1989) and the system C_ω of Da Costa (1974). Another well-known paraconsistent logic is the family of relevant logics. However, this logic is nonmonotonic (Read, 1988, p. 100). This is a problem since the idea of $ASPIC^+$ is that if the strict rules are based on a logic, this logic is monotonic. For this reason, relevance logics are not considered in this paper.

5.1 Logic of Paradox

The Logic of Paradox (Priest, 1979, 1989) is obtained by relaxing the assumption of classical propositional logic that a sentence cannot be both true and false. Sentences in the Logic of Paradox (LP) can have two truth values instead of one. The set of possible truth values is $\{\{1\}, \{0\}, \{0, 1\}\}$, where $\{0, 1\}$ is the paradoxical 'true' and 'false'.

The semantics for the propositional version of LP is as follows.

1. (a) $1 \in v(\neg A) \Leftrightarrow 0 \in v(A)$
   (b) $0 \in v(\neg A) \Leftrightarrow 1 \in v(A)$
2. (a) $1 \in v(A \wedge B) \Leftrightarrow 1 \in v(A)$ and $1 \in v(B)$
   (b) $0 \in v(A \wedge B) \Leftrightarrow 0 \in v(A)$ or $0 \in v(B)$
3. (a) $1 \in v(A \vee B) \Leftrightarrow 1 \in v(A)$ or $1 \in v(B)$
   (b) $0 \in v(A \vee B) \Leftrightarrow 0 \in v(A)$ and $0 \in v(B)$
4. (a) $1 \in v(A \rightarrow B) \Leftrightarrow 0 \in v(A)$ or $1 \in v(B)$
   (b) $0 \in v(A \rightarrow B) \Leftrightarrow 1 \in v(A)$ and $0 \in v(B)$

An interpretation is a *model* of a formula f if and only if $1 \in v(f)$ holds in that interpretation. It is a model of a set of formulas if and only if it is a model of every formula in the set. The semantical notion of logical consequence is defined as follows:

- $\Sigma \models_{LP} A \;\Leftrightarrow$ for all evaluations v either $1 \in v(A)$ or for some $B \in \Sigma$, $v(B) = \{0\}$

It has been shown that LP coincides with propositional logic on its tautologies but not on its valid inferences. In particular, although $A \wedge \neg A \rightarrow B$ is a tautology in LP, the corresponding inferences $\{A \wedge \neg A\} \models_{LP} B$ and also $\{A, \neg A\} \models_{LP} B$ are invalid. For a counterexample, consider an evaluation such that B is strictly false and $A, \neg A$ are undetermined (both true and false), so $A \wedge \neg A$ is undetermined as well. Then $0 \in v(A \wedge \neg A)$, so by the valuation postulates $1 \in v(A \wedge \neg A \rightarrow B)$. Therefore, $\{A \wedge \neg A\} \not\models_{LP} B$ and also $\{A, \neg A\} \not\models_{LP} B$. Therefore, the Logic of Paradox is a paraconsistent logic.

It turns out that the postulate of indirect consistency is not satisfied in case the strict rules of $ASPIC^+$ are instantiated with all valid inferences in the Logic of Paradox, that is, if $S \rightarrow \varphi \in \mathcal{R}_s$ iff S is finite and $S \models_{LP} \varphi$. The following counterexample was brought to our attention by Graham Priest (personal communication).

Example 5.1. Take a SAF defined by an argumentation theory with the knowledge base $\mathcal{K}_p \cup \mathcal{K}_n$, with $\mathcal{K}_n = \emptyset$ and $\mathcal{K}_p = \{a, \neg a \vee b, \neg a \vee c, \neg b \vee \neg c\}$. Further suppose that $\mathcal{R}_s$

corresponds to the Logic of Paradox and that there are no defeasible rules ($\mathcal{R}_d = \emptyset$). Finally, assume that all arguments with at least one ordinary premise are equally preferred according to the argument ordering $\preceq$ of SAF.

It is easily checked that $\mathcal{K}_p$ implies that at least one of a, b or c must be paradoxical. Therefore, there exists an argument $A_1 : a, \neg a \vee b, \neg a \vee c, \neg b \vee \neg c \rightarrow (a \wedge \neg a) \vee (b \wedge \neg b) \vee (c \wedge \neg c)$. Since tautologies are preserved in the Logic of Paradox, $\neg(b \wedge \neg b)$, $\neg(a \wedge \neg a)$ and $\neg(c \wedge \neg c)$ are also entailed by $\mathcal{K}$. This implies that there exists an argument $A_2 : a, \neg a \vee b, \neg a \vee c, \neg b \vee \neg c \rightarrow \neg((a \wedge \neg a) \vee (b \wedge \neg b) \vee (c \wedge \neg c))$. These arguments only use strict rules so they can only be attacked on their premises. However, there does not exist an argument built from $\mathcal{K}$ which has a conclusion $\neg d$ for a $d \in \mathcal{K}_p$. To show this, for each $d \in \mathcal{K}_p$ a model has to be found for which $v(d) \neq \{0\}$ holds while not $1 \in v(\neg d)$.

Model 1: to show that it is not the case that $\neg a$ follows from $\mathcal{K}_p$.
Take the model $v(a) = \{1\}, v(b) = \{1\}$ and $v(c) = \{0, 1\}$. Then it is clear that $v(a) = \{1\}, v(\neg a \vee b) = \{1\}, v(\neg a \vee c) = \{0, 1\}$ and $v(\neg b \vee \neg c) = \{0, 1\}$, but not $1 \in v(\neg a)$.

Model 2: to show that it is not the case that $\neg(\neg a \vee b)$ follows from $\mathcal{K}_p$.
Take again the model $v(a) = \{1\}, v(b) = \{1\}$ and $v(c) = \{0, 1\}$. Then it is clear that $v(a) = \{1\}, v(\neg a \vee b) = \{1\}, v(\neg a \vee c) = \{0, 1\}$ and $v(\neg b \vee \neg c) = \{0, 1\}$, but not $1 \in v(\neg(\neg a \vee b))$.

Model 3: to show that it is not the case that $\neg(\neg a \vee c)$ follows from $\mathcal{K}_p$.
Take the model $v(a) = \{1\}, v(b) = \{0, 1\}$ and $v(c) = \{1\}$. Then it is clear that $v(a) = \{1\}, v(\neg a \vee b) = \{0, 1\}, v(\neg a \vee c) = \{1\}$ and $v(\neg b \vee \neg c) = \{0, 1\}$, but not $1 \in v(\neg(\neg a \vee c))$.

Model 4: to show that it is not the case that $\neg(\neg b \vee \neg c)$ follows from $\mathcal{K}_p$.
Take the model $v(a) = \{0, 1\}, v(b) = \{0\}$ and $v(c) = \{0\}$. Then it is clear that $v(a) = \{0, 1\}, v(\neg a \vee b) = \{0, 1\}, v(\neg a \vee c) = \{0, 1\}$ and $v(\neg b \vee \neg c) = \{1\}$, but not $1 \in v(\neg(\neg b \vee \neg c))$.

This means that there are no arguments which defeat one of the arguments A_1 and A_2, so A_1 and A_2 are elements of a complete extension E. This means that $(a \vee \neg a) \vee (b \vee \neg b) \vee (c \vee \neg c)$ and $\neg((a \vee \neg a) \vee (b \vee \neg b) \vee (c \vee \neg c))$ are elements of $\texttt{Conc}(E)$. Therefore, this argumentation theory does not satisfy the postulate of direct consistency.

5.2 Da Costa's Basic C-system: C_ω

The system C_ω of Da Costa (1974) adds the axioms $\neg\neg A \rightarrow A$ and $A \vee \neg A$ to positive logic, a negation free first-order logic (these added axioms are called the 'Dialectic Double Negation' (DDN) and 'Exclusive Middle' (EM) respectively). C_ω is in certain aspects the dual of intuitionistic logic, since in intuitionistic logic the axiom EM is invalid and the axiom Non-Contradiction (NC, $\neg(A \wedge \neg A)$) is valid. In C_ω, the axiom EM is valid and NC is invalid. Intuitionistic logic tolerates incomplete situations to avoid inconsistency, while the C-systems admit inconsistent situations, but incomplete situations are removed. For example, in C_ω it is possible that all three sentences A, $\neg A$, $\neg\neg A$ are true. However unlike in the Logic of Paradox, sentences can only have one truth value. Next the semantics and the proof theory are given which are sound and complete with respect to each other.

The propositional version of C_ω has the following bivalent valuation for formulas built from a logical language $\mathcal{L}$.

1. $v(A \wedge B) = 1 \Leftrightarrow v(A) = 1$ and $v(B) = 1$
2. $v(A \vee B) = 1 \Leftrightarrow v(A) = 1$ or $v(B) = 1$

3. $v(A \to B) = 1 \Leftrightarrow v(A) = 0$ or $v(B) = 1$
4. $v(\neg A) = 1 \Leftarrow v(A) = 0$
5. $v(A) = 1 \Leftarrow v(\neg\neg A) = 1$

An interpretation of a formula f by its valuation form is a *model* if and only if $v(f) = 1$ in that interpretation. An interpretation is a model of a set of formulas if and only if it is a model of every formula in the set. The semantical logical consequence:

- $\Sigma \models_{C_\omega} A \;\Leftrightarrow$ for all evaluations v either $v(A) = 1$ or for some $B \in \Sigma$, $v(B) = 0$

It is easy to show that $\{A, \neg A\} \not\models_{C_\omega} B$, since the following valuation function can be chosen: $v(A),\ v(\neg A) = 1$ while $v(B) = 0$.

Replacing LP with C_ω as the source of strict rules in $ASPIC^+$ still yields counterexamples to direct consistency. (In this example, $\supset$ denotes the material implication).

Example 5.2. Suppose that the knowledge base is $\mathcal{K} = \mathcal{K}_p \cup \mathcal{K}_n$, with $\mathcal{K}_n = \emptyset$ and $\mathcal{K}_p = \{a, a \supset b, a \supset \neg b\}$ with the following valuation: $v(a) = 1$, $v(a \supset b) = 1$ and $v(a \supset \neg b) = 1$. Further suppose that $\mathcal{R}_s$ corresponds to C_ω and that there are no defeasible rules ($\mathcal{R}_d = \emptyset$). Finally, assume that all basic fallible arguments are equally strong. Then the following two arguments exist: $A_1 : a, a \supset b \to b$ and $A_2 : a, a \supset \neg b \to \neg b$. These are shown in Figure 4. These two arguments both use a strict rule. This means that these arguments can only be defeated on their premises. However, none of $\neg a$, $\neg(a \supset b)$, $\neg(a \supset \neg b)$ are in $Cl_{\mathcal{R}_s}(\mathcal{K})$, so there are no arguments which defeat A_1 or A_2 on their premises. Therefore, A_1 and A_2 will be elements of a complete extension E, which means that $b, \neg b \in \texttt{Conc}(E)$.

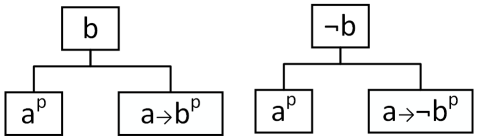

Figure 4: Arguments of Example 5.2

6. Another Attempt: Instantiating *ASPIC*$^+$ with Weak Consequence

In this section we investigate the use of another paraconsistent consequence notion, the so-called weak consequence relation originally proposed by Rescher and Manor (1970). Its basic idea is that a sentence weakly follows from a set S of sentences if it classically follows from at least one consistent subset of S. This notion is clearly related to classical argumentation, as we will formally show in Section 9. It also inspired by early consistency-based approaches to argumentation, such as Krause, Ambler, Elvang-Gøransson, and Fox (1995). To our knowledge, we are the first to use it in a system with defeasible rules. We first discuss the weak-consequence notion and then define how it can be used to overcome the trivialisation problem in $ASPIC^+$.

6.1 Weak Consequence

Weak consequence over a standard propositional language is formally defined as follows.

Definition 6.1 (**Weak consequence relation,** $\vdash_W$)**.** $\Gamma \vdash_W \alpha$ if and only if there is a maximal consistent subset Δ of Γ such that $\Delta \vdash \alpha$ in classical logic.

Note that the word 'maximal' is in fact not required, since according to Lindenbaum's Lemma every consistent set of formulas can be extended into a maximally consistent one.

It is easy to see that $\{a, \neg a\} \vdash_W b$ does not hold, because $\{a, \neg a\}$ is not a maximal consistent subset of $\{a, \neg a\}$. Therefore, this consequence relation is paraconsistent.

We next discuss three other common properties.

[**Reflexivity**] If $\alpha \in \Gamma$, then $\Gamma \vdash_W \alpha$.

This property holds for $\vdash_W$. Each $\alpha \in \Gamma$ belongs to some maximally consistent subset Δ of Γ. In classical logic, it holds that if $\alpha \in \Delta$, then $\Delta \vdash \alpha$. Therefore, it obviously holds that $\Gamma \vdash_W \alpha$.

[**Monotonicity**] $\Gamma \vdash_W \alpha$, then $\Gamma, \Pi \vdash_W \alpha$.

The monotonicity property can be proven for $\vdash_W$ as follows. There must be a maximal consistent subset Δ of Γ such that $\Delta \vdash \alpha$. Since $\Delta \subseteq \Gamma \cup \Pi$, there must exist a maximal consistent extension of Δ in $\Gamma \cup \Pi$, Δ', such that $\Delta' \vdash \alpha$. Therefore, $\Gamma, \Pi \vdash_W \alpha$.

[**Cut**] $\Gamma, \alpha \vdash_W \beta$ and $\Gamma \vdash_W \alpha$, then $\Gamma \vdash_W \beta$.

This rule does not hold. For a counterexample, consider the set $\Gamma = \{a, \neg a \wedge b\}$. Then $\Gamma \vdash_W b$ and $\Gamma, b \vdash_W a \wedge b$, while it is not the case that $\Gamma \vdash_W a \wedge b$.

Since the **Cut** rule does not hold, a naive instantiation of *ASPIC*$^+$'s strict rules with this logic W would not avoid explosion, as shown in the following example:

Example 6.1. Consider the following knowledge base $\mathcal{K}_p = \{p, \neg p, r\}$, $\mathcal{K}_n = \emptyset$, instantiate the strict rules with all valid inferences from finite sets in the logic W and let $\mathcal{R}_d = \emptyset$. Then the following arguments can be constructed:

$A_1 : p$ $\quad$ $A_2 : A_1 \rightarrow p \vee \neg r$
$B : \neg p$ $\quad$ $C : A_2, B \rightarrow \neg r$
$D : r$

Argument C concludes with $\neg r$.

The underlying reason for this problem is that the **Cut** rule does not hold for $\vdash_W$, so that in our example $\mathcal{K}_p \nvdash_W \neg r$. So if we want *ASPIC*$^+$'s strict part to behave according to $\vdash_W$, chaining of strict rules should be excluded.[3] In Example 6.1, the argument C is not allowed

3. A similar idea was suggested by Martin Caminada in personal communication. We will discuss his idea in Section 10.

since A_2 already has a strict top rule. The prohibition of the chaining of strict rules will prevent trivialisation. To this end, the $ASPIC^+$ notion of an argument must be redefined, which results in the $ASPIC^\star$ framework.

6.2 The $ASPIC^\star$ Framework

We now change the $ASPIC^+$ framework into the $ASPIC^\star$ framework by disallowing the chaining of strict rules in arguments. We first need to change the definition of an argument:

Definition 6.2. [**Argument* in $ASPIC^\star$**] An *argument** A on the basis of a knowledge base $\mathcal{K} = (\mathcal{K}, \preceq)$ in an argumentation system $(\mathcal{L}, \mathcal{R}, n, \preceq')$ is:

1. φ if $\varphi \in \mathcal{K}$ with
   $\texttt{Prem}(\varphi) = \{\varphi\}$,
   $\texttt{Conc}(\varphi) = \varphi$,
   $\texttt{Sub}(A) = \{\varphi\}$,
   $\texttt{TopRule}(A) =$ undefined.

2. $A_1, \ldots, A_n \to \psi$ if $A_1, \ldots, A_n$ are arguments* with a defeasible top rule or are from $\mathcal{K}$ and such that there exists a strict rule $\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \to \psi$ in $\mathcal{R}_s$.
   $\texttt{Prem}(A) = \texttt{Prem}(A_1) \cup \ldots \cup \texttt{Prem}(A_n)$,
   $\texttt{Conc}(A) = \psi$,
   $\texttt{Sub}(A) = \texttt{Sub}(A_1) \cup \ldots \cup \texttt{Sub}(A_n) \cup \{A\}$,
   $\texttt{TopRule}(A) = \texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \to \psi$.

3. $A_1, \ldots, A_n \Rightarrow \psi$ if $A_1, \ldots, A_n$ are arguments* such that there exists a defeasible rule $\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \Rightarrow \psi$ in $\mathcal{R}_d$.
   $\texttt{Prem}(A) = \texttt{Prem}(A_1) \cup \ldots \cup \texttt{Prem}(A_n)$,
   $\texttt{Conc}(A) = \psi$,
   $\texttt{Sub}(A) = \texttt{Sub}(A_1) \cup \ldots \cup \texttt{Sub}(A_n) \cup \{A\}$,
   $\texttt{TopRule}(A) = \texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \Rightarrow \psi$.

Arguments* are just a special case of 'normal' arguments. Therefore, all definitions for (sets of) arguments are the same in case the term argument can be replaced by argument* without problems. *Attack* and *defeat* are just the same for arguments*. *Structured* and *abstract argumentation frameworks* for the $ASPIC^\star$ framework are just the same except that they only contain arguments*. Accordingly, the notions of justified and defensible arguments* and conclusions are still defined as in Section 2.

The new $ASPIC^\star$ framework motivates a new interpretation of the strict-closure postulate in case $\mathcal{R}_s$ corresponds to a logic L. The fact that the weak-consequence notion $\vdash_W$ does not satisfy the **Cut** rule shows that closure under $\mathcal{R}_s$ does not in general coincide with closure under the consequence notion of the logic to which $\mathcal{R}_s$ corresponds. In fact, Example 6.1 can be easily extended to a counterexample for $\vdash_W$, since $\neg r$ is in the strict closure of $\{p, \neg p\}$ while $\{p, \neg p\} \nvdash_W \neg r$. This in turn gives a reason to doubt whether full closure under strict rules is always desirable. Arguably this desirability depends on the properties of the logic to which $\mathcal{R}_s$ corresponds: if this logic does not satisfy the **Cut** rule, then full strict closure is not desirable. Instead, what is desirable in such cases is that extensions are closed under consequence in the adopted logic for $\mathcal{R}_s$. This is what we will

prove for $ASPIC^\star$ when its strict rules correspond to $\vdash_W$. To this end we will below restrict the strict-closure postulate to 'allowed' applications of strict rules and we will, moreover, introduce a new rationality postulate of *logical closure*.

We think that this analysis is compatible with Caminada and Amgoud's (2007) reason for proposing the strict-closure postulate, since arguably their implicit assumption behind this postulate was that formulas in the strict closure are always 'reachable' in that arguments for them can be constructed from the arguments in the extension:

> The idea of closure is that the answer of an argumentation-engine should be closed under strict rules. That is, if we provide the engine with a strict rule $a \rightarrow b$ (...), together with various other rules, and our inference engine outputs a as justified conclusion, then it should also output b as justified conclusion. *Consequently, b should also be supported by an acceptable argument* (emphasis added by current authors). (Caminada & Amgoud, 2007, p. 294).

This quote is compatible with our new account of strict closure and our new logical-closure postulate, since if $\mathcal{R}_s$ is based on $\vdash_W$, the implicit assumption that 'being in the strict closure of an extension' equates 'supportable by an acceptable argument' is not satisfied.

To formalise our new interpretation of the strict-closure postulate, new notions of strict closure and indirect consistency are needed. We first explain some notation. Recall from Section 2 that for any set S of arguments, $FA(S)$ denotes the set of basic fallible arguments in S and $NP(S)$ denotes the set of all necessary premises of any argument in S. The corresponding notions for $ASPIC^\star$ will be denoted by $FA^\star(S)$ and $NP^\star(S)$. Next, for any set S of arguments* let $S^\#$ be defined as $FA^\star(S) \cup NP^\star(S)$. The set $S^\#$ contains those arguments* in S that have no strict top rule, so they are the arguments* in S to the conclusions of which a strict rule can be applied to form new arguments*. Then strict closure and indirect consistency can be redefined as follows.

Definition 6.3. [**Closure***] For any $X \subseteq \mathcal{L}$, let the *closure** of X under strict rules, denoted $Cl^\star_{\mathcal{R}_s}(X)$, be the smallest set containing X and the consequent of any strict rule in $\mathcal{R}_s$ whose antecedents are in X. The set of arguments* S is said to be *closed** *under strict rules* if and only if $\texttt{Conc}(S) = Cl^\star_{\mathcal{R}_s}(\texttt{Conc}(S^\#))$.

The new strict closure notion amounts to one-steps application of strict rules.

Definition 6.3 is clarified in the following example.

Example 6.2. Suppose we have the following sets: $\mathcal{K}_n = \{p\}$, $\mathcal{K}_p = \{q, t\}$ and $\mathcal{R}_d = \{q \Rightarrow r, t \Rightarrow \neg s\}$ and $\mathcal{R}_S$ corresponds to classical propositional consequence. Then the following set S of arguments* can be constructed.

$A_1 =$	p	$A_5 =$	$A_3 \Rightarrow \neg s$
$A_2 =$	q	$A_6 =$	$A_1, A_4 \rightarrow p \wedge r$
$A_3 =$	t	$A_7 =$	$A_5 \rightarrow \neg s \vee v$
$A_4 =$	$A_2 \Rightarrow r$		

In this example $S^\# = \{A_1, \ldots, A_5\}$ because it is allowed to apply strict rules to these arguments*. S is not closed* under strict rules, since for example $p \wedge q \notin \texttt{Conc}(S)$ while argument* $A_8 = A_1, A_2 \rightarrow p \wedge q$ can be constructed, so $p \wedge q \in Cl^\star_{\mathcal{R}_s}(\texttt{Conc}(S^\#))$.

Definition 6.4. [**Indirect consistency***] A set $X \subseteq \mathcal{L}$ is *indirectly consistent** if there is not a $\varphi \in \mathcal{L}$ such that $\varphi, \neg\varphi \in Cl^\star_{\mathcal{R}_s}(X)$. Otherwise it is *indirectly inconsistent**. A set of arguments* S is said to be *indirectly consistent** if $\texttt{Conc}(S^\#)$ is indirectly consistent*.

Henceforth by *consistent** we mean *indirectly consistent**. The rest of the *ASPIC** framework is the same as in the *ASPIC*$^+$ framework.

7. Verifying the Postulates for *ASPIC** with Weak Consequence

In this section we investigate a class of instantiations of the just-defined *ASPIC** framework in which the language $\mathcal{L}$ is a (nonempty) propositional language and the set $\mathcal{R}_s$ of strict rules corresponds to Rescher and Manor's (1970) notion of weak consequence over this language. More precisely:

- $\mathcal{R}_s = \{S \rightarrow \varphi \mid S \vdash_W \varphi \text{ and } S \text{ is finite}\}$

Below we will speak of such instantiations of *ASPIC** as '*ASPIC** SAFs with $\vdash_W$'. The theorem below states that SAFs avoid trivialisation in the sense of Definition 4.1.

Theorem 7.1. No *ASPIC** SAFs with $\vdash_W$ is defined by an argumentation theory with a trivialising argumentation system.

It remains to be investigated whether this class of instantiations of *ASPIC** satisfies Caminada and Amgoud's (2007) rationality postulates and the newly proposed postulate of logical closure. To this end, we first formally specify these postulates for *ASPIC**.[4]

Definition 7.1. [**Rationality postulates for *ASPIC****] Let $\Delta = (\mathcal{A}, \mathcal{C}, \preceq)$ be an *ASPIC** structured argumentation framework defined by an *ASPIC** *AT* with $AS = (\mathcal{L}, \mathcal{R}, n)$ and $\mathcal{K} = \mathcal{K}_n \cup \mathcal{K}_p$. Let *AF* be the abstract argumentation framework corresponding to Δ and let $T \in \{$complete, preferred, grounded, stable$\}$. Then:

- Δ satisfies the *closure under subarguments postulate* iff for all T-extensions E of *AF* it holds that if an argument* A is in E then all subarguments* of A are in E;
- Δ satisfies the *consistency postulate* iff for all T-extensions E of *AF* it holds that $\texttt{Conc}(E)$ is *consistent**;
- Δ satisfies the *strict closure postulate* iff for all T-extensions E of *AF* it holds that $\texttt{Conc}(E) = Cl^\star_{\mathcal{R}_s}(\texttt{Conc}(E^\#))$.
- If $\mathcal{R}_s$ corresponds to logic L, then Δ satisfies the *L-closure postulate* iff for all T-extensions E of *AF* and all $\varphi \in \mathcal{L}$ it holds that if $\texttt{Conc}(E) \vdash_L \varphi$ then $\varphi \in \texttt{Conc}(E)$.

Since all grounded (preferred, stable) extensions are also complete extensions, it suffices the prove the postulates for complete extensions. Now one way to prove the first three postulates is to try to adapt the proofs of Modgil and Prakken (2013) for *ASPIC*$^+$ to

4. Caminada and Amgoud (2007) also propose postulates for the intersection of extensions and their conclusion sets, but since their satisfaction directly follows from satisfaction of the postulates for individual extensions, these postulates will below be ignored.

ASPIC*. However, a problem here is that in general, if $\mathcal{R}_s$ corresponds to $\vdash_W$, then closure under transposition and contraposition do not hold in *ASPIC**. We first give a counterexample to closure under transposition.

Example 7.1. Since $\{a, b, c\} \vdash_W a \wedge b$ and because of the monotonicity of the logic W, it holds that $\{a, b, c, \neg a\} \vdash_W a \wedge b$. This means that $a, b, c, \neg a \rightarrow a \wedge b$ is in $\mathcal{R}_s$. However, there is no maximal consistent subset of $\{a, b, \neg(a \wedge b), \neg a\}$ that proves $\neg c$ in classical logic. Therefore $\{a, b, \neg(a \wedge b), \neg a\} \nvdash_W \neg c$ and so $a, b, \neg(a \wedge b), \neg a \rightarrow \neg c \notin \mathcal{R}_s$. This means that if the strict rules $\mathcal{R}_s$ in an argumentation system AS of the argumentation theory $AT = (AS, \mathcal{K})$ are instantiated with the valid inferences in the logic W, then the argumentation theory AT is not closed under transposition.

A similar counterexample can be given to closure under contraposition.

Example 7.2. Consider a knowledge base with $\mathcal{K}_n = \emptyset$ and $\mathcal{K}_p = \{a, b, c, \neg a\}$. Since $\{a, b, c, \neg a\} \vdash_W a \wedge b$ and $\{a, b, \neg(a \wedge b), \neg a\} \nvdash_W \neg c$, it follows that $a \wedge b \in Cl_{\mathcal{R}_s}(\{a, b, c, \neg a\})$, but not $\neg c \in Cl_{\mathcal{R}_s}(\{a, b, \neg(a \wedge b), \neg a\})$ (because chaining of strict rules is not allowed). Therefore, if the strict rules $\mathcal{R}_s$ in an argumentation system AS of the argumentation theory AT are instantiated with the valid inferences in the logic W, then AT is not closed under contraposition.

Therefore, the results of Modgil and Prakken (2013) cannot be directly used for our purposes. However, in a somewhat different formal setting, Dung and Thang (2014) provide weaker conditions for satisfying the rationality postulates, which are implied by but do not imply closure under transposition or contraposition. We will therefore use their results as a guidance in verifying the postulates for the just-defined class of instantiations of *ASPIC**. Dung and Thang (2014) formulate their results in terms of an adaptation of Amgoud and Besnard's (2013) abstract-logic approach to abstract argumentation with abstract attack and support relations between arguments. After defining their adaptation they apply it to what they call "rule-based systems". It turns out that for our purposes we do not need Dung and Thang's general abstract framework but that we can instead adapt their definitions for their rule-based instantiation to *ASPIC**. In doing so, we will for each of our definitions and results indicate its counterpart for Dung and Thang's rule-based systems. Our definitions below implicitly assume a given *ASPIC** structured argumentation framework.

Definition 7.1. [**Base of an argument*, (cf. Dung & Thang, 2014, Def. 6)**] Let A be an argument* and BA be a finite set of subarguments* of A. BA is a base of A if

- $\texttt{Conc}(A) \in Cl^\star_{\mathcal{R}_s}(\texttt{Conc}(BA))$;
- For each argument* C, C defeats A if and only if C defeats BA.

The following example shows the intuitive idea of a base.

Example 7.3. Take $\mathcal{K}_n = \emptyset$, $\mathcal{K}_p = \{a, b\}$, $\mathcal{R}_s = \{c \rightarrow d\}$ and $\mathcal{R}_d = \{a, b \Rightarrow c\}$. Then the following arguments* can be constructed: $A_1 : a$, $A_2 : b$, $A_3 : A_1, A_2 \Rightarrow c$ and $A_4 : A_3 \rightarrow d$. See Figure 5.

A_4 can only be attacked on its subarguments* A_1, A_2, or A_3 because of the strict top rule. Every argument* that attacks A_1 or A_2 also attacks A_3, so every argument* that

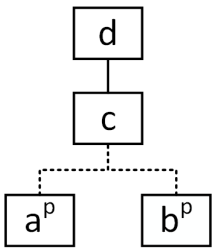

Figure 5: Arguments* of Example 7.3

attacks A_4 also attacks A_3. It is easy to see that every argument* that attacks A_3 also attacks A_4. $\texttt{Conc}(A_4) \subseteq Cl_{\mathcal{R}_s}(\texttt{Conc}(A_3))$, so $\{A_3\}$ is a base of A_4. The same kind of reasoning applies to the fact that the set $\{A_1, A_2, A_3\}$ is also a base of A_4.
However note that the set $\{A_1, A_2\}$ is not a base of A_4, because A_3 can be rebutted without A_1 or A_2 being attacked.

Definition 7.2. [**Generation of arguments*, (cf. Dung & Thang, 2014, Def. 7)**] An argument* A is said to be *generated* by a set of arguments* S, if there is a base B of A such that $B \subseteq \texttt{Sub}(S)$. The set of all arguments* generated by S is denoted by $GN(S)$.

The following lemma follows by definition of $GN(S)$.

Lemma 7.2. [**(cf. Dung & Thang, 2014, Lemma 1(2))**] For every set of arguments* S, $\texttt{Sub}(S) \subseteq GN(S)$.

Theorem 7.3. [**(cf. Dung & Thang, 2014, Thm. 1)**] Let E be a complete extension, then $GN(E) = E$.

Note that Lemma 7.2 and Theorem 7.3 immediately imply the closure* under subarguments* postulate since for every complete extension E: $\texttt{Sub}(E) \subseteq GN(E) = E$.

Theorem 7.4. Each *ASPIC** *SAF* satisfies the closure* under subarguments* postulate.

Definition 7.3. [**Compact, (cf. Dung & Thang, 2014, Def. 8)**] An *ASPIC** *SAF* is *compact* if for each set of arguments* S, $GN(S)$ is closed* under strict rules. This is equal to $\texttt{Conc}(GN(S)) = Cl^\star_{\mathcal{R}_s}(\texttt{Conc}(GN(S)^\#))$.

The following two theorems can later be combined for proving closure* under subarguments* postulate.

Theorem 7.5. [**(cf. Dung & Thang, 2014, Thm. 2)**] Each compact *ASPIC** *SAF* satisfies the closure* under strict rules postulate.

Theorem 7.6. Each *ASPIC** *SAF* with $\vdash_W$ is compact.

Definition 7.4. [**Cohesive, (cf. Dung & Thang, 2014, Def. 9)**] An *ASPIC** *SAF* is *cohesive*, if for each inconsistent* set of arguments* S, $GN(S)$ is conflicting (attacks itself).

Theorem 7.7. [**(cf. Dung & Thang, 2014, Thm. 3)**] Each cohesive *ASPIC** *SAF* satisfies the consistency* postulate.

The next two definitions are needed for proving cohesiveness.

Definition 7.5. [**Self-contradiction axiom, (cf. Dung & Thang, 2014, Def. 14)**] An *ASPIC** *SAF* is said to satisfy the *self-contradiction axiom*, if for each minimal inconsistent set $X \subseteq \mathcal{L}$: $\neg X \subseteq Cl^\star_{\mathcal{R}_s}(X)$.

Definition 7.6. [**Axiom consistent***] An *ASPIC** *SAF* is *axiom consistent** if and only if $Cl^\star_{\mathcal{R}_s}(\mathcal{K}_n)$ is consistent*.

Theorem 7.8. [**(cf. Dung & Thang, 2014, Thm. 5)**] If a compact, axiom consistent* *ASPIC** *SAF* has a reasonable argument ordering and satisfies the self-contradiction axiom, then *SAF* is cohesive.

Theorem 7.9. Each *ASPIC** *SAF* with $\vdash_W$ satisfies the self-contradiction axiom.

Combining Theorem 7.5, 7.6, 7.7, 7.8 and 7.9 results in the following important conclusion.

Theorem 7.10. Each *ASPIC** *SAF* with $\vdash_W$ which is axiom consistent* and has a reasonable argument ordering satisfies the strict-closure* and consistency* postulates.

Finally, satisfaction can be proved of the newly proposed postulate of logical closure. Below, $\vdash_{CL}$ denotes classical consequence.

Lemma 7.11. Each *ASPIC** *SAF* with $\vdash_W$ satisfies the following property: for any set S of arguments* it holds that if $\texttt{Conc}(S) \vdash_{CL} \varphi$, then $\texttt{Conc}(S^{\#}) \vdash_{CL} \varphi$.

Lemma 7.12. Each *ASPIC** *SAF* with $\vdash_W$ satisfies the following property: for any set E of arguments* it holds that if $\texttt{Conc}(E)$ is strictly closed* and consistent*, then $\texttt{Conc}(E)$ is classically consistent.

Theorem 7.13. Each *ASPIC** *SAF* with $\vdash_W$ which is axiom consistent* and has a reasonable argument ordering satisfies the logical closure postulate.

It can be concluded that we have identified a class of instantiations of the new *ASPIC** framework with Rescher and Manor's (1970) weak consequence notion that satisfies the consistency and closure postulates while preventing trivialisation in case of rebutting arguments. In order to obtain these results, the *ASPIC*$^+$ framework had to be adapted by prohibiting chaining of strict rules, resulting in the new *ASPIC** framework and new notions of strict closure and indirect consistency, plus a new postulate of closure under logical consequence.

8. Minimality of Arguments

In this section we address two aspects of minimality of arguments. We first explain the issue of inference rules with non-minimal sets of antecedents in more detail, we then investigate the effects of requiring strict rules to have minimal antecedent sets, and we finally study the issue of non-circular arguments.

8.1 The Issue of Minimality of Antecedent Sets

As said in the introduction, deductive approaches to argumentation (e.g., Besnard & Hunter, 2008; Gorogiannis & Hunter, 2011; Amgoud & Besnard, 2013) require arguments to have a subset-minimal set of premises. For example, Gorogiannis and Hunter define an argument as any (S, p) where S is a set of well-formed propositional formulas and p a well-formed propositional formula, such that:

1. S is consistent in classical propositional logic
2. S implies p in classical propositional logic
3. no proper subset of S implies p in classical propositional logic

(Hunter 2007 explores relaxations of these properties with his notion of 'approximate arguments'.) However, if arguments can apply defeasible inference rules, then the third requirement is undesirable, since defeasible rules that are based on more specific information may well be stronger. Consider the following example:

Example 8.1. Consider an argumentation system with $\mathcal{R}_s = \emptyset$, $\mathcal{R}_d = \{p \Rightarrow q; p, r \Rightarrow q; r \Rightarrow \neg q\}$ and consider a knowledge base with $\mathcal{K}_n = \{p, r\}$ and $\mathcal{K}_p = \emptyset$. The following arguments can be constructed.

A_1: p
A_2: $A_1 \Rightarrow q$
B_1: r
B_2: $A_1, B_1 \Rightarrow q$
C: $B_1 \Rightarrow \neg q$

For a real-world example of the two defeasible rules for q consider again the example from the introduction: *Observations done in ideal circumstances are usually correct* is on any reasonable account of rule strength stronger than *Observations are usually correct*. Now consider an argument ordering where arguments are compared on specificity. Then we have that $C \prec B_2$ while A_2 and C are incomparable. Then in all semantics a unique extension $\{A_1, A_2, B_1, B_2\}$ results, which is the intuitively correct outcome. However, if all arguments are required to have subset-minimal premises, then B_2 cannot be constructed and the outcome is different: the grounded extension is $\{A_1, B_1\}$ while there are two preferred and stable extensions, namely, $\{A_1, B_1, A_2\}$ and $\{A_1, B_1, C\}$.

The same analysis holds for defeasible versus strict rules. Consider a defeasible rule $p \Rightarrow q$ and a strict rule $p, r \rightarrow q$: then we clearly do not want to rule out an argument for q with premises p and r, since it could well be stronger than the defeasible argument for q with premise p. However, as noted above, this analysis does not apply to strict inference rules, since any strict inference guarantees its conclusion given its premises. So it still makes sense that strict inference rules should only be applied to a subset-minimal set of formulas. With this requirement, the number of arguments that can be generated can be significantly reduced in case this restriction is introduced, which can result in a more efficient system. The following example illustrates the problems that arise without the requirement that strict rules have subset-minimal sets of antecedents.

Example 8.2. Suppose the strict rules of an argumentation system are instantiated with classical logic and consider a knowledge base with $\mathcal{K}_n = \emptyset$ and $\mathcal{K}_p = \{p, q, r, s, t, u\}$. Then among other things, argument $A_1 : p, q \rightarrow p \wedge q$ can be constructed. Since classical logic is monotonic, all following arguments (and more) for $p \wedge q$ can also be constructed.

$A_2 : p, q, r \rightarrow p \wedge q$ $\quad$ $A_3 : p, q, s \rightarrow p \wedge q$
$A_4 : p, q, t \rightarrow p \wedge q$ $\quad$ $A_5 : p, q, u \rightarrow p \wedge q$
$A_6 : p, q, r, s \rightarrow p \wedge q$ $\quad$ $A_7 : p, q, t, u \rightarrow p \wedge q$

All arguments in the table of Example 8.2 can be considered as redundant given A_1. Recall that for defeasible rules this is different, since more specific defeasible rules for the same conclusion may well be stronger, as just explained. The problem then is to adapt the minimality requirement to the setting with defeasible rules and to investigate the extent to which this affects the conclusions that can be drawn.

8.2 Minimal Arguments* for the *ASPIC** Framework

We next investigate whether excluding strict inferences from non-subset-minimal sets of formulas makes a difference. In line with the focus in this paper we will only prove results for *ASPIC**, but the proofs for their *ASPIC*$^+$ counterparts would be entirely similar. In particular, we want to know whether the conclusions that can be drawn from an argumentation framework are affected in case arguments* are required to be minimal. We will prove that under a rather weak condition on the argument* ordering the conclusions are the same in both cases.

First, the above described idea of *minimal arguments* is formally defined. The main difference with Definition 6.2 is that clause (2) now disallows application of strict rules with a non-minimal antecedent set.

Definition 8.1. [**Minimal argument***] A *minimal argument** A on the basis of a knowledge base $\mathcal{K} = (\mathcal{K}, \preceq)$ in an argumentation system $(\mathcal{L}, \mathcal{R}, n, \preceq')$:

1. φ if $\varphi \in \mathcal{K}$ with
   $\texttt{Prem}(\varphi) = \{\varphi\}$,
   $\texttt{Conc}(\varphi) = \varphi$,
   $\texttt{Sub}(A) = \{\varphi\}$,
   $\texttt{TopRule}(A) =$ undefined.

2. $A_1, \ldots, A_n \rightarrow \psi$ if $A_1, \ldots, A_n$ are minimal arguments* with a defeasible top rule or are from $\mathcal{K}$ and such that there exists a strict rule $\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \rightarrow \psi$ in $\mathcal{R}_s$ and does not exist a strict rule $a_1, \ldots, a_i \rightarrow \psi$ for $\{a_1, \ldots, a_i\} \subset \texttt{Conc}(\{A_1, \ldots, A_n\})$ in $\mathcal{R}_s$.
   $\texttt{Prem}(A) = \texttt{Prem}(A_1) \cup \ldots \cup \texttt{Prem}(A_n)$,
   $\texttt{Conc}(A) = \psi$,
   $\texttt{Sub}(A) = \texttt{Sub}(A_1) \cup \ldots \cup \texttt{Sub}(A_n) \cup \{A\}$,
   $\texttt{TopRule}(A) = \texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \rightarrow \psi$.

3. $A_1, \ldots, A_n \Rightarrow \psi$ if $A_1, \ldots, A_n$ are arguments* such that there exists a defeasible rule $\texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \Rightarrow \psi$ in $\mathcal{R}_d$.
   $\texttt{Prem}(A) = \texttt{Prem}(A_1) \cup \ldots \cup \texttt{Prem}(A_n)$,
   $\texttt{Conc}(A) = \psi$,
   $\texttt{Sub}(A) = \texttt{Sub}(A_1) \cup \ldots \cup \texttt{Sub}(A_n) \cup \{A\}$,
   $\texttt{TopRule}(A) = \texttt{Conc}(A_1), \ldots, \texttt{Conc}(A_n) \Rightarrow \psi$.

Recall that each of the functions defined in this definition can also be defined for sets of arguments*.

It is easy to see that our minimality constraint does not exclude the construction of argument B_2 in Example 8.1, as desired.

In order to show that non-minimal arguments* are not required to obtain the right extensions, we first define the notions of a minimal and an extended version of an argument*. Informally, a minimal version of an argument* A is an argument* A^- that is the same as A except that in A^- any non-minimal strict rule in A is replaced with a minimal version of the strict rule, i.e., a strict rule with the same consequent but a subset-minimal antecedent set. This may lead to the deletion of subarguments* from A, namely, subarguments* for the deleted antecedents. Conversely, an extended version of a minimal argument* A^- is any argument that is the same as A^- except that in A zero or more strict rules in A^- are replaced with non-minimal version of these strict rules. This may lead to the addition of subarguments*, namely, subarguments* for the added antecedents.

Definition 8.2. [A^- **and** S^-] For any argument* A, an argument* A^- is a *minimal argument** of A iff:

- $A \in \mathcal{K}$ and $A^- = A$; or
- A is of the form $A_1, \ldots, A_n \rightarrow \varphi$ and $A^- = A_i^-, \ldots, A_j^- \rightarrow \varphi$ such that:
  - $\texttt{Conc}(\{A_i, \ldots, A_j\})$ is a minimal subset of $\texttt{Conc}(\{A_1, \ldots, A_n\})$ such that $\texttt{Conc}(A_i), \ldots, \texttt{Conc}(A_j) \rightarrow \varphi \in \mathcal{R}_s$; and
  - For every $k \in \{i, \ldots, j\}$ it holds that A_k^- is a minimal argument* of A_k;

  or
- A is of the form $A_1, \ldots, A_n \Rightarrow \varphi$ and $A^- = A_1^-, \ldots, A_n^- \Rightarrow \varphi$ such that for every $1 \leq i \leq n$ it holds that A_i^- is a minimal argument* of A_i.

For a set of arguments* S, define S^- as all minimal arguments* of S.

Below, when we write A^-, we mean that argument* A^- is a minimal argument* of A.

Note that A^- is not guaranteed to be unique. For example, argument* $A : p \wedge q, q \wedge p \rightarrow p$ has two minimal variants, namely $A_1 : p \wedge q \rightarrow p$ and $A_2 : q \wedge p \rightarrow p$.

Obviously, the following structured argumentation frameworks are *ASPIC** frameworks that only contain arguments*.

Definition 8.3. [**Minimal** SAF**,** SAF^-] For a $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$, let SAF^- be the *minimal* SAF with $SAF^- = (\mathcal{A}^-, \mathcal{C}^-, \preceq^-)$. Where $\mathcal{C}^-$ is defined as $\mathcal{C} \cap (\mathcal{A}^- \times \mathcal{A}^-)$ and $\preceq^- = \preceq \cap (\mathcal{A}^- \times \mathcal{A}^-)$.

Definition 8.4. [**Extended argument*,** A^+] For any argument* A, an argument* A^+ is an *extended argument** of A iff:

- $A \in \mathcal{K}$ and $A^+ = A$; or
- $A = A_1, \ldots, A_n \rightarrow \varphi$, and $A^+ = A'_1, \ldots, A'_m \rightarrow \varphi$ such that $m \geq n$ and for every A'_i $(1 \leq i \leq n)$ it holds that A'_i is an extended argument* A_i^+ of A_i; or
- $A = A_1, \ldots, A_n \Rightarrow \varphi$, and $A^+ = A'_1, \ldots, A'_n \Rightarrow \varphi$ such that for every A'_i $(1 \leq i \leq n)$ it holds that A'_i is an extended argument* A_i^+ of A_i.

Below, whenever we write A^+, we mean that argument* A^+ is an extended argument* of A.

Note that A is also an A^+ and that A is an A^{-+}. In general, A^+ is not unique.
The following example clarifies the definitions given above.

Example 8.3. Consider a SAF with the following arguments:

$A_1 : p \quad A_2 : p \rightarrow q$
$A_3 : r \quad A_4 : p, r \rightarrow q$

Then A_1, A_2 and A_3 are minimal arguments*, so SAF^- contains these three arguments*. A_2 is the minimal argument* corresponding to A_4, so $A_2 = A_4^-$. Now, it is also easy to see that indeed A_4 is an A_2^+. Furthermore A_4 is one example of an A_2^+.

For some results that follow, it is needed that for any argument* A, A^+ cannot be strictly preferred in the argument ordering over A and that A cannot be strictly preferred over A^-. This is not implied by the current definition of a reasonable argument* ordering. This is illustrated with an example.

Example 8.4. Consider again Example 8.3, assume that p and r are in $\mathcal{K}_n$ and assume the following argument* ordering (where $x \approx y$ means as usual that $x \preceq y$ and $y \preceq x$): $A_1 \approx A_2$; $A_3 \approx A_4$, $A_2 \prec A_4$. This argument* ordering satisfies all properties of a reasonable argument* ordering but is nevertheless counterintuitive, since the only difference between A_2 and A_4 is that A_4 contains a non-minimal version of a strict rule applied in A_2: since strict rules guarantee their consequent given their antecedent, A_4 should intuitively not be strictly preferred over A_2.

Therefore, we introduce the definition of a *tolerable argument* ordering.*

Definition 8.5. [**Tolerable argument* ordering**] $\preceq$ is a *tolerable argument* ordering* if and only if:

- For every A^+ of A, $A^+ \preceq A$;
- For any A^- of A, $A \preceq A^-$.

Intuitively, an argument* ordering is tolerable if replacing a minimal strict rule with one of its non-minimal versions cannot make an argument* stronger and if replacing a non-minimal strict rule with its minimal version cannot make an argument* weaker. For strict inference rules, which are meant to capture deductively valid inferences, this is a very natural property, since these operations always amount to adding, respectively, deleting attackable subarguments*.

The next lemmas are all needed for proving the equivalence of the conclusions that are drawn from minimal and non-minimal structured argumentation frameworks.

Lemma 8.1. For any argument* A and any extended argument* A^+ the following holds: for any $A' \in \texttt{Sub}(A)$ there is an argument* $A'' \in \texttt{Sub}(A^+)$ such that $A'' = A'^+$.

This lemma is clarified with the example below.

Example 8.5. Take the following arguments*:

$$A_1 : p \to p \vee q \qquad A_2 : A_1 \Rightarrow s$$
$$B_1 : p, r \to p \vee q \qquad B_2 : B_1 \Rightarrow s$$

Then it is obvious that B_2 is an A_2^+. Lemma 8.1 states that for every subargument* A' of A_2 there is a subargument* B' of A_2^+ such that $B' = A'^+$. For example, take A' to be A_1. Then B_1 is the subargument* of B_2 such that $B_1 = A_1^+$.

The preceding result is needed for proving that if argument* A attacks/defeats B, then the minimal argument* corresponding to A attacks/defeats every extended version of B.

Lemma 8.2. If $\preceq$ is a tolerable argument* ordering and argument* A defeats/attacks B, then every A^- defeats/attacks every B^+.

The following lemma follows from Lemma 8.2.

Lemma 8.3. If $\preceq$ is a tolerable argument* ordering, then for all complete extensions E:

1. If $A \in E$, then $A^- \in E$ for every A^-;
2. If $B \notin E$, then $B^+ \notin E$ for all B^+.

The following lemma states that Dung's characteristic function (see Definition 2.1) is a monotonic bijection from all complete extensions of a SAF^- onto all complete extensions of SAF. This lemma is based on the results of Dung, Toni, and Mancarella (2010) for the ABA framework, which will be discussed below. First, the following definition is needed.

Definition 8.6. [S^-] For a set of arguments S, let S^- be the set of arguments in S such that each argument is minimal.

Lemma 8.4. Let $SAF^- = (\mathcal{A}^-, \mathcal{C}^-, \preceq^-)$ be the minimal structured argumentation framework corresponding to $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ (for the *ASPIC** framework), and let AF be the abstract argumentation framework corresponding to SAF. Let $\preceq$ be a tolerable argument* ordering. Also, let C and C^- be the sets of complete extensions of SAF and SAF^- respectively and let F_{AF} be Dung's characteristic function of AF. Then:

1. For each $E \in C^- : F_{AF}(E)^- = E$.

2. For each $E \in C : F_{AF}(E^-) = E$ and $E^- \in C^-$.

Clause (1) says that the set of acceptable minimal arguments* w.r.t. an extension of a minimal SAF does not change if also non-minimal arguments* are considered. Clause (2) says that the set of acceptable arguments* w.r.t. an extension of a non-minimal SAF does not change if only minimal arguments* are considered.

Below, the main result is stated for the conclusions that can be drawn from SAFs and SAF^-s. This theorem is also based on the results of Dung et al. (2010) for the ABA framework.

Theorem 8.5. Let $SAF^- = (\mathcal{A}^-, \mathcal{C}^-, \preceq^-)$ be the minimal structured argumentation framework corresponding to $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ (for the $ASPIC^\star$ framework). Let $\preceq$ be a tolerable argument* ordering. Take $T \in \{$complete, grounded, preferred, stable$\}$ and F as defined in Definition 2.1, then:

1. Let E be a T extension in SAF, then E^- is a T extension in SAF^-.

2. Let E be a T extension in SAF^-, then $F(E)$ is a T extension in SAF.

By combining all these results it can be concluded that the conclusions that can be drawn from an $ASPIC^\star$ structured argumentation framework are not affected in case arguments* are required to be minimal.

Our results generalise those of Modgil and Prakken (2013), who prove the same result for the special case of arguments with minimal sets of premises (their Proposition 28). Other related work is Dung et al. (2010), which study minimal arguments for assumption-based argumentation (ABA) as reconstructed by Dung et al. (2007) in terms of Dung's (1995) abstract argumentation frameworks. As mentioned above, this version of ABA can be reconstructed as a special case of the $ASPIC^+$ framework without preferences and defeasible rules. In fact, Dung et al. (2010) define the notion of a *non-redundant argument*, which is more general than the notion of a *minimal argument* defined in this paper. A non-redundant argument is in turn defined in terms of a *less redundant relation*. Then Dung et al. (2007) prove similar results for ABA as our Theorem 8.5 for $ASPIC^\star$. Our above constructions and proofs are clearly inspired by the work of Dung et al. (2010). However, a purely formal link with $ASPIC^\star$ cannot be established, since unlike ABA, $ASPIC^\star$ does not allow chaining of strict rules. For this reason we will not make a detailed formal comparison here.

8.3 Disallowing Repetition of Conclusions

We now address a second aspect of minimality, by studying a modification of $ASPIC^\star$ in which circular arguments* are avoided, that is, in which an argument* cannot have the same conclusion as one of its subarguments*. This requirement is not part of $ASPIC^+$ as defined by Prakken (2010) and Modgil and Prakken (2013) but it is part of the system of Vreeswijk (1997), from which the $ASPIC$-style definition of an argument originates. In argumentation theory, circular arguments are generally regarded as fallacious, so it seems a good idea to exclude them. In addition, this has computational benefits, as will be shown in this section. First Definition 6.2 of an argument* is modified as follows.

Definition 8.7. [**Strongly minimal arguments* in *ASPIC****] Strongly minimal *arguments** are defined as *arguments** in Definition 8.1 except that the following condition is added to clauses (2) and (3):

$$\psi \notin \texttt{Conc}\{(A_1, \ldots, A_n)\}.$$

Below we call a structured argumentation framework *strongly minimal* if its set of arguments* is defined with Definition 8.7.

Disallowing repetition of conclusions can in general change the extensions, as the following example shows.

Example 8.6. Consider an argumentation theory with $\mathcal{K}_n = \mathcal{K}_p = \emptyset$ and $\mathcal{R}_d = \{\Rightarrow p, \Rightarrow \neg p, p \Rightarrow p\}$. Then there are at least two arguments* for p and at least one for $\neg p$:

A_1: $\Rightarrow p$
A_2: $A_1 \Rightarrow p$
B: $\Rightarrow \neg p$

If we have an ordering on rules $\mathcal{R}_d$ such that $\Rightarrow p <\Rightarrow \neg p < p \Rightarrow p$ (where $x < y$ means that y is strictly preferred over x), then with the last-link ordering as defined by Modgil and Prakken (2013), B defeats A_1 while A_2 defeats B, which yields a grounded extension that contains neither of these three arguments*. However, if A_2 cannot be constructed, then the grounded extension contains B.

As just noted, excluding circular arguments* not only avoids fallacies but also has computational benefits. In particular, it can be shown that if $\mathcal{K}$ and $\mathcal{R}_d$ are finite, then each argument* has at most a finite number of attackers. In the words of Dung (1995) this means that the induced abstract argumentation framework is 'finitary'. As shown by Dung, finitary AFs have a number of computational benefits. Among other things, the grounded extension can be constructed by iterative application of the F operator (see Definition 2.1 above) on the empty set.

To prove this result, we first introduce some notation relative to a given AT.

- For any r which is defined as $S \to \varphi \in \mathcal{R}_s$ or as $S \Rightarrow \varphi \in \mathcal{R}_d$ let $\texttt{Cons}(r) = \varphi$.
- For any set $T \subseteq \mathcal{R}$ let $\texttt{Cons}(T) = \{\varphi \mid \varphi = \texttt{Cons}(r) \text{ for some } r \in T\}$.
- Let $\mathcal{X} = \{S \to \varphi \in \mathcal{R}_s \mid S \subseteq \mathcal{K} \cup \texttt{Cons}(\mathcal{R}_d)\}$. Informally, $\mathcal{X}$ is the set of all strict rules that are potentially applicable in the AT, that is, of which the antecedents all belong to $\mathcal{K}$ or are a consequent of a defeasible rule. Note that this set does not have to be equal to $\mathcal{R}_s$, since it might be that all rules in $\mathcal{R}_s$ apply to some set of well-formed formulas from $\mathcal{L}$ while yet $\varphi \notin \mathcal{K}$and there is no defeasible rule with consequent φ.
- For any $Y \subseteq \mathcal{L}$ define $\mathcal{X}^Y = \{S \to \varphi \in \mathcal{X} \mid \varphi \in Y\}$. Informally, $\mathcal{X}^Y$ is the set of all potentially applicable rules for a formula in Y.

We next prove the following lemma.

Lemma 8.6. For any finite $Y \subseteq \mathcal{L}$ it holds that $\mathcal{X}^Y$ is finite.

Then the following result can be proven.

Theorem 8.7. Let $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ be a strongly minimal structured argumentation framework corresponding to an argumentation theory with a finite $\mathcal{K}$ and $\mathcal{R}_d$. Then for any $\varphi \in \mathcal{L}$ the set $\{A \in \mathcal{A} \mid \texttt{Conc}(A) = \varphi\}$ is finite.

Note that this result cannot be proved without the exclusion of arguments that chain strict rules, since otherwise infinite sets of arguments for the same conclusion p can be generated by constructing arguments as follows for any n, which provides a counterexample to Theorem 8.7 in case such arguments are not excluded:

$A_1 = p$
$A_2 = A_1 \rightarrow \neg\neg p$
$A_3 = A_2 \rightarrow \neg\neg\neg\neg p$
$\ldots$
$A_n = A_{n-1} \rightarrow p$

It remains to verify the rationality postulates for strongly minimal SAFs. It turns out that the only result of Section 7 that needs to be reproved is the only-if half of Theorem 7.6.

Theorem 8.8. Each strongly minimal *ASPIC** *SAF* with $\vdash_W$ is compact.

Combined with the other results of Section 7 this implies that

Theorem 8.9. Each strongly minimal *ASPIC** *SAF* with $\vdash_W$ which is axiom consistent* and has a reasonable argument* ordering satisfies the closure* and consistency* postulates.

9. *ASPIC** as a Generalisation of Classical Argumentation

In this section we explain that by combining the two main contributions so far, a class of instantiations of *ASPIC** is obtained that is a proper generalisation in three respects of two versions of classical argumentation defined by Besnard and Hunter (2008) and Gorogiannis and Hunter (2011)[5]. The two versions for which this holds are of particular interest since they are by Gorogiannis and Hunter proven to be the only two of seven versions of classical argumentation that satisfy the consistency postulates. The observation can be explained as follows. Modgil and Prakken (2013) prove that classical argumentation with two forms of premise attack called *direct undercut* and *direct defeat* can be reconstructed as the following class of instantiations of *ASPIC*$^+$: no defeasible rules, no preference relations between arguments, only ordinary premises, $\mathcal{L}$ has a negation symbol as defined above in Definition 2.3, all arguments have indirectly consistent premise sets, and the strict rules are instantiated with all classically valid inferences from finite sets of premises. Modgil and Prakken (2013) also prove that requiring arguments to have subset-minimal premises implying their conclusion does not change this result. Then it should be shown that for such classical-logic instantiations the prohibition in *ASPIC** to chain strict rules does not make a difference with *ASPIC*$^+$ either. If we can show this, then we have shown that instantiations of *ASPIC*$^+$ in which $\mathcal{R}_s$ corresponds to Rescher and Manor's (1970) notion

5. We thank Sanjay Modgil for suggesting this to us in personal communication

of weak consequence and in which all arguments* are minimal are proper generalisations of classical argumentation. We will actually prove this for the case with nontrivial preferences but where the preferences are fully determined by the premises; then the result for the case with an empty preference relation follows as a special case.

Below with a minimal $ASPIC^+$ or $ASPIC^\star$ argumentation theory we mean an argumentation theory where all arguments or all arguments* have a subset-minimal set of premises. Clearly, for arguments* this means that they are also minimal in the sense of Definition 8.1. Moreover, a c-structured argumentation framework is a notion from Modgil and Prakken (2013): restricted to the present context it amounts to the requirement on SAFs that all arguments have classically consistent premises. The notion of a minimal c-structured argumentation framework is defined accordingly.

Theorem 9.1. Let AS be an argumentation system with $\mathcal{L}$ a classical-logic language and $\mathcal{R}_s$ corresponding to classical logic, and let $\mathcal{K} = \mathcal{K}_p$ be a knowledge base in $\mathcal{L}$. Let AS^* be obtained from AS by removing from R_s all inference rules that are invalid according to $\vdash_W$. Let $AT = (AS, \mathcal{K})$ be a minimal $ASPIC^+$ argumentation theory and $AT^* = (AS^*, \mathcal{K})$ the corresponding minimal $ASPIC^\star$ argumentation theory. Then let $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ and $SAF^* = (\mathcal{A}^*, \mathcal{C}^*, \preceq^*)$ be, respectively, the minimal c-structured argumentation framework defined by AT in $ASPIC^+$ and the minimal structured argumentation framework defined by AT^* in $ASPIC^\star$ such that:

- For all $A, B, C \in \mathcal{A}$: if $\texttt{Prem}(A) = \texttt{Prem}(B)$ then ($C \preceq A$ iff $C \preceq B$ and $A \preceq C$ iff $B \preceq C$). Likewise for $\preceq^*$.
- For all A^*, $B^* \in \mathcal{A}^*$: $A^* \preceq B^*$ iff $A^* \preceq^* B^*$.[6]

Then for $T \in$ {complete, grounded, preferred, stable} it holds that:

1. Let E be a T extension in SAF, then $E \cap \mathcal{A}^*$ is a T extension in SAF^* such that $\texttt{Conc}(E) = \texttt{Conc}(E \cap \mathcal{A}^*)$.
2. Let E be a T extension in SAF^*, then $F(E)$ is a T extension in SAF such that $\texttt{Conc}(E) = \texttt{Conc}(F(E))$.

A special case of this result is that for the case without defeasible rules, necessary premises and preferences $ASPIC^\star$ with $\vdash_W$ gives the same conclusion sets as classical argumentation. Then the first proper extension of classical argumentation is with preferences. The second proper extension is with the necessary premises $\mathcal{K}_n$ while the third proper extension is to the case with defeasible rules, by observing that in $ASPIC^\star$ any strict rule that is applied to the conclusion of at least one defeasible subargument* is applied to a subset-minimal and classically consistent set of formulas that classically implies the consequent of the strict rule.

10. Summary, Discussion and Conclusion

In this section we summarise and discuss our results and put them in the context of related work.

6. This is well-defined since by construction of SAF and SAF^* and the fact that $\vdash_W$ draws no inferences from inconsistent sets it holds that $\mathcal{A}^* \subseteq \mathcal{A}$.

10.1 Summary of the Results

In this paper we tackled several related issues concerning relevance in structured argumentation. We carried out our investigations in the context of the $ASPIC^+$ framework, which consolidates and extends one of the main AI approaches to argumentation: modelling combined reasoning with strict and defeasible inference rules. One main contribution of this paper was to solve the long-standing trivialisation problem first identified by Pollock (1994, 1995). The problem is to tame the trivialising effect of the Ex Falso principle in classical logic in a way that preserves consistency and closure properties. To solve this problem, we instantiated the strict rules of $ASPIC^+$ with Rescher and Manor's (1970) paraconsistent logic W. To make this work, we had to disallow chaining of strict rules in arguments since the logic W does not satisfy the Cut rule; this resulted in the adapted framework $ASPIC^\star$ and in a new view on the postulate of strict closure. We argued that what is important is not whether conclusion sets are closed under strict rules but whether they are closed under the consequence notion of the logic to which the strict rules correspond. Accordingly, we modified the notion of strict closure and we introduced a new rationality postulate of logical closure. We then proved new versions of the consistency and closure postulates for $ASPIC^\star$. We also investigated whether two other well-known paraconsistent logics, the system C_ω of Da Costa (1974) and the Logic of Paradox of Priest (1979, 1989), are suitable sources of strict rules. We showed that in both cases this would lead to violation of indirect consistency. In future research we want to consider other paraconsistent logics and we want to more generally study the properties that a paraconsistent logic should satisfy to be useful as a source for strict rules in $ASPIC^\star$.

The second issue studied in this paper was minimality of arguments. We first showed that under a natural assumption on the argument ordering, restricting strict-rule application to subset-minimal sets of formulas does not affect the conclusions drawn in $ASPIC^\star$. This will in many cases make the reasoning more efficient by ignoring irrelevant information. We also noted that this result can be easily adapted to $ASPIC^+$. We then disallowed circular arguments in $ASPIC^\star$ and showed that this may change the status of arguments but does not affect the satisfaction of the rationality postulates. In addition, we proved that with a finite set of defeasible rules and a finite knowledge base $ASPIC^\star$ without circular arguments has the computationally attractive property that the induced abstract argumentation frameworks are finitary in the sense of Dung (1995). This latter result cannot be adapted to $ASPIC^+$ since it crucially relies on the prohibition to chain strict rules. All these results on minimality hold independently of the choice of strict rules.

Finally, we proved that by combining the contributions of this paper a version of $ASPIC^\star$ is obtained that is a proper generalisation of two versions of classical argumentation with premise attack defined by Besnard and Hunter (2008) and Gorogiannis and Hunter (2011). These two versions are of particular interest since they are by Gorogiannis and Hunter proven to be the only two of seven versions of classical argumentation that satisfy the consistency postulates.

It should be noted that in the course of our investigations we have changed the $ASPIC^+$ framework as originally defined by Prakken (2010). In fact, we are not the first to do so. Modgil and Prakken (2013) consider four variants of the $ASPIC^+$ framework. First, they consider versions with and without the constraint that arguments have consistent premises

and then for both of these variants they define a variant in which conflict-freeness of sets of arguments is not defined relative to the defeat relation but to the attack relation. Furthermore, Caminada et al. (2014) study a variant of $ASPIC^+$ in which arguments can also be rebutted on conclusions derived by strict rules, provided that at least one subargument has an ordinary premise or applies a defeasible rule. Finally, Wu and Podlaszewski (2015) study a variant of $ASPIC^+$ in which the set of all conclusions of all subarguments of an argument has to be consistent. Thus the work on $ASPIC^+$ has resulted in a family of frameworks based on the same core ideas but making different design choices at specific points, and new members of this family may result in the future. We think that this is a healthy situation, since it amounts to a systematic investigation of the effects of different design choices within a common approach, which each may be applicable to certain kinds of problems.

10.2 Discussion of Related Work

We next discuss related work.

10.2.1 The Additional Rationality Postulates of Caminada et al. (2012)

As mentioned in the introduction, Caminada et al. (2012) formulate a new set of rationality postulates in addition to those of Caminada and Amgoud (2007), to characterise cases under which the trivialisation problem is avoided (called the postulates of non-interference and crash-resistance). Wu (2012) and Wu and Podlaszewski (2015) prove for their adaptation of $ASPIC^+$ with consistent arguments and no preferences that these new postulates are satisfied for complete semantics. They did not investigate other semantics. However, we will not attempt to prove Caminada et al.'s postulates, for two reasons. First, we want to obtain results for the case with preferences and for other semantics as well and, second, it seems to us that Caminada et al.'s postulates in fact capture a stronger intuitive notion than the one we study in this paper, so that proving satisfaction of the new postulates would be more than required for the purposes of this paper.

The intuition of Caminada et al.'s notion of trivialisation is as follows. They consider "knowledge bases", which are pairs of sets of formulas from $\mathcal{L}$ and sets of defeasible rules. Two knowledge bases are defined to be *disjoint* if they are composed from disjoint sets of atomic formulas from $\mathcal{L}$. Then a knowledge base $KB_1 = (K_1, D_1)$ is *contaminating* if for every knowledge base $KB_2 = (K_2, D_2)$ the set of extensions (under a given semantics) is the same for KB_1 as for $KB_1 \cup KB_2$ (where $(K_1, D_1) \cup (K_2, D_2) = (K_1 \cup K_2, D_1 \cup D_2)$. Then a system is said to satisfy *crash resistance* if there does not exist a contaminating knowledge base in the system.

Now consider an $ASPIC^+$ or $ASPIC^\star$ instantiation with $\mathcal{K}_p = \emptyset, \mathcal{K}_n = \{p\}$ and a single defeasible rule $\Rightarrow \neg d$, where $n(\Rightarrow \neg d) = d$. This instantiation has no stable extensions, since the argument $A = \Rightarrow \neg d$ defeats itself and is not defeated by any other argument. Then the knowledge base $(\emptyset, \{\Rightarrow \neg d\})$ is contaminating, since no syntactically disjoint knowledge base to which it is added will have extensions. More generally, this situation can arise with any knowledge base with which there are no stable extensions.

However, in this paper we were not interested in excluding such situations but only in taming the contaminating effect of the Ex Falso principle. To this end we introduced a simpler definition of trivialisation in Definition 4.1 and managed to avoid trivialisation

as thus defined even for stable semantics, since the just-given example is not a case of trivialisation in the sense of Definition 4.1. We conclude from this that Caminada et al.'s postulates capture a stronger notion than the notion of paraconsistency (the focus of our paper). In the future it would be interesting to study whether our class of instantiations of *ASPIC** satisfies Caminada et al.'s postulates, but this would ideally be preceded by a study of what exactly is captured by these postulates.

10.2.2 Wu (2012), Wu and Podlaszewski (2015)

In an alternative attempt to solve the trivialisation problem, Wu and Podlaszewski (2015) introduced the inconsistency-cleaned *ASPIC Lite* system. This system is similar to the argumentation formalism treated by Caminada and Amgoud (2007) and can be seen as a system specified in *ASPIC*$^+$ in which all arguments are equally preferred. Wu and Podlaszewski define an argument to be *consistent* if the set of conclusions of all its subarguments is directly consistent. An argumentation framework is *inconsistency-cleaned* if all inconsistent arguments are removed. Wu and Podlaszewski prove that an inconsistency-cleaned version of the *ASPIC Lite* satisfies both the original rationality postulates of Caminada and Amgoud (2007) and the new postulates of Caminada et al. (2012). So it solves the trivialisation problem while retaining known results on closure and consistency. However, Wu (2012) and Wu and Podlaszewski (2015) provide a counterexample (originally due to Leon van der Torre) to satisfaction of the consistency and strict-closure postulates in case preferences are added and the last-link argument ordering of Prakken (2010) is applied. This example was originally presented in the *ASPIC Lite* system, but here it is translated into the *ASPIC*$^+$ framework.

Example 10.1. [(Wu, 2012; Wu & Podlaszewski, 2015)] Given the knowledge base $\mathcal{K} = \emptyset$, $\mathcal{R}_d = \{\Rightarrow p; p \Rightarrow q; \Rightarrow \neg p \vee \neg q\}$ and $\mathcal{R}_s$ is instantiated with all valid inferences in classical logic. Assume that $\Rightarrow p$ has priority 1 (lowest), $\Rightarrow \neg p \vee \neg q$ has priority 2 (middle) and $p \Rightarrow q$ has priority 3 (highest). In that case, we can construct the following arguments with associated (last-link principle) preferences. Table 1 depicts some arguments that can be generated[7] and Figure 6 shows the defeat relation between these arguments.

Argument	**Preference**
$A_1 : \Rightarrow p$	(1)
$A_2 : \Rightarrow \neg p \vee \neg q$	(2)
$A_3 : A_1 \Rightarrow q$	(3)
$A_4 : A_1, A_2 \rightarrow \neg q$	(1)
$A_5 : A_1, A_3 \rightarrow \neg(\neg p \vee \neg q)$	(1)
$A_6 : A_2, A_3 \rightarrow \neg p$	(2)

Table 1: Six arguments

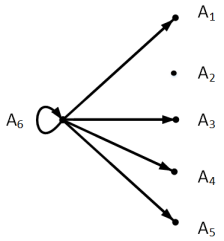

Figure 6: Partial abstract AF

7. Note that classical reasoning allows the generation of infinitely more arguments but they are all irrelevant to any of these six arguments.

Argument A_6 is an inconsistent argument. So according to the solution proposed by Wu and Podlaszewski for the case without preferences, A_6 needs to be deleted from the argumentation framework. Figure 7 shows the resulting argumentation framework.

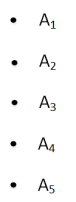

Figure 7: Inconsistency-cleaned version

There is a complete extension $E = \{A_1, A_2, A_3, A_4, A_5\}$. It does not satisfy closure under strict rules because A_2 and A_3 are in E and A_6 is not in E. Moreover, direct consistency is also not satisfied since A_3 and A_4 are both in the complete extension E, but the conclusions q and $p \wedge \neg q$ are not consistent.

Arguments A_3 and A_4 have opposite conclusions so without preferences A_4 would defeat A_3. However, with the preference ordering chosen in the counterexample, A_4 is weaker than A_3 so A_4 cannot defeat A_3. A_3 and A_4 are also not being attacked on their subarguments. The fact that these arguments are both in the same complete extension causes the problem. Every argument that concludes with $\neg p$ uses A_1 as a subargument, so this implies that it is an inconsistent argument and it has to be removed from the framework. Therefore, A_1 is not defeated by any argument, which means that A_3 and A_4 are both in a complete extension. It can be concluded that an inconsistent argument like A_6 is really needed to defeat A_3 and A_4.

Furthermore, it can be observed that if A_6 is not deleted, there are no problems at all. Figure 6 shows that in that case there is only one complete extension $E = \{A_2\}$. and it satisfies both consistency and strict closure. Therefore, in this example, it is undesirable that A_6 is removed.

Consider next the *ASPIC** framework as defined in this paper. The same arguments can be constructed as in the original framework. This way, there is only one complete extension $\{A_2\}$ and, as explained above, the rationality postulates are all satisfied. The approach with *ASPIC** is therefore more general than the solution of Wu and Podlaszewski (2015), since it applies to frameworks that include preferences and defeasible rules.

It should be noted that our idea to forbid chaining of strict rules was earlier suggested to us by Martin Caminada (personal communication). However, he combined his suggestion with the idea to disallow inconsistent arguments; in that case, the above counterexample to consistency and strict closure can still be constructed. Apart from this, one could say that the logic W, which does not satisfy the Cut rule, provides a theoretical foundation for the idea to disallow chaining of strict rules.

At first sight, it would seem that a system that allows inconsistent arguments is flawed even if it satisfies the consistency and closure postulates. However, note that such arguments will never be in any extension. Although as explained above and by Caminada (2005), inconsistent arguments can sometimes prevent other arguments from being in an extension, this is only a problem if such arguments are based on the Ex Falso principle, since that

principle holds as a matter of logic. Among other things, this means that allowing arguments based on Ex Falso would dramatically increase the number of (counter)arguments, which would lead to computational problems. By contrast, inconsistent arguments as argument A_6 in Example 10.1 arise because of specific modeling choices in $\mathcal{R}_d$ not dictated by logic and they do not proliferate, so from a logical or computational point of view there is no need to exclude them.

10.3 Dung's (2014) Rule-Based Systems

Dung (2014) introduces a formalism of 'rule-based systems' (further studied in Dung, 2016), which essentially is a notational variant of $ASPIC^+$ restricted to literal languages and empty knowledge bases (necessary or ordinary facts are represented as strict or defeasible rules with empty antecedents). Dung introduces three new rationality postulates. His postulate for *attack monotonicity* informally says that strengthening an argument cannot eliminate an attack of that argument on another. His postulate of *credulous cumulativity* informally means that changing a conclusion of an argument in some extension to a necessary fact cannot eliminate that extension. Finally, his property of *irrelevance of redundant defaults* says that adding redundant defaults should not change the set of extensions.

Dung then investigates the argument orderings studied by Modgil and Prakken (2013) on whether they satisfy these and the consistency postulates, with both positive and negative results. While these results are valuable, Dung (2014) unfortunately, somewhat confuses matters by referring to the orderings studied by Modgil and Prakken as "the $ASPIC^+$ semantics". Thus he overlooks the distinction between $ASPIC^+$ as a general framework and instantiations of the framework. The argument orderings studied by Modgil and Prakken are not inherent to the $ASPIC^+$ framework but are just some example orderings. The $ASPIC^+$ framework and its variants leave every room for other ways to define the argument ordering. It should be noted that Dung (2016) does not refer to Modgil and Prakken's orderings as "the $ASPIC^+$ semantics" any more and thus respects that these orderings are not inherent to $ASPIC^+$.

For present purposes Dung's findings are strictly speaking irrelevant since we have not studied particular argument orderings. In future research it would be interesting to investigate whether the use of particular argument orderings in $ASPIC^\star$ satisfies Dung's postulates of attack monotonicity and irrelevance of redundant defaults. However, we disagree with Dung (2014) that credulous cumulativity would be a desirable property. On the contrary, with Prakken and Vreeswijk (2002, section 4.4) we believe that this property is instead undesirable, since strengthening a defeasible conclusion to an indisputable fact may make arguments stronger than before. This can give them the power to defeat other arguments that they did not have before. This may well result in the loss of the extension from which the conclusion was promoted to an indisputable fact.

10.4 Conclusion

In this paper we have successfully addressed some open issues concerning relevance in structured argumentation. We have solved the trivialisation problems that arise when argumentation includes full classical logic and we have created the prospects for reducing computational complexity by enforcing minimality and non-circularity of arguments while ensuring

closure and consistency results. All this was done in the context of the *ASPIC* approach, resulting in a new variant of the $ASPIC^+$ framework called $ASPIC^\star$ and a well-behaved class of instantiations of the new framework. This class of instantiations was shown to be a proper generalisation of classical argumentation with preferences and defeasible rules. It is this class of instantiations and its properties that are the main contribution of this paper. In our paper we have employed a flexible attitude towards design choices within the *ASPIC* approach. We have thus shown that this approach is a fruitful one, provided the distinction between frameworks and their instantiations is kept in mind.

Acknowledgement

We thank all three JAIR reviewers for their many useful comments on earlier versions of our paper.

Appendix A. Proofs

This appendix contains the proofs of all results reported in the paper.

A.1 Proofs for Section 7

Theorem 7.1 No $ASPIC^\star$ SAFs with $\vdash_W$ is defined by an argumentation theory with a trivialising argumentation system.

Proof. To this end, we must show that for any argumentation system AS with a propositional language and with $\mathcal{R}_s$ as just defined, a knowledge base $\mathcal{K}$ in AS can be defined such that for some $\{\varphi, \neg\varphi\} \subseteq \mathcal{K}$ and for some $\psi \in \mathcal{L}$ it does not hold that an argument* with conclusion ψ can be constructed on the basis of $\mathcal{K}$ in AS. Consider any such AS. We choose $\mathcal{K} = \mathcal{K}_n \cap \mathcal{K}_p$ where $\mathcal{K}_p = \emptyset$ and $\mathcal{K}_n = \{\varphi, \neg\varphi\}$ for any formula φ from $\mathcal{L}$ (φ is guaranteed to exist since $\mathcal{L}$ is assumed nonempty). Then by definition of $\vdash_W$ it clearly holds that $\mathcal{K} \nvdash_W (\varphi \wedge \neg\varphi)$ since no consistent subset of $\mathcal{K}$ can classically imply a contradiction. So there exists no strict argument* for $(\varphi \wedge \neg\varphi)$ on the basis of $\mathcal{K}$ in AS. □

Theorem 7.3 Let E be a complete extension, then $GN(E) = E$.

Proof. First note that according to Definition 7.2 for each set S of arguments* $\texttt{Sub}(S) \subseteq GN(S)$, therefore $E \subseteq GN(E)$.
Suppose now that an argument* C defeats an argument* $A \in GN(E)$. Let BA be a base of A such that $BA \subseteq \texttt{Sub}(E)$, then C defeats BA. Hence C defeats $\texttt{Sub}(E)$ and so it defeats E. Since E is a complete extension, every defeat against E is counter defeated by E. A is defended by E, so $A \in E$. Therefore $GN(E) \subseteq E$. □

Theorem 7.5 Each compact $ASPIC^\star$ SAF satisfies the closure* under strict rules postulate.

Proof. Let E be a complete extension. The compactness* implies that $GN(E)$ is closed* under strict rules. From Theorem 7.3 E is closed* under strict rules. □

Theorem 7.6 Each $ASPIC^\star$ SAF with $\vdash_W$ is compact.

Proof.
[$\texttt{Conc}(GN(S)) \supseteq Cl^{\star}_{\mathcal{R}_s}(GN(S)^{\#})$]
Let S be a set of arguments* and $\sigma \in Cl^{\star}_{\mathcal{R}_s}(GN(S)^{\#})$. It needs to be shown that $\sigma \in \texttt{Conc}(GN(S))$. Let X be a minimal subset of $\texttt{Conc}(GN(S)^{\#})$ such that $\sigma \in Cl^{\star}_{\mathcal{R}_s}(X)$. Hence there is a strict argument* A_0 over X with conclusion σ. Further let S_X be a minimal set of arguments* from $GN(S)^{\#}$ s.t. $\texttt{Conc}(S_X) = X$. Let A be the argument* obtained by replacing each leaf in A_0 (viewed as a directed acyclic graph) labelled by a literal α from X by an argument* with conclusion α from S_X. Note that this is possible since all arguments* in S_X are basic fallible arguments* or are just necessary premises. It is obvious that the conclusion of A is σ. It is shown that S_X is a base of A. Suppose B is an argument* defeating A. Since A_0 is a strict argument* over X, B must defeat a basic fallible subargument* in S_X. Hence B defeats S_X. Thus $A \in GN(S)$. Hence $\sigma \in \texttt{Conc}(GN(S))$.
[$\texttt{Conc}(GN(S)) \subseteq Cl^{\star}_{\mathcal{R}_s}(GN(S)^{\#})$]
Suppose $\sigma \in \texttt{Conc}(GN(S))$, then it has to be shown that $\sigma \in Cl^{\star}_{\mathcal{R}_s}(GN(S)^{\#})$. $\sigma \in \texttt{Conc}(GN(S))$ means that there is an argument* $A \in GN(S)$ with $\texttt{Conc}(A) = \sigma$.
Suppose A is of the form $\Rightarrow \sigma$ or $\sigma \in \mathcal{K}_p$, then $A \in FA^{\star}(GN(S))$ and thus $A \in GN(S)^{\#}$.
Suppose A is of the form $\rightarrow \sigma$ or $\sigma \in \mathcal{K}_n$, then $\sigma \in Cl^{\star}_{R_s}(\emptyset)$ or $\sigma \in NP^{\star}(GN(S))$ respectively, so $\sigma \in Cl^{\star}_{\mathcal{R}_s}(GN(S)^{\#})$.
Suppose A is of the form $A_1, \ldots, A_n \Rightarrow \sigma$, then $A \in FA^{\star}(GN(S))$ and thus $A \in GN(S)^{\#}$.
Finally, suppose A is of the form $A_1, \ldots, A_n \rightarrow \sigma$, then since $A_1, \ldots, A_n$ are basic fallible arguments* $A_1, \ldots, A_n \in GN(S)^{\#}$. Therefore $\sigma \in Cl^{\star}_{\mathcal{R}_s}(GN(S)^{\#})$.
It can be concluded that $\sigma \in Cl^{\star}_{\mathcal{R}_s}(GN(S)^{\#})$.
It is proven that the AF is compact. □

Theorem 7.7 Each cohesive *ASPIC** *SAF* satisfies the consistency* postulate.

Proof. Let E be a complete extension. Suppose E is inconsistent*. From cohesion, it follows that $GN(E)$ is conflicting. Theorem 7.3 states that then E is conflicting. This is a contradiction since E is a complete extension, so E has to be consistent*. □

Theorem 7.8 If a compact, axiom consistent* *ASPIC** *SAF* has a reasonable argument* ordering and satisfies the self-contradiction axiom, then SAF is cohesive.

Proof. Let S be an inconsistent* set of arguments* and take a minimal inconsistent* subset S' of $\texttt{Sub}(S)$. Definition 6.4 combined with axiom consistency* and the minimality of S' causes that $S' \neq \emptyset$ and only contains basic fallible arguments* or necessary premises. Remark that S' cannot consist of only necessary premises, because of axiom consistency*. Further note that $\texttt{Conc}(S')$ is a minimal inconsistent set. Since AF satisfies the self-contradiction axiom, for all $\sigma \in \texttt{Conc}(S')$ it holds that $\neg\sigma \in Cl^{\star}_{\mathcal{R}_s}(\texttt{Conc}(S'))$. Let B be the weakest argument* of S' with $\texttt{Conc}(B) = \sigma$. Note that B cannot be a necessary premise because of the reasonable argument* ordering and the fact that S' must contain basic fallible arguments*. By construction of S' it holds that $S' \subseteq GN(S')^{\#}$. Therefore $\neg\sigma \in Cl^{\star}_{\mathcal{R}_s}(\texttt{Conc}(GN(S')^{\#}))$. Because of the compactness of AF it follows that $\neg\sigma \in \texttt{Conc}(GN(S'))$. Therefore, there is an argument* $A \in GN(S')$ such that $\texttt{Conc}(A) = \neg\sigma$. Hence A attacks B. The base of A is S', so it can be concluded that all basic fallible subarguments* of A are in S'. B is the weakest argument* of S', so because of the

reasonable argument* ordering and the fact that all basic fallible subarguments* of A are in S' implies that $A \not\prec B$. This means that A defeats B. Since $B \in S' \subseteq GN(S') \subseteq GN(S)$, $GN(S)$ is conflicting. Therefore, AF is cohesive. □

Theorem 7.9 Each *ASPIC** *SAF* with $\vdash_W$ satisfies the self-contradiction axiom.

Proof. It has to be proved that for every minimal inconsistent* set $X \subseteq \mathcal{L}$ it holds that for each $\sigma \in X$, $\neg\sigma \in Cl^\star_{\mathcal{R}_s}(X)$. Let X be a minimally inconsistent* set and take $S = X \backslash \sigma$. Note that S is a maximal consistent* subset of X and that $S, \sigma \vdash \bot$ (where $\vdash$ denotes classical entailment). By the deduction theorem for classical logic $S \vdash \sigma \supset \bot$, which implies $S \vdash \neg\sigma$. Since S is a maximal consistent* subset of X, $X \vdash_W \neg\sigma$. This holds for every $\sigma \in X$, so $\neg\sigma \in Cl^\star_{\mathcal{R}_s}(X)$. It can be concluded that AF satisfies the self-contradiction axiom. □

Lemma 7.11 Each *ASPIC** *SAF* with $\vdash_W$ satisfies the following property: for any set S of arguments* it holds that if $\texttt{Conc}(S) \vdash_{CL} \varphi$, then $\texttt{Conc}(S^\#) \vdash_{CL} \varphi$.

Proof. Suppose that $\texttt{Conc}(S) \vdash_{CL} \varphi$. Consider any argument* $T_i \rightarrow \varphi_i$ in $S \setminus S^\#$. By definition of $\vdash_W$ and choice of $\mathcal{R}_s$ it holds that $\texttt{Conc}(T_i) \vdash_{CL} \varphi_i$ for any i. Then since $\vdash_{CL}$ satisfies the **Cut** rule, we have that $\texttt{Conc}(S^\#) \vdash_{CL} \varphi$. □

Lemma 7.12 Each *ASPIC** *SAF* with $\vdash_W$ satisfies the following property: for any set E of arguments* it holds that if $\texttt{Conc}(E)$ is strictly closed* and consistent*, then $\texttt{Conc}(E)$ is classically consistent.

Proof. Assume that $\texttt{Conc}(E)$ is strictly closed* and consistent* and suppose for contradiction that $\texttt{Conc}(E) \vdash_{CL} \varphi, \neg\varphi$. Consider any subset-minimal $S \subseteq E$ such that $\texttt{Conc}(S) \vdash_{CL} \varphi, \neg\varphi$. Then by Lemma 7.11 we have that $\texttt{Conc}(S^\#) \vdash_{CL} \varphi, \neg\varphi$. Consider any minimal $T \subseteq S^\#$ such that $\texttt{Conc}(T) \vdash_{CL} \varphi, \neg\varphi$. Note that T cannot be empty. Then for any $\psi \in \texttt{Conc}(T)$ it holds that $\texttt{Conc}(T) \backslash \{\psi\} \vdash_{CL} \neg\psi$. By choice of T it holds that $\texttt{Conc}(T) \backslash \{\psi\}$ is classically consistent. But then since $\texttt{Conc}(E)$ is strictly closed* and no argument* in $S^\#$ has a strict top rule, there exists an argument* $T' \rightarrow \neg\psi$ in E for some $T' \subset T$ such that $\texttt{Conc}(T') = \texttt{Conc}(T) \setminus \{\psi\}$. But then $\texttt{Conc}(E)$ is not consistent*. □

Theorem 7.13 Each *ASPIC** *SAF* with $\vdash_W$ which is axiom consistent* and has a reasonable argument ordering satisfies the logical closure postulate.

Proof. Suppose $\texttt{Conc}(E) \vdash_W \varphi$ for some complete extension E and consider any subset-minimal $S \subseteq E$ such that $\texttt{Conc}(S) \vdash_W \varphi$. Then by definition of $\vdash_W$ and choice of $\mathcal{R}_s$ it holds that $\texttt{Conc}(S)$ is classically consistent and $\texttt{Conc}(S) \vdash_{CL} \varphi$. Then by Lemma 7.11 we have that $\texttt{Conc}(S^\#) \vdash_{CL} \varphi$. Moreover, $\texttt{Conc}(S^\#)$ is classically consistent because of Lemma 7.12 and the fact that any subset of a classically consistent set is also classically consistent (note that by Theorem 7.10 $\texttt{Conc}(E)$ is strictly closed* and consistent*, so the conditions of Lemma 7.12 are fulfilled). But then $\texttt{Conc}(S^\#) \vdash_W \varphi$ and by choice of R_s there exists a strict rule $\texttt{Conc}(S^\#) \rightarrow \varphi$. Since no argument* in $S^\#$ has a strict top rule, $\varphi \in \texttt{Conc}(E)$ by strict closure* of E. □

A.2 Proofs for Section 8.2

Lemma 8.1 For any argument* A and any extended argument* A^+ the following holds: for any $A' \in \texttt{Sub}(A)$ there is an argument* $A'' \in \texttt{Sub}(A^+)$ such that $A'' = A'^+$.

Proof. This proof is a proof by induction on the height of argument* A (viewed as a directed acyclic graph). Suppose A is an element of $\mathcal{K}$ (so the height is 1), then A' and A^+ have to be equal to A. This means that A'' is also equal to A and it is easy to see that $A'' = A'^+$. Suppose that the lemma holds for all arguments* of height i for an $i \in \{1, 2, \ldots\}$. Now it has to be proven for arguments* of height $i+1$. Take an arbitrary argument* A of height $i+1$ and take a subargument* A' of A. Note that A cannot be an element of $\mathcal{K}$ since the height is greater than 1. Therefore, A has to be of the form $A_1, \ldots, A_n \rightarrow / \Rightarrow \varphi$. Then there are two possibilities: either (i) A' is a subargument* of one of the arguments* $A_1, \ldots, A_n$, or (ii) A' is equal to A.
(i). There is a $j \in \{1, \ldots, n\}$ such that A' is a subargument* of A_j. A_j has a height of i, so there must be an $A'' \in \texttt{Sub}(A_j^+)$ such that $A'' = A'^+$. According to Definition 8.4, A'' is a subargument* of A^+.
(ii). A' is equal to A. Take A'' to be A^+, then it follows that $A'' = A'^+$ and $A'' \in \texttt{Sub}(A^+)$.
Now it is proved that for every argument* the lemma holds. □

Lemma 8.2 If $\preceq$ is a tolerable argument* ordering and argument* A defeats/attacks B, then every A^- defeats/attacks every B^+.

Proof.
Attack
(i). Suppose that A undercuts B, so $\texttt{Conc}(A) = \neg n(r)$ for a defeasible top rule r of a $B' \subseteq \texttt{Sub}(B)$. By definition of A^-, the conclusion for every A^- is the same as for A. By Lemma 8.1 it holds that for every B^+, there is a subargument* $B'' \in \texttt{Sub}(B^+)$ such that $B'' = B'^+$. Note that B' and B'' both have the same defeasible top rule, since by definition of B^+, only the strict rules can mutate. Then it follows that every A^- undercuts B^+ on B''.
(ii). Suppose now that A undermines B, so $\neg\texttt{Conc}(A)$ is the ordinary premise B'. Every B^+ also has this ordinary premise. By definition of A^- (Definition 8.2), the conclusion of every A^- is the same as for A. Therefore any A^- undermines every B^+ on B'.
(iii). Suppose that A rebuts B, so $\neg\texttt{Conc}(A)$ is the conclusion of some basic fallible argument* $B' \in \texttt{Sub}(B)$. The conclusion for every A^- is the same as for A. By Lemma 8.1, it holds that for every B^+, there is a subargument* $B'' \in \texttt{Sub}(B^+)$ such that $B'' = B'^+$. Note that the conclusions of B' and B'' are the same and that B'' also has a defeasible top rule. Therefore any A^- rebuts B^+ on B''.
Defeat
Suppose argument* A defeats B. This means that A attacks B on a subargument* B'. Then (i) A undercuts B, or (ii) $A \not\prec B'$.
(i). Suppose that A undercuts B, then it follows that every A^- undercuts B^+ on B'' (see reasoning above for attack (i)). This implies that any A^- defeats every B^+.
(ii). Otherwise it has to be the case that $A \not\prec B'$. Note that, because of the tolerable argument* ordering, $A^- \succeq A$. In the proof for attack it is shown that every A^- attacks

every B^+ on B'', where $B'' = B'^+$. The tolerable argument* ordering causes $B'' \preceq B'$. Therefore for every A^- and B'' it holds that $A^- \not\prec B''$, so every A^- defeats every B^+. □

Lemma 8.3 If $\preceq$ is a tolerable argument* ordering, then for all complete extensions E:

1. If $A \in E$, then $A^- \in E$ for every A^-;
2. If $B \notin E$, then $B^+ \notin E$ for all B^+.

Proof. (1). Suppose $A \in E$ and let B be an argument* defeating an A^-. Then by Lemma 8.2, B defeats A. Therefore, there must be an argument* $C \in E$ such that C defeats B. Hence, A^- is defended by E and thus $A^- \in E$.
(2). Suppose $B \notin E$. Then there exists an argument* A such that A defeats B and there does not exist a $C \in E$ that defeats A. According to Lemma 8.2, A defeats every B^+, so B^+ is not defended by E and hence $B^+ \notin E$. □

Lemma 8.4 Let $SAF^- = (\mathcal{A}^-, \mathcal{C}^-, \preceq^-)$ be the minimal structured argumentation framework corresponding to $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ (for the *ASPIC** framework), and let AF be the abstract argumentation framework corresponding to SAF. Let $\preceq$ be a tolerable argument* ordering. Also, let C and C^- be the sets of complete extensions of SAF and SAF^- respectively and let F_{AF} be Dung's characteristic function of AF. Then:

1. For each $E \in C^- : F_{AF}(E)^- = E$.
2. For each $E \in C : F_{AF}(E^-) = E$ and $E^- \in C^-$.

Proof. Take two sets of arguments* X and Y such that $X \subseteq Y$. Suppose $F_{AF}(Y) \subset F_{AF}(X)$, then there must be an argument* A that is defended by X but not by Y. Since $X \subseteq Y$, A has to be defended by Y. Therefore, the monotonicity of $F_{AF}(E)$ with respect to set inclusion is obvious.
(1). It will be shown that $F_{AF}(E)$ is indeed a function from C^- into C such that $F_{AF}(E)^- = E$ by showing that $F_{AF}(E)$ is a complete extension in SAF, if E is a complete extension in SAF^-. Let E be a complete extension in SAF^-. First it has to be shown that E is an admissible set in SAF.
E is conflict-free in $\mathcal{A}^-$, so it has to be conflict-free in $\mathcal{A}$. It also defeats each minimal argument* defeating E and it contains every minimal argument* defended by E.
For any $A \in E$ and any $B \in \mathcal{A}$ that defeats A, take a B^-. Then $B^- \in \mathcal{A}^-$ defeats A by Lemma 8.2, so some $C \in E$ defeats B^-. But then C also defeats B by Lemma 8.2 combined with the fact that B is a B^{-+}. This means that $A \in \mathcal{A}$ is defended by E. It can be concluded that E defeats every argument* that defeats E. It was already shown that E is conflict-free, so $E \subseteq \mathcal{A}$ is an admissible set.
For each argument* $A \in F_{AF}(E)$, any A^- has to be defended by E in $\mathcal{A}^-$, so every A^- is in E. Therefore, $(F_{AF}(E) - E)^- = \emptyset$, thus $F_{AF}(E)^- = E$.
Now, it has to be shown that $F_{AF}(E)$ is a complete extension. Let A be defended by $F_{AF}(E)$ and let B defeat A. Hence, there is a $C \in F_{AF}(E)$ defeating B. $C \in F_{AF}(E)$ means that C is defended by E, thus E defeats every argument* defeating C. Suppose an argument* D defeats C^-, then D also defeats C by Lemma 8.2. Therefore, E must defeat D, so C^- is defended by E. Since E is a complete extension of SAF^-, any C^- is in E. C^-

defeats B (Lemma 8.2). Therefore, E defeats B. Thus, A is defended by E, and therefore $A \in F_{AF}(E)$. As a consequence, $F_{AF}(E)$ is complete.
(2). Let E be a complete extension in SAF. It will be shown that E^- is complete in SAF^-. First it has to be shown that E^- is an admissible set of $\mathcal{A}^-$. Since E is conflict-free, E^- has to be conflict-free as well.
From Lemma 8.3 and the fact that E is complete in SAF, each minimal version of the arguments* in E belongs to E. Let $A \in \mathcal{A}^-$ defeat E^-. Hence, there is $B \in E$ defeating A. According to Lemma 8.3 $B^- \in E$, so $B^- \in E^-$. Hence, B^- defeats A (Lemma 8.2). Thus, E^- is admissible.
Each minimal argument* defended by E^- is defended by E and hence belongs to E and so to E^-. E^- is therefore complete.
Since $E^- \subseteq E$ and E is complete, it is clear that $F_{AF}(E^-) \subseteq F_{AF}(E) = E$. It is now shown that each argument* defended by E is also defended by E^-. Let A be an argument* defended by E in SAF and let B be an argument* defeating A. Hence, there is an argument* $C \in E$ defeating B and so each $C^- \in E$ defeats B. Hence E^- defeats B. Thus A defended by to E^- in SAF. It can be concluded that $F_{AF}(E^-) \supseteq F_{AF}(E) = E$, i.e. $F_{AF}(E^-) = E$.
[Injective] Take $X, Y \in C^-$ such that $F_{AF}(X) = F_{AF}(Y)$. It is obvious that $F_{AF}(X)^- = F_{AF}(Y)^-$. Then according to the proof for point (1) $F_{AF}(X)^- = X$ and $F_{AF}(Y)^- = Y$. Therefore, it follows that $X = Y$.
[Surjective] It has to be shown that for all $Y \in C$ there is an $X \in C^-$ such that $F_{AF}(X) = Y$. Now take X to be Y^-, then the proof for point (2) provides that $X \in C^-$ and that $F_{AF}(X) = Y$.
[Bijective] Injectivity and surjectivity provides that F_{AF} is a bijection from C^- onto C. □

Theorem 8.5 Let $SAF^- = (\mathcal{A}^-, \mathcal{C}^-, \preceq^-)$ be the minimal structured argumentation framework corresponding to $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ (for the *ASPIC** framework). Let $\preceq$ be a tolerable argument* ordering. Take $T \in$ {complete, grounded, preferred, stable}, then:

1. Let E be a T extension in SAF, then E^- is a T extension in SAF^-.
2. Let E be a T extension in SAF^-, then $F(E)$ is a T extension in SAF.

Proof.
[$T =$ complete]
Let C and C^- be the sets of complete extensions of SAF and SAF^-, respectively. Lemma 8.4 states that F_{AF} is a bijection from C^- to C. This immediately provides that $F(E) \in C$. The second point of Lemma 8.4 states that $E^- \in C^-$.
[$T \in$ {grounded, preferred}]
From Lemma 8.4 it follows immediately that for each $E \in C^-$, $F_{AF}(E)$ is minimal or maximal with respect to set inclusion in E if and only if E is minimal or maximal respectively in C^-. Hence E is grounded or preferred in SAF^- if and only if $F_{AF}(E)$ is grounded or preferred in SAF, respectively.
[$T =$ stable]
(1). Take E to be a stable extension in SAF. Suppose for contradiction that E^- is not a stable extension in SAF^-. Then there must be an argument* $A \in \mathcal{A}^-$, $A \notin E^-$ such that A is not defeated by any argument* in E^-. However, $A \notin E^-$ implies $A \notin E$ since A is minimal. E is stable, thus there must be an argument* $B \in E$ such that B defeats A. But

then any $B^- \in E$ (Lemma 8.3) defeats A. It is clear that $B \in E^-$, so this is a contradiction with the fact that E^- does not defeat A. Therefore, E^- is a stable extension in SAF^-.
(2). Take E to be a stable extension of SAF^-. Suppose for contradiction that $F_{AF}(E)$ is not a stable extension of SAF. Then there must be an argument* $A \in \mathcal{A}$ such that $A \notin F_{AF}(E)$ and $F_{AF}(E)$ does not defeat A. $A \notin F_{AF}(E)$ means that A is defended by E. Therefore, there must be an argument* $B \in \mathcal{A}$ that defeats A such that there is not a $C \in E$ that defeats B. Since E is stable in SAF^-, $B \in E$ so E defeats A. This implies that $F_{AF}(E)$ defeats A, which is a contradiction with the fact that $F_{AF}(E)$ does not defeat A. Therefore, $F_{AF}(E)$ has to be a stable extension of SAF. □

A.3 Proofs for Section 8.3

Lemma 8.6 For any finite $Y \subseteq \mathcal{L}$ it holds that $\mathcal{X}^Y$ is finite.

Proof. Note that for any φ there is at most a single strict rule $\rightarrow \varphi$, while since both $\mathcal{K}$ and $\mathcal{R}_d$ are finite, the set of all sets that equal the set of antecedents of a strict rule for φ in $\mathcal{X}$ is also finite. □

Theorem 8.7 Let $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ be a strongly minimal structured argumentation framework corresponding to an argumentation theory with a finite $\mathcal{K}$ and $\mathcal{R}_d$. Then for any $\varphi \in \mathcal{L}$ the set $\{A \in \mathcal{A} \mid \texttt{Conc}(A) = \varphi\}$ is finite.

Proof. We prove the result by iteratively constructing of the set of all finite arguments* constructible from $\mathcal{K}$ in AS.

- $\mathcal{A}_0 = \mathcal{K}$ and $\mathcal{A}_0^d = \mathcal{A}_0^s = \emptyset$;
- $\mathcal{A}_{i+1} = \mathcal{A}_i \cup \mathcal{A}_{i+1}^d = \mathcal{A}_{i+1}^s$ where
  - $\mathcal{A}_{i+1}^d = \{A = Q \Rightarrow \varphi \mid Q \subseteq \mathcal{A}_i$ and $\texttt{Conc}(Q) \Rightarrow \varphi \in \mathcal{R}_d$ and $\varphi \notin \texttt{Conc}(Q)\}$.
  - $\mathcal{A}_{i+1}^s = \{A = Q \rightarrow \varphi \mid Q \subseteq \mathcal{A}_i$ and $\texttt{Conc}(Q) \rightarrow \varphi \in \mathcal{R}_s$ and $\varphi \notin \texttt{Conc}(Q)$ and for no $B \in Q$ it holds that $\texttt{TopRule}(B) \in \mathcal{R}_s\}$.

It is easy to see that $\bigcup_{i=0}^n \mathcal{A}_i$ is the set of all constructible finite arguments*.

To prove the theorem, note first that $\mathcal{A}_0$ is finite since $\mathcal{K}$ is finite and $\mathcal{A}_i^d$ is finite for any i since $\mathcal{R}_d$ is finite. By contrast, $\mathcal{A}_i^s$ is infinite for any i since given our choice of $\mathcal{R}_s$, to any set of wff an infinite number of strict rules applies. However, it follows from Lemma 8.6 that for any i and any $\varphi \in \mathcal{L}$ there exists at most a finite number of arguments* in $\mathcal{A}_i^s$ with conclusion φ. Together these observations imply that for any i and any $\varphi \in \mathcal{L}$ there exists at most a finite number of arguments* with conclusion φ in $\mathcal{A}_i$.

It is left to prove that the construction is finite. This follows from the fact that both $\mathcal{K}$ and $\mathcal{R}_d$ are finite and the fact that by Definition 8.7 rules cannot be repeated in an argument*. So at some point, $\mathcal{A}_i^d = \emptyset$. But then also $\mathcal{A}_{i+1}^s = \emptyset$ since strict rules cannot be chained in an argument*. So the construction of the set of all finite arguments* constructible from $\mathcal{K}$ in AS is finite, while for each $\varphi \in \mathcal{L}$ at each step in the construction only a finite set of arguments* with conclusion φ is created. This proves the theorem. □

Theorem 8.8 Each strongly minimal *ASPIC** *SAF* with $\vdash_W$ is compact.

Proof.
[$\texttt{Conc}(GN(S)) \supseteq Cl^\star_{\mathcal{R}_s}(GN(S)^\#)$]
Let S be a set of arguments* and $\sigma \in Cl^\star_{\mathcal{R}_s}(GN(S)^\#)$. It needs to be shown that $\sigma \in \texttt{Conc}(GN(S))$. Let X be a minimal subset of $\texttt{Conc}(GN(S)^\#)$ such that $\sigma \in Cl^\star_{\mathcal{R}_s}(X)$. Hence there is a strict argument* A_0 over X with conclusion σ. Further let S_X be a minimal set of arguments* from $GN(S)^\#$ s.t. $\texttt{Conc}(S_X) = X$. There is no problem with repetition of conclusions here, since if $\sigma \in X$ then $A_0 = \sigma$. Let A be the argument* obtained by replacing each leaf in A_0 (viewed as a directed acyclic graph) labelled by a literal α from X by an argument* with conclusion α from S_X. Note that if A repeats conclusion σ then we do not need A since then clearly $\sigma \in \texttt{Conc}(GN(S))$ and we are done. Otherwise, the construction of A is possible since, firstly, all arguments* in S_X are basic fallible arguments* or are just necessary premises and, second, by definition of $GN(S)^\#$ it holds that A will not repeat any conclusion in X. It is obvious that the conclusion of A is σ. It is shown that S_X is a base of A. Suppose B is an argument* defeating A. Since A_0 is a strict argument* over X, B must defeat a basic fallible subargument* in S_X. Hence B defeats S_X. Thus $A \in GN(S)$. Hence $\sigma \in \texttt{Conc}(GN(S))$.
[$\texttt{Conc}(GN(S)) \subseteq Cl^\star_{\mathcal{R}_s}(GN(S)^\#)$]
As for Theorem 7.6.
It is proven that the strongly minimal *SAF* is compact. □

A.4 Proofs for Section 9

Theorem 9.1 Let *AS* be an argumentation system with $\mathcal{L}$ a classical-logic language and $\mathcal{R}_s$ corresponding to classical logic, and let $\mathcal{K} = \mathcal{K}_p$ be a knowledge base in $\mathcal{L}$. Let AS^* be obtained from *AS* by removing from R_s all inference rules that are invalid according to $\vdash_W$. Let $AT = (AS, \mathcal{K})$ be a minimal *ASPIC*$^+$ argumentation theory and $AT^* = (AS^*, \mathcal{K})$ the corresponding minimal *ASPIC** argumentation theory. Then let $SAF = (\mathcal{A}, \mathcal{C}, \preceq)$ and $SAF^* = (\mathcal{A}^*, \mathcal{C}^*, \preceq^*)$ be, respectively, the minimal c-structured argumentation framework defined by *AT* in *ASPIC*$^+$ and the minimal structured argumentation framework defined by AT^* in *ASPIC** such that:

- For all $A, B, C \in \mathcal{A}$: if $\texttt{Prem}(A) = \texttt{Prem}(B)$ then $C \preceq A$ iff $C \preceq B$ and $A \preceq C$ iff $B \preceq C$. Likewise for $\preceq^*$.
- For all A^*, $B^* \in \mathcal{A}^*$: $A^* \preceq B^*$ iff $A^* \preceq^* B^*$.

Then for $T \in$ {complete, grounded, preferred, stable} it holds that:

1. Let E be a T extension in *SAF*, then $E \cap \mathcal{A}^*$ is a T extension in SAF^* such that $\texttt{Conc}(E) = \texttt{Conc}(E \cap \mathcal{A}^*)$.
2. Let E be a T extension in SAF^*, then $F(E)$ is a T extension in *SAF* such that $\texttt{Conc}(E) = \texttt{Conc}(F(E))$.

Proof. First, it is easy to see that $\mathcal{A}^* \subseteq \mathcal{A}$. Next, we show by induction on the structure of arguments* that for any argument* $A \in \mathcal{A} \setminus \mathcal{A}^*$ there exists an argument* $B \in \mathcal{A}^*$ such that $\texttt{Prem}(A) = \texttt{Prem}(B)$ and $\texttt{Conc}(A) = \texttt{Conc}(B)$.

The base case is that $A \in \mathcal{K}$. Then $B = A$. For the inductive case consider $A = B_1, \ldots, B_n \rightarrow \varphi$. By the induction hypothesis the arguments* $B_1^*, \ldots, B_n^*$ satisfy the conditions. Then for any $B_i^* (1 \leq i \leq n)$ of the form $B_i^* = B_k^{*i}, \ldots, B_m^{*i} \rightarrow \varphi_i$ the induction hypothesis yields that $B_k^{*i}, \ldots, B_m^{*i}$ are from $\mathcal{K}$. Then replacing all such B_i^* with $B_k^{*i}, \ldots, B_m^{*i}$ in $B_1^*, \ldots, B_n^* \rightarrow \varphi$ yields an argument* B satisfying the conditions. Note that the corresponding strict inference rule is in $\mathcal{R}_s$ since classical logic satisfies the cut rule and all arguments* in $\mathcal{A}$ have consistent premises.

Then it is easy to prove along the lines of the proof of Lemma 8.2 for any argument* $A \in \mathcal{A}$ that any argument* attacked/defeated by A is also attacked/defeated by any B and vice versa, and any argument* attacking/defeating A also attacks/defeats any B and vice versa. Then the proof of Theorem 9.1 can be completed along the lines of the proof of Theorem 8.5. □

References

Amgoud, L., & Besnard, P. (2013). Logical limits of abstract argumentation frameworks. *Journal of Applied Non-classical Logics*, *23*, 229–267.

Bench-Capon, T., Prakken, H., & Visser, W. (2011). Argument schemes for two-phase democratic deliberation. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Law*, pp. 21–30, New York. ACM Press.

Besnard, P., & Hunter, A. (2008). *Elements of Argumentation*. MIT Press, Cambridge, MA.

Bondarenko, A., Dung, P., Kowalski, R., & Toni, F. (1997). An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, *93*, 63–101.

Brewka, G. (1991). *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge University Press, Cambridge.

Caminada, M. (2005). Contamination in formal argumentation systems. In *Proceedings of the Seventeenth Belgian-Dutch Conference on Artificial Intelligence (BNAIC-05)*, Brussels, Belgium.

Caminada, M., & Amgoud, L. (2007). On the evaluation of argumentation formalisms. *Artificial Intelligence*, *171*, 286–310.

Caminada, M., Carnielli, W., & Dunne, P. (2012). Semi-stable semantics. *Journal of Logic and Computation*, *22*, 1207–1254.

Caminada, M., Modgil, S., & Oren, N. (2014). Preferences and unrestricted rebut. In Parsons, S., Oren, N., Reed, C., & Cerutti, F. (Eds.), *Computational Models of Argument. Proceedings of COMMA 2014*, pp. 209–220. IOS Press, Amsterdam etc.

Caminada, M., Sa, S., Alcantara, J., & Dvorak, W. (2015). On the difference between assumption-based argumentation and abstract argumentation. *IFCoLog Journal of Logic and its Applications*, *2*, 15–34.

Da Costa, N. (1974). On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, *15*(4), 497–510.

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

Dung, P. (2014). An axiomatic analysis of structured argumentation for prioritized default reasoning. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI 2014)*, pp. 267–272.

Dung, P. (2016). An axiomatic analysis of structured argumentation with priorities. *Artificial Intelligence*, *231*, 107–150.

Dung, P., Mancarella, P., & Toni, F. (2007). Computing ideal sceptical argumentation. *Artificial Intelligence*, *171*, 642–674.

Dung, P., & Thang, P. (2014). Closure and consistency in logic-associated argumentation. *Journal of Artificial Intelligence Research*, *49*, 79–109.

Dung, P., Toni, F., & Mancarella, P. (2010). Some design guidelines for practical argumentation systems. In Baroni, P., Cerutti, F., Giacomin, M., & Simari, G. (Eds.), *Computational Models of Argument. Proceedings of COMMA 2010*, pp. 183–194. IOS Press, Amsterdam etc.

Garcia, A., & Simari, G. (2004). Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming*, *4*, 95–138.

Ginsberg, M. (1994). AI and nonmonotonic reasoning. In Gabbay, D., Hogger, C., & Robinson, J. (Eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*, pp. 1–33. Clarendon Press, Oxford.

Gorogiannis, N., & Hunter, A. (2011). Instantiating abstract argumentation with classical-logic arguments: postulates and properties. *Artificial Intelligence*, *175*, 1479–1497.

Grooters, D., & Prakken, H. (2014). Combining paraconsistent logic with argumentation. In Parsons, S., Oren, N., Reed, C., & Cerutti, F. (Eds.), *Computational Models of Argument. Proceedings of COMMA 2014*, pp. 301–312. IOS Press, Amsterdam etc.

Hunter, A. (2007). Real arguments are approximate arguments. In *Proceedings of the 22nd National Conference on Artificial Intelligence (AAAI-07)*, pp. 66–71.

Jakobovits, H., & Vermeir, D. (1999). Robust semantics for argumentation frameworks. *Journal of Logic and Computation*, *9*, 215–261.

Krause, P., Ambler, S., Elvang-Gøransson, M., & Fox, J. (1995). A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, *11*(1), 113–131.

Lin, F., & Shoham, Y. (1989). Argument systems. A uniform basis for nonmonotonic reasoning. In *Principles of Knowledge Representation and Reasoning: Proceedings of the First International Conference*, pp. 245–255, San Mateo, CA. Morgan Kaufmann Publishers.

Modgil, S., & Prakken, H. (2013). A general account of argumentation with preferences. *Artificial Intelligence*, *195*, 361–397.

Modgil, S., & Prakken, H. (2014). The ASPIC+ framework for structured argumentation: a tutorial. *Argument and Computation*, *5*, 31–62.

Pollock, J. (1987). Defeasible reasoning. *Cognitive Science*, *11*, 481–518.

Pollock, J. (1990). A theory of defeasible reasoning. *International Journal of Intelligent Systems*, *6*, 33–54.

Pollock, J. (1992). How to reason defeasibly. *Artificial Intelligence*, *57*, 1–42.

Pollock, J. (1994). Justification and defeat. *Artificial Intelligence*, *67*, 377–408.

Pollock, J. (1995). *Cognitive Carpentry. A Blueprint for How to Build a Person.* MIT Press, Cambridge, MA.

Prakken, H. (2010). An abstract framework for argumentation with structured arguments. *Argument and Computation*, *1*, 93–124.

Prakken, H. (2012). Some reflections on two current trends in formal argumentation. In *Logic Programs, Norms and Action. Essays in Honour of Marek J. Sergot on the Occasion of his 60th Birthday*, pp. 249–272. Springer, Berlin/Heidelberg.

Prakken, H., & Modgil, S. (2012). Clarifying some misconceptions on the ASPIC+ framework. In Verheij, B., Woltran, S., & Szeider, S. (Eds.), *Computational Models of Argument. Proceedings of COMMA 2012*, pp. 442–453. IOS Press, Amsterdam etc.

Prakken, H., & Sartor, G. (1997). Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-classical Logics*, *7*, 25–75.

Prakken, H., & Vreeswijk, G. (2002). Logics for defeasible argumentation. In Gabbay, D., & Günthner, F. (Eds.), *Handbook of Philosophical Logic* (Second edition)., Vol. 4, pp. 219–318. Kluwer Academic Publishers, Dordrecht/Boston/London.

Prakken, H., Wyner, A., Bench-Capon, T., & Atkinson, K. (2015). A formalisation of argumentation schemes for legal case-based reasoning in ASPIC+. *Journal of Logic and Computation*, *25*, 1141–1166.

Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, *8*(1), 219–241.

Priest, G. (1989). Reasoning about truth. *Artificial Intelligence*, *39*(2), 231–244.

Read, S. (1988). *Relevant logic.* Blackwell, Oxford.

Rescher, N., & Manor, R. (1970). On inference from inconsistent premises. *Journal of Theory and Decision*, *1*, 179–219.

Simari, G., & Loui, R. (1992). A mathematical treatment of defeasible argumentation and its implementation. *Artificial Intelligence*, *53*, 125–157.

Vreeswijk, G. (1997). Abstract argumentation systems. *Artificial Intelligence*, *90*, 225–279.

Walton, D. (1996). *Argumentation Schemes for Presumptive Reasoning.* Lawrence Erlbaum Associates, Mahwah, NJ.

Wu, Y. (2012). *Between Argument and Conclusion. Argument-based Approaches to Discussion, Inference and Uncertainty.* Doctoral Dissertation Faculty of Sciences, Technology and Communication, University of Luxemburg.

Wu, Y., & Podlaszewski, M. (2015). Implementing crash-resistence and non-interference in logic-based argumentation. *Journal of Logic and Computation*, *25*, 303–333.