Artificial Intelligence 231 (2016) 107–150

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

An axiomatic analysis of structured argumentation with priorities

Phan Minh Dung

Computer Science and Information Management, Asian Institute of Technology, PO Box 4, Klong Luang, Pathumthani 12120, Thailand

ARTICLE INFO

Article history:
Received 9 November 2014
Received in revised form 14 September 2015
Accepted 26 October 2015
Available online 1 December 2015

Keywords:
Argumentation
Structured argumentation with priorities
Ordinary and normal attack relation assignments
Normal form

ABSTRACT

Several systems of semantics have been proposed for structured argumentation with priorities. As the proposed semantics often sanction contradictory conclusions (even for skeptical reasoners), there is a fundamental need for guidelines for understanding and evaluating them, especially their conceptual foundation and relationship.

In this paper, we present an axiomatic analysis of the semantics of structured defeasible argumentation both with and without preferences by introducing a class of ordinary attack relations satisfying a set of simple and intuitive properties. We show that there exists a "normal form" for ordinary attack relations in the sense that stable extensions wrt any ordinary attack relation are stable extensions wrt the normal attack relations.

We relate the ordinary semantics to other approaches, especially to the ASPIC+ framework and the prioritized approaches in logic programming.

1. Introduction

Prioritized defeasible reasoning has been studied extensively [35,34,8,15,44,40,10,26,24,43]. Distinct semantics are proposed that could give different (even contradictory) answers to the same query as the following example illustrates.

Example 1.1. Consider a knowledge base K (adapted from [10,11]), consisting of three defeasible rules

$$d_1 : \mathit{Dean} \Rightarrow \mathit{Professor} \quad d_2 : \mathit{Professor} \Rightarrow \mathit{Teach} \quad d_3 : \mathit{Administrator} \Rightarrow \neg \mathit{Teach}$$

and two strict rules

$$r : \mathit{Dean} \to \mathit{Administrator} \quad r' : \neg \mathit{Administrator} \to \neg \mathit{Dean}$$

with $d_1 \prec d_3 \prec d_2$[1] and $d_i \preceq d_i$, $i = 1..3$.

Suppose we know some Dean. The question is *whether the dean teaches.*[2]

Proposed semantics in literature deal with this example differently.

An influential and important approach to structured argumentation is the ASPIC+ framework. Modgil and Prakken [34] proposed four attack relations based on the last- or weakest-link principles coupled with the elitist- or democratic-orderings

E-mail address: dung.phanminh@gmail.com.

[1] $d \prec d'$ means that d is less preferred than d'.

[2] The relevant arguments concerning this question are given in Figs. 3, 4.

http://dx.doi.org/10.1016/j.artint.2015.10.005
0004-3702/

that are arguably the most prominent attack relations proposed until now in the ASPIC+ framework. One of these four attack relations, the one based on the weakest link and elitist ordering leads to semantics with respect to which *the dean does not teach* while the other three as well as the non-argument-based approach of Brewka and Eiter [10] lead to conclusion that *the dean does teach*. □

As the proposed approaches to defeasible reasoning with priorities [35,34,8,15,44,40,10,26,24,43] often sanction contradictory conclusions (even for skeptical reasoners) there is a fundamental need for guidelines for understanding and evaluating them, especially their conceptual foundation and relationship when a user applies prioritized defeasible reasoning in reality.

A key property for evaluating the semantics of structured argumentation is the attack monotonicity. For a quick illustration of this property imagine you have a lively dancing bird in your garden and you know that it is a penguin.[3,4] Suppose some neighbour tells you that the bird is most likely a penguin.[5] Will it change anything in your beliefs about your bird? Of course not. This is an example of the property of irrelevance of redundant defaults stating that adding redundant defaults into your knowledge base does not change your beliefs. This simple and natural property follows from the property of attack monotonicity. Proposed semantics in literature behave differently wrt these properties.

Example 1.2 *(A Sherlock Holmes investigation).* Sherlock Holmes is investigating a case involving three persons P_1, P_2 and S together with the dead body of a big man. The case could be represented by the following knowledge base.

1. The knowledge that one of the persons is the murderer is represented by three strict rules:

$$r_1: \ Inno(P_1), Inno(S) \to \neg Inno(P_2)^{6}$$
$$r_2: \ Inno(P_2), Inno(S) \to \neg Inno(P_1)$$
$$r_3: \ Inno(P_1), Inno(P_2) \to \neg Inno(S)$$

2. S is a small child who cannot kill a big man. This fact is captured in the base of evidence $BE = \{Inno(S)\}$.
3. The legal principle that people are considered innocent until proven otherwise could be represented in two ways:
   - By three defeasible rules

$$d_1: \ \Rightarrow Inno(P_1) \quad d_2: \ \Rightarrow Inno(P_2) \quad d: \ \Rightarrow Inno(S)$$

   - By two defeasible rules

$$d_1: \ \Rightarrow Inno(P_1) \quad d_2: \ \Rightarrow Inno(P_2)$$

   as S is innocent, and hence the defeasible rule $d :\Rightarrow Inno(S)$ is intuitively redundant.
4. After digging around, it becomes clear to Holmes that P_1 has a strong motive to kill the victim while there is nothing connecting P_2 to the dead man. He hence will focus his investigation on P_1. This knowledge is represented by a preference

$$d_1 \prec d_2$$

   stating that Holmes gives higher priority (in his investigation) to the scenario in which P_2 is innocent than to the other one.

Let KB_1 be the knowledge base containing the strict rules r_1, r_2, r_3, the three defaults d_1, d_2, d and the fact that S is innocent together with the preference $d_1 \prec d_2$.

Further let KB_0 be the knowledge base obtained from KB_1 by removing defeasible rule $d :\Rightarrow Inno(S)$.

Due to the fact that S is innocent, we expect that default d will have no impact on the belief sets of the knowledge base KB_1. In other words, *both KB_1 and KB_0 are expected to have identical belief sets*, concluding

$$\neg Inno(P_1), \ Inno(P_2)$$

Surprisingly, KB_0, KB_1 have different belief sets wrt the semantics based on attack relations employing the democratic order proposed and studied by Modgil and Prakken in [34] as elaborated below.

For ease of reference, we refer to the attack relations proposed and studied by Modgil and Prakken in [34] as MP-attack relations in the rest of this example.

[3] In other words, it is an undisputed fact to you that the bird is a penguin. According to Definition 3.3, $BE = \{penguin\}$.

[4] Remember Mumble, the main penguin character in the animated movie Happy Feet?

[5] In other words, you add a defeasible rule $\Rightarrow penguin$ to your knowledge base.

[6] Inno stands for Innocent.

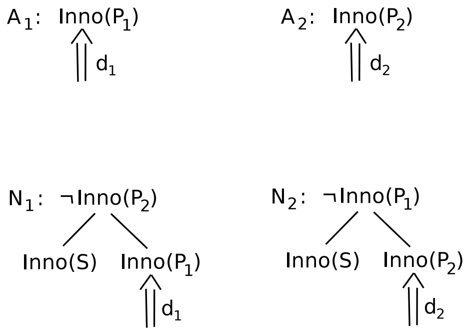

Fig. 1. Arguments of KB_0.

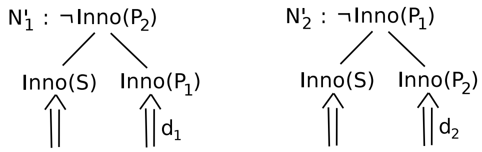

Fig. 2. Extra arguments for KB_1.

- Relevant arguments concerning the innocence of P_1, P_2 wrt KB_0 are given in Fig. 1. Due to the preference of d_2 over d_1, N_2 attacks A_1 but N_1 does not attack A_2 wrt all four MP-attack relations in the ASPIC+ framework. Therefore N_2 also attacks N_1. The unique stable extension for KB_0 thus contains A_2, N_2. Hence

$$\neg Inno(P_1), Inno(P_2) \quad \text{are skeptically justified (as expected) for } KB_0$$

- There are two new relevant arguments concerning the innocence of P_1, P_2 wrt KB_1 (illustrated in Fig. 2). According to the MP-attack relations based on the democratic order, N'_1 attacks A_2. Hence N'_1 also attacks N_2, N'_2. Therefore there is a new stable extension containing A_1, N'_1, N_1 justifying:

$$Inno(P_1),\ \neg Inno(P_2), \ \text{a counter-intuitive set of beliefs}$$

A closer look reveals that MP-semantics based on democratic order behave counter-intuitively because they violate the principle of attack monotonicity as follows:

N_1 could be obtained from N'_1 by replacing the defeasible evidence $\Rightarrow Inno(S)$ by the hard evidence $Inno(S)$. Hence N_1 should be "stronger" than N'_1. Therefore if N'_1 attacks A_2, we expect N_1 to also attack A_2, which is not the case according to the MP-attack relations based on the democratic order in [34].

MP-semantics based on the elitist ordering as well as the approaches of Brewka and Eiter [10] and Delgrande, Schaub, Tompits, Wang [15,44] all provide the expected conclusion in this case. □

Brewka, Niemelä and Truszczynski [11] have persuasively argued that the intuition of defeasible reasoning is about finding *(justified) belief sets that give the most accurate picture of reality assuming the world is as normal as possible* (see also Reiter, Reiter and Criscuolo [41,42] and Delgrande [14]). Stable belief sets could be viewed as providing such pictures of the reality where the beliefs are supported and defended by arguments that are grounded in the undisputed facts about the reality and based on the world's "normal patterns" represented by defeasible rules. The assumption that the world is "as normal as possible" is realized by taking into account all possible arguments that could be built based on the given facts and normal patterns of the world. The grounding of considered arguments in the undisputed facts ensures that the accepted beliefs are grounded in the reality. Therefore the more facts we have about the real world, the more pro and con arguments we have about the world and hence the more accurate pictures we have about the reality.

Suppose all beliefs in a stable belief set S of a knowledge base K indeed represent facts about the real world. Therefore the updated knowledge base $K' = K + \Omega$ obtained by adding the facts in a subset $\Omega \subseteq S$ to K, contains more facts about the real world and hence should provide a more accurate picture of the reality. Since S is an accurate picture of the reality and K' should provide a more accurate picture of the reality then K, it should be intuitive and sensible to expect that S is also a belief set of the updated knowledge base K'.

We refer to the property stating that expanding the base of undisputed facts and observation of a knowledge base K by new facts that belong to a stable belief set S of K results in an updated knowledge base of which S is still a stable belief set as the property of credulous cumulativity. The following Example 1.3 shows that credulous cumulativity helps shed useful insights into the different semantics of structured argumentation.

Example 1.3. Consider the knowledge base K in Example 1.1 where it is known that Dean holds.

According to the MP-last link and democratic ordering semantics,[7] $\{D, A, P, T\}$[8] is the unique stable belief set of K while $\{D, A, P, \neg T\}$ is the unique stable belief set of K according to the MP-weakest link and elitist ordering.

Suppose that we also know that the dean is a professor and updating the knowledge base with this fact results in $K' = K + \{P\}$. According to the property of credulous cumulativity, $\{D, A, P, T\}$ (resp. $\{D, A, P, \neg T\}$) should be a stable belief set of K' wrt the MP-last link and democratic ordering semantics (resp. MP-weakest link and elitist ordering semantics). As we will see later, $\{D, A, P, T\}$ is indeed a stable belief set of K' wrt the MP-last link and democratic ordering semantics while $\{D, A, P, \neg T\}$ is not a stable belief set of K' wrt MP-weakest link and elitist ordering.

We will show in section 7.4 that credulous cumulativity is satisfied wrt MP-semantics (as studied in [34]) based on democratic ordering and also the non-argument-based approach of Brewka and Eiter while it is not satisfied by the MP-semantics based on the elitist ordering. □

Rule preferences have been studied extensively in logic programming [8,10,15], which is arguably the conceptually closest framework to structured argumentation. As the intuition underlining the credulous cumulativity property is conceptually not specifically bound to argument-based approaches, it is natural to ask whether this property is also satisfied by the semantics of logic programming with and without rule preferences. In sections 7.1, 7.2, 7.3 we show that credulous cumulativity is indeed embraced, although implicitly, by the two most well-known and well-studied approaches to preference handling in logic programming: the approach advocated by Delgrande, Schaub, Tompits, Wang and others [15,44] where preferences are viewed as a prescription of the application order of rules and the Brewka and Eiter approach [10], which is based on two principles referred to as BE-principles that preference semantics should follow. We study the relation between structured argumentation and logic programming with rule preferences in two different directions: In sections 7.1, 7.2, we demonstrate that the conceptual ideas and principles underlining the logic programming approaches could be naturally incorporated into our technical framework of structured argumentation. In the other direction, we show directly within the technical contexts of logic programming with rule preferences (section 7.3) that the credulous cumulativity property is satisfied.

There are distinct approaches to study the semantics of rule preferences. Some are based on specificity while others could be based on social values, laws or just common customs [6,18,15,11,10]. Though the underlining intuitions for the introduction of rule preferences could be different, they all share a basic and natural interpretation of the preference of a defeasible rule d over another rule d', that in a situation when each of rules d, d' is applicable (i.e. the premises of both rules follow from the factual evidence and the strict rules in the knowledge base), but both could not be applied together, then d should be applied. We adopt this basic and natural view of rule preference and capture it by the property of effective rebuts stating simply that a defeasible argument containing exactly one rule attacks another defeasible single-rule argument only if the rule of the former is not less preferred than the one of the latter.

A key contribution of this paper is the introduction and study of a set of properties for analyzing and evaluating the semantics of structured argumentation. The properties could be divided into three groups. The first group consists of the properties of credulous cumulativity and attack monotonicity together with the property of context-independence. The second one consists of two pretty simple properties: one is the property of subargument structure stating that any attack against a subargument is also an attack against the entire argument, and the other is the property of attack closure stating intuitively that attacks are either based on undercuts (and hence preference-independent) or based on contradicting arguments (and hence preference-dependent if the preference relation is not empty). These two groups of axioms together determine the semantics of knowledge bases without preferences between rules. The semantics of rule preferences are dealt with by two new properties. One is the property of effective rebuts and the other is the link-oriented property expressing the intuition that attacks are directed towards identifying "culprit link" within arguments.

There is some tension between the properties of credulous cumulativity and attack monotonicity. When some beliefs justified by some arguments in a stable extension are confirmed in the reality and hence put into the base of evidence, many arguments may get stronger due to the attack monotonicity property.

If the arguments belonging to the stable extension get stronger, they will strengthen the defense of the stable extension and hence there is no counter-effect on the credulous cumulativity property. But if an argument not belonging to the stable extension gets stronger, its attacking power also grows. The question is whether such stronger arguments could threaten the stability of the stable extension. The answer is "no" if the attack relation assignment satisfies the link-oriented property stating intuitively that attacks are directed against "culprit links" within arguments.

If the link-oriented property is not satisfied then the growing power of the strengthened arguments could destroy the credulous cumulativity property. For an illustration, consider the arguments (see Figs. 3, 4) wrt knowledge bases K, K' in Examples 1.1, 1.3. A_1 is a subargument of A_2 and A'_2 is obtained from A_2 by replacing the subargument A_1 with its conclusion. It is clear that A_3 does not attack A_1. Suppose A_3 attacks A_2.[9] Because A_3 does not attack A_1,[10] the link-oriented property states that A_3 should attack A'_2.

[7] We refer to the attack relations and associated semantics studied by Modgil and Prakken in [34] wrt the ASPIC+ framework as MP-attack relations.

[8] D, P, T, A stand for Dean, Professor, Teach and Administrator respectively.

[9] Hence the attack is directed against some link in A_2.

[10] Hence the link in A_1 is not the "culprit" causing the attack from A_3 against A_2 implying that the "culprit link" is in A'_2.

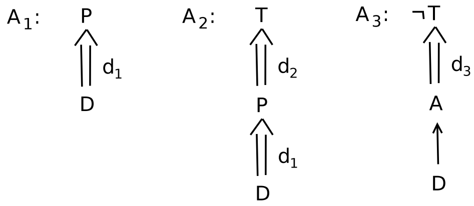

Fig. 3. Arguments of Example 1.1.

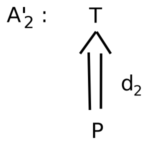

Fig. 4. Extra arguments for Example 1.3.

It turns out that wrt the MP-weakest link principle and the elitist ordering, A_3 attacks A_2 but A_3 attacks neither A_1 nor A'_2. Therefore the link-oriented property is violated. Not surprisingly, the credulous cumulativity is violated as we have discussed in Example 1.3.

The presented properties in this paper could be viewed as an axiom system for evaluating, understanding and comparing existing approaches to structured argumentation systems with preferences. They could also serve as guidelines when defining new argumentation systems. The stable semantics of argument systems obeying such axioms is characterized by a normal form such that the stable extensions of any argument system obeying the axiom system are also stable extensions wrt the normal form which is itself also obeying the axioms.

Another important contribution of this paper is the application of the presented axiom system to analyze and evaluate prominent approaches to defeasible reasoning with preferences. We especially analyze and evaluate the four prominent attack relations proposed by Modgil and Prakken for the ASPIC+ framework. As illustrated in Example 1.2, the two attack relations based on democratic ordering do not satisfy the attack monotonicity which could lead to counter-intuitive conclusions. We will show later in section 7.4 that the other ones based on the elitist ordering do not satisfy consistency postulate (and hence the credulous cumulativity property). These insights suggest that a revision of prominent MP-attack relations together with much more research may be needed to understand the semantics of ASPIC+. The axiom system studied in this paper could play a significant role here.

The paper also makes a relevant contribution in the exploration of the relations between argument-based and non-argument-based approaches to defeasible reasoning with priorities by showing that the semantics proposed by Delgrande, Schaub and Tompits [15] and Brewka and Eiter [10], arguably the most prominent ones in logic programming with priorities, satisfy the credulous cumulativity property. To our knowledge, this is the first time a deep insight between these two prominent approaches to reasoning with priorities has been gained. This insight is even more relevant from the point of view of a user who is interested in the application of defeasible reasoning with priorities in reality as logic programming offers arguably one of the most sophisticated industrial-strength development environments for reasoning systems. Such a user may be well-advised to use logic programming as her/his development base (saving very valuable resources for the implementation of new systems for reasoning with preferences) if she/he knows that there is a significant convergence between the argument-based and non-argument-based approaches. Of course more work needs to be done to reach this stage of knowledge and we believe that this paper provides a significant first step for further research in this direction.

The paper is organized as follows. We recall quickly the basic notions of abstract argumentation in section 2. In the following section, key notions of knowledge bases, rule preferences, arguments and sensible classes of knowledge bases are recalled from the literature or introduced. In section 4, we introduce a novel notion of attack relation assignments together with the basic properties for knowledge bases with or without rule preferences. We then introduce in section 5 a basic attack relation assignment and show that it represents a "normal form" of attack relation assignments satisfying the basic properties and that each stable extension wrt any attack relation assignments satisfying the basic properties is also a stable extension wrt the basic attack relation assignment. In section 6, we introduce two new properties for dealing with rule preferences. We define a class of ordinary attack relation assignments as those satisfying the basic properties together with the new properties for handling rule preferences. We then introduce a normal attack relation assignment and show that it represents a "normal-form" of ordinary attack relation assignments in the sense that it is itself an ordinary attack relation assignment and each stable extension wrt any ordinary attack relation assignments is also a stable extension wrt the normal attack relation assignment. We study in section 7 the relationship between our approach and related approaches in the literature. We first show that the semantics, based on the operational reading of the preference relation between rules as a specification of their application order [15], could be captured in our framework. We then show that the principles underlining Brewka and Eiter's semantics of prioritized logic programming [10] are also satisfied by the ordinary semantics in our framework. We also show that extended logic programming with and without preferences satisfies the property of credulous cumulativity. We end this section by an in depth study of the attack relations of Modgil and Prakken [34]. We discuss possible future works and conclude in section 8.

This paper is both a follow-up and an extension of our paper in [17]. It is an extension as it provides an expanded presentation and proofs of the properties presented in [17]. It is a follow-up as it offers significant new results, especially new properties of attack closure and link-orientation and new theorems showing that the normal attack relation assignment represents indeed a normal form of ordinary attack relation assignments and the credulous cumulativity also holds for complete extension semantics. We also provide an extensive study of related literature.

2. Preliminaries

An abstract argumentation framework [16] is defined simply as a pair (AR, att) where AR is a set of arguments and $att \subseteq AR \times AR$. $(A, B) \in att$ means that A *attacks* B. A set of argument S attacks (or is attacked by) an argument A (or a set of arguments R) if some argument in S attacks (or is attacked by) A (or some argument in R); S is *conflict-free* if it does not attack itself. S is *conflicting* if S is not conflict-free. A set of arguments S *defends* an argument A if S attacks each attack against A. S is *admissible* if S is conflict-free and defends each argument in it. The semantics of abstract argumentation is defined by various notions of extensions. A *complete extension* is an admissible set of arguments containing each argument it defends. A *stable extension* is a conflict-free set of arguments that attacks every argument not belonging to it. It is well-known that stable extensions are complete but not vice versa.

3. Defeasible knowledge bases

We assume a non-empty set $\mathcal{L}$ of ground atoms and their classical negations. An atom is also called a positive literal while a negative literal is the negation of a positive literal. A set of literals is said to be **contradictory** if it contains a pair $a, \neg a$ of an atom a and its negation $\neg a$.

Atoms in $\mathcal{L}$ are distinguished between *domain atoms* representing propositions about the concerned domains and *non-domain atoms* of the form ab_d representing the non-applicability of defeasible rules d.[11]

Following ASPIC+ [34,35], Garcia and Simari [21], Nute [36], Nute and Lewis [37], Loui [31], Pollock [38], Gelfond and Son [24] and Vreeswijk [46], we distinguish between strict and defeasible rules.

Definition 3.1.

1. A defeasible rule is of the form

$$b_1, \ldots, b_n \Rightarrow h$$

where $b_1, \ldots, b_n, h$ are domain literals.
2. A strict rule is of the form

$$b_1, \ldots, b_n \rightarrow h$$

where
- $b_1, \ldots, b_n$ are domain literals, and
- h could be a domain literal or a non-domain atom of the form ab_d where d is a defeasible rule.[12,13] □

Notation 3.1.

1. For a rule r of the form $b_1, \ldots, b_n \rightarrow / \Rightarrow h$, the set $\{b_1, \ldots, b_n\}$ (resp. the literal h) is referred to as the body (resp. head) of r, denoted by bd(r) (resp. hd(r)).
2. For a set of rules R, denote hd(R) = $\{hd(r) \mid r \in R\}$. □

Definition 3.2. A rule-based system is defined as a triple

$$RBS = (RS, RD, \preceq)$$

where following conditions are satisfied:

1. RS is a set of strict rules.

[11] Intuitively, the atom ab_d represents an "abnormal situation" for defeasible rule d where d should not be applied even if its premises hold. This common use of ab_d is borrowed from logic programming [5].

[12] A strict rule $b_1, \ldots, b_n \rightarrow ab_d$ states that defeasible rule d must not be applied when $b_1, \ldots, b_n$ hold.

[13] We thank an anonymous referee of our ECAI 2014 paper [17] for suggesting the elegant notation ab_d.

2. RD is a finite set of defeasible rules.
3. $\preceq$ is a binary preference relation over RD that is transitive.
4. If RS contains a rule of the form $bd \to ab_d$ then $d \in RD$. □

We write $d \prec d'$ iff $d \preceq d'$ and $d' \npreceq d$.

A **base of evidence** BE is a set of ground domain literals representing unchallenged observations, facts etc.

Definition 3.3. A **knowledge base** is a pair (RBS, BE) of a rule-based system $RBS = (RS, RD, \preceq)$ and a base of evidence BE.

A knowledge base is said to be **basic**[14] if its preference relation is empty, i.e. $\preceq = \emptyset$. □

For an illustration, consider the knowledge base KB_1 in Example 1.2 where $\{r_1, r_2, r_3\}$ is the set of strict rules, $\{d_1, d_2, d\}$ is the set of defeasible rules, $\{Inno(S)\}$ is the base of evidence and $d_1 \prec d_2$ represents the preference relation.

Knowledge bases whose preference relations are preorders[15] are studied extensively within the ASPIC+ framework [34,35] while basic knowledge bases are studied in [12].[16]

Remark 3.1. For convenience, we often write $K = (RS, RD, \preceq, BE)$ instead of $K = (RBS, BE)$ with $RBS = (RS, RD, \preceq)$.

We recall below the key notion of arguments from [46,12,34,35].

Definition 3.4 *(Arguments).* Let $K = (RBS, BE)$ be a knowledge base. An **argument** wrt K is defined inductively as follows:

1. For each $\alpha \in BE$, $[\alpha]$ is an argument with conclusion α.[17]
2. Let r be a rule of the forms $\alpha_1, \ldots, \alpha_n \to / \Rightarrow \alpha$, $n \geq 0$, from RBS. Further suppose that $A_1, \ldots, A_n$ are arguments with conclusions α_i, $1 \leq i \leq n$, respectively. Then

$$A = [A_1, \ldots, A_n \to / \Rightarrow \alpha] \quad \text{(also denoted by } A = [A_1, \ldots, A_n, r]\text{)}$$

is an argument with conclusion α and last rule r denoted by cnl(A) and last(A) respectively.
3. Each argument wrt K is obtained by applying the above steps 1, 2 finitely many times. □

For an illustration, consider the arguments in Example 1.3 where $A_1 = [[D] \Rightarrow P] = [[D], d_1]$, $A_2 = [A_1 \Rightarrow T] = [A_1, d_2]$.

In Example 1.2, the formal representation of A_1, N_1 are as follows: $A_1 = [\Rightarrow Inno(P_1)] = [d_1]$, $N_1 = [[Inno(S)], A_1 \to \neg Inno(P_2)] = [[Inno(S)], A_1, r_1]$.

Notation 3.2.

1. The set of all arguments wrt a knowledge base K is denoted by AR_K. The set of the conclusions of arguments in a set $S \subseteq AR_K$ is denoted by $cnl(S)$.
2. A **strict** argument is an argument containing no defeasible rule. An argument is **defeasible** iff it is not strict. The set of defeasible rules appearing in an argument A is denoted by $dr(A)$.
A defeasible argument A is called **basic defeasible** iff $last(A)$ is defeasible.
3. An argument B is a **subargument** of an argument A iff $B = A$ or $A = [A_1, \ldots, A_n, r]$ and B is a subargument of some A_i.
B is a **proper subargument** of A if B is a subargument of A and $B \neq A$. □

For any set of literals $X \subseteq \mathcal{L}$, X_{dom} denotes the set of all domain literals in X. Let $K = (RBS, BE)$ with $RBS = (RS, RD, \preceq)$.

Definition 3.5. The **closure** of a set of literals $X \subseteq \mathcal{L}$ wrt the set of strict rules RS, denoted by $CN_{RS}(X)$, is the union of X and the set of conclusions of all strict arguments wrt knowledge base (RBS, X_{dom}).[18]

X is said to be **closed** iff $X = CN_{RS}(X)$. X is said to be **inconsistent** iff its closure $CN_{RS}(X)$ is contradictory.[19] X is **consistent** iff it is not inconsistent. □

---

[14] Also called preference-free in [17].

[15] A preorder is a reflexive and transitive relation [13].

[16] Note that preorders are reflexive, i.e. they always contain $d \preceq d$ for any $d \in RD$. Therefore if RD is not empty, the preorder $\preceq$ is not empty. It follows that the preference relations of basic knowledge bases with non-empty set of defeasible rules are not preorders.

[17] Note that only domain literals are contained in BE. Hence there are no arguments of the form $[ab_d]$ or $[\neg ab_d]$ wrt K.

[18] X_{dom} acts as a base of evidence. It is straightforward to see that if X contains only domain literals, then $CN_{RS}(X)$ is the set of conclusions of all strict arguments wrt (RBS, X). In general, $CN_{RS}(X) = X \cup CN_{RS}(X_{dom})$. Further it is not difficult to see that for any $X \subseteq \mathcal{L}$, $CN_{RS}(X) = CN_{RS}(CN_{RS}(X))$ (a simple proof is given in Appendix K).

[19] Note that a set X could be inconsistent but not contradictory. For example, let $RS = \{a \to \neg b\}$. The set $X = \{a, b\}$ is inconsistent as its closure $CN_{RS}(X) = \{a, b, \neg b\}$ is contradictory though X itself is not contradictory.

Notation 3.3. We often write $CN(X)$ or $CN_K(X)$ for $CN_{RS}(X)$ if there are no possibilities for misunderstanding. We also often write $X \vdash l$ (or $X \vdash_K l$) iff $l \in CN(X)$. □

In Example 1.2, $RS = \{r_1, r_2, r_3\}$. The closure of $X = \{Inno(P_1), Inno(S)\}$ wrt RS is $CN_{RS}(X) = \{Inno(P_1), Inno(S), \neg Inno(P_2)\}$. It is clear that for $Y = \{Inno(P_1), Inno(S), Inno(P_2)\}$, the closure of Y, $CN_{RS}(Y) = \{Inno(P_1), Inno(S), Inno(P_2), \neg Inno(P_2)\}$ is contradictory. Therefore Y is inconsistent though Y is not contradictory.

Definition 3.6. Let $K = (RS, RD, \preceq, BE)$ be a knowledge base.

1. K is said to be consistent iff its base of evidence is consistent wrt its set of strict rules.[20]
2. K is said to be closed under transposition iff for each strict rule of the form $b_1, \ldots, b_n \to h$ in K s.t. h is a domain literal, all the rules of the forms $b_1, \ldots, b_{i-1}, \neg h, b_{i+1}, \ldots, b_n \to \neg b_i$, $1 \le i \le n$, also belong to K.
3. K is said to be closed under contraposition iff for each set of domain literals S, each domain literal λ, if $S \vdash_K \lambda$ then for each $\sigma \in S$, $S \setminus \{\sigma\} \cup \{\neg\lambda\} \vdash_K \neg\sigma$.
4. K is said to satisfy the self-contradiction property iff for each minimal inconsistent set of domain literals $X \subseteq \mathcal{L}$, for each $x \in X$, it holds: $X \vdash_K \neg x$. □

The properties of closure under transposition or contraposition are introduced in [12,39,34] while the self-contradiction property together with the following lemma can be found in [19].

Lemma 3.1. *If K is closed under transposition or contraposition then K satisfies the self-contradiction property.*

Proof. To keep the paper self-contained, we recall the proof in Appendix A. □

The classes of knowledge bases closed under transposition or contraposition or satisfying the self-contradiction property are prominent examples of sensible classes of knowledge bases characterized by their rule-based systems.

Definition 3.7 *(Sensible classes of knowledge bases).* A class $\mathcal{K}$ of knowledge bases is said to be **sensible** iff following conditions are satisfied:

1. $\mathcal{K}$ is not empty.
2. Every knowledge base in $\mathcal{K}$ is consistent.
3. If a knowledge base $K = (RBS, BE)$ belongs to $\mathcal{K}$ then all consistent knowledge bases of the form (RBS, BE') also belong to $\mathcal{K}$. □

For an illustration, several sensible classes of knowledge bases are given in the following example.

Example 3.1.

1. The class of all consistent basic knowledge bases closed under transposition is sensible. All three conditions in Definition 3.7 are obviously satisfied.
   Similarly, the class of all consistent knowledge bases satisfying the self-contradiction property is also sensible.
2. Another example is the class $\mathcal{K} = \{\, (RS, RD, \preceq, BE) \mid BE \text{ is not contradictory}\,\}$ where $RS = \emptyset$, $\preceq = \emptyset$ and $RD = \{d_0, d_1\}$ with $d_0 :\Rightarrow a$ and $d_1 :\Rightarrow \neg a$.
   It is easy to verify that $\mathcal{K}$ is sensible. □

We conclude this section with another helpful notation on argument structure.

Notation 3.4.

1. The basic defeasible subarguments of an argument A that are not proper subarguments of other basic defeasible subarguments of A play a prominent role in our later exposition and are referred to as maximal basic defeasible subarguments of A and formally defined by:

$$mbd(A) = \begin{cases} \emptyset & \text{if } A = [\alpha] \text{ for } \alpha \in BE \\ \{A\} & \text{if A is basic defeasible} \\ mbd(A_1) \cup \ldots \cup mbd(A_n) & \text{if } A = [A_1, \ldots, A_n, r] \text{ and r is strict} \end{cases}$$

[20] I.e. the set $CN_{RS}(BE)$ is not contradictory.

2. For an argument A, the set of defeasible rules appearing last in A is defined by:

$$ldr(A) = \{ last(X) \mid X \in mbd(A) \} \qquad \square$$

For an illustration, consider again arguments in Figs. 1, 2. It is not difficult to see $mbd(A_1) = mbd(N_1) = \{A_1\}$, $mbd(N'_1) = \{A_1, B\}$ where $B = [\Rightarrow Inno(S)]$. Therefore $ldr(A_1) = ldr(N_1) = \{d_1\}$, $ldr(N'_1) = \{d_1, d\}$.

Maximal basic defeasible sub-arguments are special cases of the maximal fallible sub-arguments in [34]. While the later are defined for any ASPIC+ knowledge bases, the formers are defined only for the knowledge bases considered in this paper.

4. Basic properties of attack relation assignments

We introduce in this section the basic properties of attack relation assignments.

The semantics of structured defeasible argumentation systems are determined by appropriate attack relations where strict arguments are not attacked by any other arguments as they are considered reflecting the reality and hence beyond any doubt.

Definition 4.1 *(Attack relation assignment).* An attack relation assignment att defined for a sensible class $\mathcal{K}$ of knowledge bases is a function assigning to each knowledge base $K \in \mathcal{K}$ an attack relation $att(K) \subseteq AR_K \times AR_K$ such that there is no attack against strict arguments, i.e. for each strict argument $B \in AR_K$, there is no argument $A \in AR_K$ such that $(A, B) \in att(K)$.

For convenience, we often say A attacks B wrt att(K) for $(A, B) \in att(K)$. $\square$

Notation 4.1.

- From now on, whenever we refer to a knowledge base K without any specific information, we mean one of the form $K = (RBS, BE)$ with $RBS = (RS, RD, \preceq)$.
- For any finite set Ω of domain literals, define $K + \Omega = (RBS, BE \cup \Omega)$. $\square$

Definition 4.2 *(Belief sets).* A set $S \subseteq \mathcal{L}$ is said to be a stable (resp. complete) belief set of knowledge base K wrt an attack relation assignment att iff $att(K)$ is defined and there is a stable (resp complete) extension E of $(AR_K, att(K))$ such that $S = cnl(E)$. $\square$

4.1. Credulous cumulativity

We begin with the introduction of the formal definition of the credulous cumulativity property.

Definition 4.3 *(Credulous cumulativity).* Let $\mathcal{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathcal{K}$. We say att satisfies the property of credulous cumulativity for $\mathcal{K}$ if and only if for each $K \in \mathcal{K}$, for each stable belief set S of K wrt att and for each finite subset $\Omega \subseteq S$ of domain literals,

1. $K + \Omega$ is a consistent knowledge base (i.e. $K + \Omega$ belongs to $\mathcal{K}$), and
2. S is a stable belief set of $K + \Omega$ wrt att. $\square$

For an illustration, consider the knowledge base K in Example 1.3. It is obvious that $K' = K + \{P\}$ is consistent and hence belongs to any sensible class of knowledge bases to which K belongs. Let att be an attack relation assignment where $att(K) = \{(A_2, A_3)\}$ and $att(K') = \{(A_2, A_3), (A'_2, A_3)\}$.[21] It is easy to see that $S = \{D, A, P, T\}$ is a stable belief set of K wrt $att(K)$. It is also straightforward to see that S is a stable belief set of K' wrt $att(K')$.

There are different versions of the property of credulous cumulativity according to different types of belief sets. In this paper, we focus on stable semantics to facilitate the comparison of argument-based and non-argument-based prioritized default reasoning. We discuss the credulous cumulativity property wrt complete extension semantics in section 8.

Notation 4.2. Slightly abusing notation for convenience, we often simply say that

an attack relation assignment att satisfies some property P (like the property of credulous cumulativity) for a sensible class of knowledge bases $\mathcal{K}$

assuming implicitly that att is defined for $\mathcal{K}$. $\square$

[21] Note that att(K) is different to the attack relation assignment based on the weakest link and the elitist ordering for K, K' in Example 1.3.

The following lemma shows that credulous cumulativity implies Caminada and Amgoud's postulates of consistency and closure.

Lemma 4.1. *Suppose an attack relation assignment att satisfies the credulous cumulativity property for a sensible class $\mathcal{K}$ of knowledge bases. Then for each $K \in \mathcal{K}$, the stable belief sets of K (wrt att) are both consistent and closed.*

Proof. Let $K \in \mathcal{K}$ and S be a stable belief set of K wrt att. It is clear that there are no negative literals of the form $\neg ab_d$ in S.

Suppose S is inconsistent. Therefore, there is a finite subset $\Omega \subseteq S$ of domain literals such that Ω is inconsistent wrt set of strict rules of K. Hence $K + \Omega$ is not a consistent knowledge base, contradicting the first condition in the definition of credulous cumulativity axiom.

Let $S \vdash_K \alpha$ for a literal α. We need to show $\alpha \in S$. If $\alpha \in S$, there is nothing to prove. Suppose $\alpha \notin S$. Therefore there is a finite set $\Omega \subseteq S$ of domain literals such that $\Omega \vdash_K \alpha$. Therefore there is a strict argument A of $K' = K + \Omega$ such that $cnl(A) = \alpha$. As there is no attack against A wrt $att(K')$, A belongs to any stable extension of $(AR_{K'}, att(K'))$. Hence α belongs to any stable belief set of K' wrt att. As S is also a stable belief set of K', $\alpha \in S$. Contradiction. Hence the case $\alpha \notin S$ cannot occur. □

4.2. Attack monotonicity

We proceed further with the introduction of the property of attack monotonicity stating that when some piece of defeasible information on which an argument is based is confirmed by unchallenged observations, the argument is strengthened in the sense that whatever is attacked by the original argument should also be attacked by the strengthened one, and whatever attacks the strengthened one, attacks the original one. In other words, the more hard evidence your arguments are based on, the stronger your arguments become.

Let $A \in AR_K$ and $\Omega \subseteq BE$ be a finite set of literals. The strengthening of A wrt Ω denoted by $A \uparrow \Omega$, is the set of arguments obtained by replacing zero, one or more subarguments of A by their conclusions provided that these conclusions belong to Ω.

Definition 4.4 *(Strengthening operation).* Let $A \in AR_K$ and $\Omega \subseteq BE$ be a finite set of domain literals. The **strengthening** of A wrt Ω denoted by $A \uparrow \Omega$ is defined inductively as follows:

$$A \uparrow \Omega = \begin{cases} \{[\alpha]\} & \text{if } A = [\alpha] \text{ and } \alpha \in BE \\ AS \cup \{[hd(r)]\} & \text{if } A = [A_1, \ldots, A_n, r] \text{ and } hd(r) \in \Omega \\ AS & \text{if } A = [A_1, \ldots, A_n, r] \text{ and } hd(r) \notin \Omega \end{cases}$$

where $AS = \{[X_1, \ldots, X_n, r] \mid \forall i : X_i \in A_i \uparrow \Omega\}$. □

For illustration, in Example 1.3,

$$[D] \uparrow \{P\} = \{[D]\}, \quad A_1 \uparrow \{P\} = \{[P], A_1\}, \quad A_2 \uparrow \{P\} = \{A_2, A_2'\}.$$

In Example 1.2, $N_1' \uparrow \{Inno(S)\} = \{N_1', N_1\}$.

Lemma 4.2. *Let K be a knowledge base, $A \in AR_K$ and $\Omega \subseteq BE$ be finite. The following assertion holds:*

X' is a subargument of an argument $X \in A \uparrow \Omega$ iff there exists a sub-argument A' of A such that $X' \in A' \uparrow \Omega$.

Proof. See Appendix C. □

For an illustration of Lemma 4.2, consider a basic knowledge base where $BE = \{a\}$ and $RS = \emptyset$ and RD consist of following defeasible rules

$$d_1: \ \Rightarrow a \quad d_2: \ a \Rightarrow b \quad d_3: \ b \Rightarrow c$$

Consider the arguments A, A', X, X' in Fig. 5 and $\Omega = \{a\}$. It is clear that A', X' are subarguments of A, X respectively and $X \in A \uparrow \{a\}$ and $X' \in A' \uparrow \{a\}$.

Definition 4.5 *(Attack monotonicity).* Let att be an attack relation assignment defined for a sensible class $\mathcal{K}$ of knowledge bases. We say att satisfies the property of attack monotonicity for $\mathcal{K}$ iff for each knowledge base $K \in \mathcal{K}$, for each finite subset $\Omega \subseteq BE$, for all $A, B \in AR_K$ and for each $X \in A \uparrow \Omega$, the following assertions hold:

1. If $(A, B) \in att(K)$ then $(X, B) \in att(K)$.
2. If $(B, X) \in att(K)$ then $(B, A) \in att(K)$. □

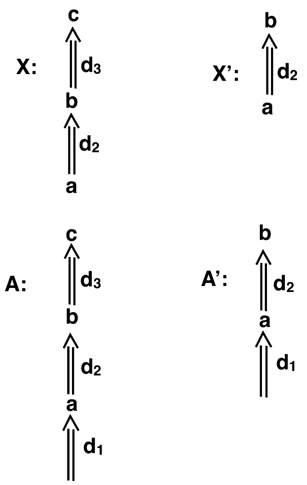

Fig. 5. Strengthening-subargument.

As we have discussed in the introduction, a natural property of knowledge bases inherently related to the property of attack monotonicity is the property of irrelevance of redundant defeasible rules. We introduce below the formal definition of this property and a theorem showing that it follows from the property of attack monotonicity.

Notation 4.3. For any defeasible rule d, denote

$$K + d = (RS, RD \cup \{d\}, \preceq, BE)$$

where $K = (RS, RD, \preceq, BE)$. □

For ease of reference, for any evidence $\omega \in BE$, we denote the default "$\Rightarrow \omega$" by d_ω.

Definition 4.6 *(Irrelevance of redundant defaults).* Let $\mathcal{K}$ be a sensible class of knowledge bases such that for each $K = (RSB, BE) \in \mathcal{K}$, for each evidence $\omega \in BE$, $K + d_\omega$ belongs to $\mathcal{K}$. Further let att be an attack relation assignment defined for $\mathcal{K}$.

We say the attack relation assignment att satisfies the property of irrelevance of redundant defaults for $\mathcal{K}$ iff for each knowledge base $K = (RSB, BE) \in \mathcal{K}$, for each evidence $\omega \in BE$:

1. the stable belief sets of K and $K + d_\omega$ coincide, and
2. the complete belief sets of $K, K + d_\omega$ coincide. □

The property of irrelevance of redundant defaults follows from the attack monotonicity property if the attack relations satisfy a general and natural condition of context-independence.

Definition 4.7 *(Context-independence).* Let att be an attack relation assignment defined for a sensible class $\mathcal{K}$ of knowledge bases. We say att satisfies the property of context-independence for $\mathcal{K}$ iff for any two arbitrary knowledge bases K, K' with preference relations $\preceq, \preceq'$ respectively and for any two arguments A, B belonging to $AR_K \cap AR_{K'}$ such that the restrictions of $\preceq$ and $\preceq'$ on $dr(A) \cup dr(B)$ coincide,[22] it holds that

$$(A, B) \in att(K) \quad \text{iff} \quad (A, B) \in att(K') \qquad \square$$

For an illustration of the context-independence property, consider arguments A_2, A_3 in Example 1.3. Both A_2, A_3 are arguments wrt K and K' and the preference relations in both knowledge bases are identical. Context-independence requires that the attack relation between A_2, A_3 should also be the same wrt K, K'. It is indeed the case for all attack relations defined in ASPIC+ as well as the basic or normal attack relations defined in Definitions 5.2, 6.6.

Theorem 4.1. *Let $\mathcal{K}$ be a sensible class of knowledge bases such that for each $K = (RSB, BE) \in \mathcal{K}$, for each evidence $\omega \in BE$, $K + d_\omega$ belongs to $\mathcal{K}$. Further let* att *be an attack relation assignment satisfying the properties of attack monotonicity and context-independence for $\mathcal{K}$. Then att also satisfies the property of irrelevance of redundant defaults for $\mathcal{K}$.*

[22] I.e. $\preceq \cap (D \times D) = \preceq' \cap (D \times D)$ for $D = dr(A) \cup dr(B)$.

Proof. See Appendix C. □

Example 1.2 could be viewed as an illustration of Theorem 4.1 (in a contrapositive way).

4.3. Further properties on attack structure

We introduce in this section two simple and natural properties on the structure of attack relations. We begin with the property of sub-argument structure.

Definition 4.8 *(Subargument structure).* Let $\mathcal{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathcal{K}$. Then att is said to satisfy the property of subargument structure for $\mathcal{K}$ iff for each knowledge base $K \in \mathcal{K}$, for all $A, B \in AR_K$, it holds that

$$(A, B) \in att(K) \text{ iff there is a defeasible subargument } B' \text{ of } B \text{ such that } (A, B') \in att(K) \qquad \square$$

A key consequence of the property of subargument structure is the property of subargument closure.

Lemma 4.3. *Let $\mathcal{K}$ be a sensible class of knowledge bases and* att *be an attack relation assignment satisfying the property of subargument structure for $\mathcal{K}$. Further let $K \in \mathcal{K}$. Then each complete extension of $(AR_K, att(K))$ contains all subarguments of its arguments.*

Proof. Let E be a complete extension of $(AR_K, att(K))$. Further let $A \in E$ and X be a subargument of A. If X is strict then there is no attack against X. Hence $X \in E$.

Suppose X is defeasible. From the sub-argument structure axiom, each attack against X is an attack against A. Hence each attack against X is counter-attacked by E. Therefore $X \in E$. □

Lemma 4.3 is not new. Martinez, Garcia and Simari [33] include it as a key component in their study of abstract argumentation. Within the ASPIC+ framework, it is known as Theorem 12 [34]. It is proposed as a postulate in Amgoud [1].

One way for arguments to attack each other is by undercutting [38,12,34,35] recalled below.

Definition 4.9. Let $A, B \in AR_K$ for knowledge base K.

We say A **undercuts** B (at B') iff B' is a basic defeasible subargument of B and the conclusion of A is $\mathsf{ab}_{\mathsf{last}(B')}$.

We say A **directly undercuts** B iff A undercuts B (at B). □

Contradiction between arguments represents another key source of conflicts among arguments.

We say A **contradicts** B (at B') iff B' is a sub-argument of B and the conclusions of A and B' are contradictory.

For an illustration, in Fig. 1, N_1 contradicts N_2 (at A_2) and N_2 contradicts N_1 (at A_1).

Remark 4.1. If A contradicts B (at B') then $cnl(A) = \neg cnl(B')$ and both $cnl(A)$ and $cnl(B')$ are domain literals because the base of evidence contains only domain literals and the heads of rules in the knowledge base are either domain literals or non-domain atoms.

Within our framework, if two arguments do not undercut or contradict each other then they do not have any reason to attack each other. We capture this intuition by the property of attack closure introduced below.

Definition 4.10 *(Attack closure).* Let $\mathcal{K}$ be a sensible class of knowledge bases and att be an attack relation assignment for $\mathcal{K}$.

The attack relation assignment att is said to satisfy the property of attack closure for $\mathcal{K}$ iff for each knowledge base $K \in \mathcal{K}$, for all $A, B \in AR_K$, the following conditions hold:

1. If A attacks B wrt $att(K)$ then A undercuts B or A contradicts B.
2. If A undercuts B then A attacks B wrt $att(K)$. □

The scope of attack closure should be expanded when new types of attacks need to be considered. ASPIC+ [34] allows other types of attacks, like contrary-rebut or contrary-undermine attacks. These attacks like undercuts are preference-independent in the sense that they are not affected by the preferences between rules or arguments. As we focus in this paper on understanding the semantics of preferences between rules and hence allow undercuts as the only type of preference-independent attacks, our version of attack closure reflects this focus.[23]

[23] A revised version of the attack closure property to capture other types of attacks would be something like: If A attacks B then A preference-independently attacks B or A contradicts B where A preference-independent attacks B iff A undercuts B or A contrary rebuts B or A contrary-undermines B.

Notation 4.4 *(Basic properties).* The properties of credulous cumulativity, attack monotonicity, context-independence, subargument structure and attack closure are often referred to as *basic properties*. □

5. An axiomatic semantics for basic knowledge bases

We start our exposition with recalling a notion of rebut, a special case of contradicting, from [12,34,38].

Definition 5.1. Let $A, B \in AR_K$ for a knowledge base K.
We say A **rebuts** B (at B') iff B' is a basic defeasible subargument of B and the conclusions of A and B' are contradictory.
We say A **directly rebuts** B iff A rebuts B (at B). □

It is obvious that *if A rebuts B then A also contradicts B, but not vice versa.* For an illustration, arguments N_1, N_1' in Figs. 1, 2 rebut (and hence also contradict) A_2, but A_2 does not rebut any of N_1, N_1' though A_2 contradicts both N_1, N_1'.

Pollock [38] views A as a rebut to B if A contradicts B (at B) and B is defeasible. Caminada and Amgoud [12] introduce a restricted version of rebut where B is required to be basic defeasible. Modgil and Prakken [34] view restricted rebut as an attack in their system though they referred to it simply as rebut. We follow Modgil and Prakken in Definition 5.1.

The following lemma suggests that stable semantics wrt the attack relation assignments satisfying the basic axioms could be captured by those based solely on the notions of rebut and undercut even though the attack closure property states only that attacks should be based on the notions of undercut and contradiction.

Lemma 5.1 *(Characteristic lemma). Let $\mathcal{K}$ be a sensible class of basic knowledge bases and att be an attack relation assignment satisfying all basic properties for $\mathcal{K}$. Further let $K \in \mathcal{K}$ and E be a stable extension of $(AR_K, att(K))$, $B \in AR_K \setminus E$ and $A \in E$ such that $(A, B) \in att(K)$. Then A either undercuts or rebuts B.*

Proof. Follows immediately from Lemma D.1 in Appendix D. □

Remark 5.1. Lemma 5.1 suggests that there is something like a normal form based solely on undercuts and rebuts, among the attack relation assignments satisfying the basic properties. This normal form is called the basic attack relation assignment formally introduced below.

Definition 5.2 *(Basic attack relation assignment).* Let $\mathcal{B}$ be the sensible class of all consistent basic knowledge bases. An attack relation assignment defined for $\mathcal{B}$ is said to be **basic** and denoted by att_{bs} iff it assigns to each basic knowledge base $K \in \mathcal{B}$, an attack relation

$$att_{bs}(K) = Rebut \ \cup \ Undercut$$

where $Rebut = \{ (A, B) \mid A, B \in AR_K : A \text{ rebuts } B \}$ and $Undercut = \{ (A, B) \mid A, B \in AR_K : A \text{ undercuts } B \}$. □

We proceed to show that the basic attack relation assignment att_{bs} is indeed a normal form among the attack relation assignments satisfying the basic properties in the following sense:

- *Stable extensions wrt any attack relation assignment satisfying the basic properties are also stable extensions wrt the basic attack relation assignment.*
- *The basic attack relation assignment att_{bs} satisfies all basic properties.*

The following theorem confirms the first of the above two assertions.

Theorem 5.1. *Let $\mathcal{K}$ be a sensible class of basic knowledge bases and att be an attack relation assignment defined for $\mathcal{K}$.*
It holds that if att satisfies all basic properties for $\mathcal{K}$ then for each $K \in \mathcal{K}$, each stable extension of $(AR_K, att(K))$ is also a stable extension of $(AR_K, att_{bs}(K))$.

Proof. Let E be a stable extension of $(AR_K, att(K))$. Because att satisfies the credulous cumulativity property, $cnl(E)$ is consistent (Lemma 4.1). Since E is conflict-free wrt att(K) and att satisfies the property of attack closure, E is free of undercut-attacks.[24]

We show that E is conflict-free wrt $att_{bs}(K)$.

Suppose E is not conflict-free wrt $att_{bs}(K)$. Therefore there is $A, B \in E$ s.t. $(A, B) \in att_{bs}(K)$, i.e. A either undercuts or rebuts B. Since E is free from undercut-attacks, it follows A rebuts B (at B'). Therefore $cnl(A) = \neg cnl(B')$. From Lemma 4.3, $B' \in E$. Therefore $cnl(E)$ is contradictory. This is impossible since $cnl(E)$ is consistent. Hence E is conflict-free wrt $att_{bs}(K)$.

[24] That means that there are no arguments $X, Y \in E$ s.t. X undercuts Y.

We show that E attacks every argument not belonging to it (wrt $att_{bs}(K)$). Let $B \in AR_K \setminus E$. From Lemma 5.1, it follows that there is $A \in E$ undercutting or rebutting B. Hence $(A, B) \in att_{bs}(K)$.

We have proved that E is a stable extension of $(AR_K, att_{bs}(K))$. □

To show that the basic attack relation assignment satisfies all basic properties, we first prove two relevant lemmas.

Lemma 5.2. *Let $\mathcal{K}$ be a sensible class of basic knowledge bases. Then att_{bs} satisfies the properties of attack monotonicity, context-independence, subargument structure and attack closure for $\mathcal{K}$.*

Proof. It is obvious that att_{bs} satisfies the properties of context-independence, subargument structure and attack closure for $\mathcal{K}$. We only need to show that att_{bs} satisfies the property of attack monotonicity for $\mathcal{K}$.

Let $K \in \mathcal{K}$ and $A, B \in AR_K$. Suppose A attacks B wrt $att_{bs}(K)$ and $X \in A \uparrow \Omega$.

If A undercuts B, it is clear that X also undercuts B. If A rebuts B (at B') then it is also clear that X also rebuts B (at B'). Hence X attacks B wrt $att_{bs}(K)$.

Suppose C attacks X wrt $att_{bs}(K)$, i.e. C undercut X or C rebut X (at X') for $X \in A \uparrow \Omega$. If C undercuts X, it is clear that C also undercuts A. Suppose C rebuts X (at X'). From Lemma 4.2, there is a subargument A' of A such that $X' \in A' \uparrow \Omega$. Hence $last(A') = last(X')$ and A' is basic defeasible. Hence C rebuts A (at A'). Therefore C attacks A wrt $att_{bs}(K)$. □

Lemma 5.3. *Let $\mathcal{K}$ be a sensible class of basic knowledge bases that satisfy the self-contradiction property. Then the basic attack relation assignment att_{bs} satisfies the property of credulous cumulativity for $\mathcal{K}$.*

Proof. It follows immediately from Lemmas 6.3 and 6.6. The readers are advised to delay its easy verification until after the Lemma 6.6. □

The following theorem follows immediately from Lemmas 5.2, 5.3

Theorem 5.2. *Let $\mathcal{K}$ be a sensible class of basic knowledge bases that satisfy the self-contradiction property. Then the basic attack relation assignment att_{bs} satisfies all basic properties for $\mathcal{K}$.* □

Theorems 5.2, 5.1 confirm an insight that the combination undercut + rebut captures the intuition of the attack closure property even though rebut is not explicitly specified in it.[25]

An attentive reader may wonder whether the reverse of Theorem 5.1 holds. We give two examples and a lemma below to show that it does not hold in general.

Example 5.1. Let $K = (RS, RD, \preceq, BE)$ such that

- RS consists of four strict rules:

$$r_0 : a \to c \quad r_1 : b \to \neg c \quad r_2 : c \to \neg b \quad r_3 : \neg c \to \neg a$$

- RD consists of two defeasible rules:

$$d_0 : \ \Rightarrow a \quad d_1 : \ \Rightarrow b$$

- $\preceq = \emptyset$ and $BE = \emptyset$.

Let $\mathcal{K}$ be a sensible class of basic knowledge bases defined by:

$$K' \in \mathcal{K} \quad \text{iff} \quad K' = (RS, RD, \preceq, BE') \text{ s.t. } BE' \text{ is consistent wrt } RS$$

Let A_0, A_1, A_2, A_3 be defined as in Fig. 6.

Let att be an attack relation assignment defined for $\mathcal{K}$ as follows:

$$\forall K' \in \mathcal{K}, \quad att(K') = att_{bs}(K') \cup \{ (A_0, A_1), (A_1, A_0) \}$$

Since A_1 is a defeasible subargument of A_3, from $(A_0, A_1) \in att(K')$ and $(A_0, A_3) \notin att(K')$,[26] it is clear that att does not satisfy the property of subargument structure.

[25] We will see in section 7.4 (Lemma 7.5) that all MP-attack relation assignments proposed in [34] coincide with the basic attack relations for the basic knowledge bases.

[26] Suppose $(A_0, A_3) \in att(K')$. From the definition of att, it follows $(A_0, A_3) \in att_{bs}(K')$, i.e. A_0 rebuts A_3. This is impossible according to Definition 5.1. Hence $(A_0, A_3) \notin att(K')$.

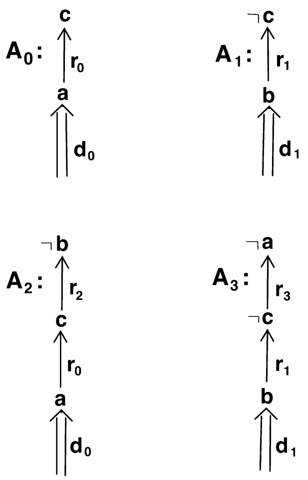

Fig. 6. Arguments of Example 5.1.

It is also obvious that for each $K' \in \mathcal{K}$, the stable extensions of $(AR_{K'}, att(K'))$ and $(AR_{K'}, att_{bs}(K'))$ coincide (see Example 5.2 and Lemma 5.4 below for a precise proof). So the reverse of Theorem 5.1 does not hold for $\mathcal{K}$. □

In Example 5.2, we present another sensible class of knowledge bases containing the one in Example 5.1 as a subclass.

Example 5.2. Let $\mathcal{K}$ be the sensible class of all consistent basic knowledge bases closed under transposition.
Let K be the knowledge base defined in Example 5.1. It is clear that $K \in \mathcal{K}$. Let $A_0, \dots, A_3$ be arguments defined as in Fig. 6.
Let *att* be an attack relation assignment defined for $\mathcal{K}$ as follows:
For each $K' \in \mathcal{K}$:

$$att(K') = \begin{cases} att_{bs}(K') & \text{if } \{A_0, A_1\} \not\subseteq AR_{K'} \\ att_{bs}(K') \cup \{(A_0, A_1), (A_1, A_0)\} & \text{otherwise} \end{cases}$$

Let $K' \in \mathcal{K}$. From Lemma 5.4 below, it follows that the stable extensions of $(AR_{K'}, att(K'))$ and $(AR_{K'}, att_{bs}(K'))$ coincide.
As $K \in \mathcal{K}$, it follows from the elaboration in Example 5.1 that att does not satisfy the property of subargument structure for $\mathcal{K}$. Therefore the reverse of Theorem 5.1 does not hold for $\mathcal{K}$. □

Lemma 5.4. *Let $\mathcal{K}$ be the sensible class of all consistent basic knowledge bases closed under transposition. Let att be the attack relation assignment defined in Example 5.2. Following assertions hold:*

1. *For each $K' \in \mathcal{K}$, the stable extensions of $(AR_{K'}, att(K'))$ and $(AR_{K'}, att_{bs}(K'))$ coincide.*
2. *The reverse of Theorem 5.1 does not hold for $\mathcal{K}$.*

Proof. See Appendix M. □

6. Axiomatic semantics for prioritized knowledge bases

There are many interpretations of the preference of one default to another. Some are due to specificity or social values while others could just be a specification of the operational order in applying the defaults [18,6,8,15,10]. Though the underlying intuitions for the introduction of rule preferences could be different, they all share a basic and natural interpretation of the preference of a defeasible rule over another rule that in a situation when each of them is applicable (i.e. the premises of both rules follow from the factual evidences and the strict rules in the knowledge base), but both could not be applied together, then the preferred one should be applied.

Example 6.1. Suppose K consists of just two defeasible rules: $d_1 :\Rightarrow a$, $d_2 :\Rightarrow \neg a$, with $d_1 \prec d_2$. The arguments $A_1 = [\Rightarrow a]$, $A_2 = [\Rightarrow \neg a]$ (illustrated in Fig. 7) rebut each other.
As d_2 is preferred to d_1, A_2 is considered an effective rebut against A_1 while A_1 is an ineffective rebut against A_2. Hence A_2 is an attack against A_1, but not vice versa. □

We adopt this basic and natural view of rule preference and capture it by the property of effective rebuts below.

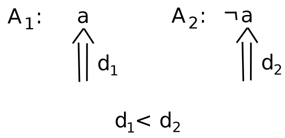

Fig. 7. Effective rebuts.

Definition 6.1 *(Effective rebut).* Let $\mathcal{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathcal{K}$. We say that att satisfies the effective rebut property for $\mathcal{K}$ iff for each knowledge base $K \in \mathcal{K}$, for all arguments $A_0, A_1 \in AR_K$ such that

- each A_i, $i = 0, 1$, contains exactly one defeasible rule d_i (i.e. $dr(A_i) = \{d_i\}$), and
- A_0 rebuts A_1,

the following assertion holds:

$$(A_0, A_1) \in att(K) \text{ iff } d_0 \not< d_1 \qquad \square$$

It is now the time to present the last property in our axiom system, of which the intuition is that attacks are directed towards "identifying" some links as the culprits in an attacked argument.

We first introduce a "weakening" operation, a kind of a "reverse" version of the strengthening operation.

Let $A \in AR_K$ and $AS \subseteq AR_K$. By $A \downarrow AS$ we denote the set of arguments obtained by replacing zero, one or more premises of A by arguments in AS whose conclusions coincide with the premises.

Definition 6.2 *(Weakening operation).*

- Let $A \in AR_K$ and $AS \subseteq AR_K$. The weakening of A by AS, denoted by $A \downarrow AS$ is defined inductively as follows:

$$A \downarrow AS = \begin{cases} \{[\alpha]\} \cup \{ X \in AS \mid cnl(X) = \alpha \} & \text{if } A = [\alpha] \text{ and } \alpha \in BE \\ \{ [X_1, \ldots, X_n, r] \mid X_i \in A_i \downarrow AS \} & \text{if } A = [A_1, \ldots, A_n, r] \end{cases}$$

- $B \in AR_K$ is said to be a weakening of A by AS iff $B \in A \downarrow AS$. $\square$

For an illustration, consider again the arguments in Example 1.3. Applying Definition 6.2 yields directly that $[P] \downarrow \{A_1\} = \{[P], A_1\}$, $A'_2 \downarrow \{A_1\} = \{A'_2, A_2\}$.

Lemma 6.1. *Let $A, B \in AR_K$, $AS \subseteq AR_K$ such that $S = cnl(AS) \subseteq BE$.*
It holds that if $X \in A \downarrow AS$ then $A \in X \uparrow S$.

Proof. By induction on the structure of A.

1. Base case: $A = [\alpha]$, $\alpha \in BE$. From Definition 6.2, it follows immediately $X \in [\alpha] \downarrow AS$ iff $X = [\alpha]$ or $X \in AS$ and $cnl(X) = \alpha$. From Definition 4.4, it follows immediately $[\alpha] \in X \uparrow cnl(AS)$ iff $X = [\alpha]$ or $cnl(X) = \alpha$ and $\alpha \in cnl(AS)$. The lemma holds obviously.
2. Inductive step. Let $A = [A_1, \ldots, A_n, r]$. Suppose $X \in A \downarrow AS$. From Definition 6.2, $A \downarrow AS = \{[A'_1, \ldots, A'_n, r] \mid A'_i \in A_i \downarrow AS\}$. Hence $X = [X_1, \ldots, X_n, r]$ where $X_i \in A_i \downarrow AS$. From induction hypothesis, $A_i \in X_i \uparrow S$. From Definition 4.4, $A \in X \uparrow S$. $\square$

Note that the reverse of the above Lemma 6.1 does not hold in general.[27]

Definition 6.3 *(Link-orientation).* Let $\mathcal{K}$ be a sensible class of knowledge bases and att be an attack relation assignment defined for $\mathcal{K}$. We say that att satisfies the link-oriented property (or property of link-orientation) for $\mathcal{K}$ iff for each knowledge base $K \in \mathcal{K}$, for all arguments $A, B, C \in AR_K$ such that C is a weakening of B by $AS \subseteq AR_K$ (i.e. $C \in B \downarrow AS$), the following assertion holds:

If A attacks C wrt att(K) and A does not attack AS wrt att(K) then A attacks B wrt att(K).[28] $\square$

[27] To see this, just let $AS = \{N_1\}$ (see Figs. 1, 2) and $X = N'_1$ and $A = [\neg inno(P_2)]$. Let $S = cnl(AS) = \{\neg inno(P_2)\}$. It is clear that $A \in X \uparrow S$. But $A \downarrow AS = \{A, N_1\}$. Therefore $X \notin A \downarrow AS$.

[28] I.e. if $(A, C) \in att(K)$ and $\forall X \in AS : (A, X) \notin att(K)$ then $(A, B) \in att(K)$.

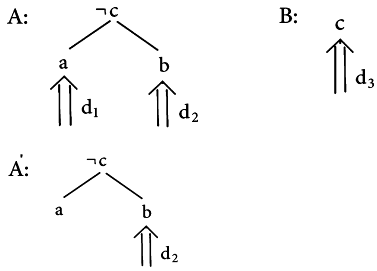

Fig. 8. Normal rebuts.

For an illustration, consider again arguments in Figs. 3, 4 (Example 1.3). As A_3 does not attack A_1, and A_2 is a weakening of A'_2 by $\{A_1\}$,[29] in any attack relation assignment att satisfying the link-oriented axiom, if A_3 attack A_2 wrt $att(K + \{P\})$ then A_3 attacks A'_2 wrt $att(K + \{P\})$.

As elaborated in Example 1.3, with respect to the ASPIC+-attack relation based on the weakest link principle and the elitist ordering, A_3 attack A_2 but A_3 does not attack A'_2. Hence this attack relation assignment does not satisfy the link-oriented property.

Definition 6.4 *(Ordinary attack relations).* An attack relation assignment att is said to be **ordinary** for a sensible class $\mathcal{K}$ of knowledge bases iff it satisfies the properties of credulous cumulativity, attack monotonicity, context-independence, subargument structure, attack closure, effective rebuts and link-orientation for $\mathcal{K}$. □

6.1. A structural insight into ordinary attack relation assignments

Apart from undercut, stable extensions wrt ordinary attack relation assignments employ a special kind of rebut, referred to as normal-rebut, to attack arguments not belonging to it. Let us first illustrate the idea.

Let *att* be an ordinary attack relation assignment for a sensible class of knowledge bases $\mathcal{K}$. Consider a knowledge base $K = (RS, RD, \preceq, BE)$ in $\mathcal{K}$ such that the arguments A and B (in Fig. 8) belong to its set of arguments. Note that the bodies of defaults d_1, d_2, d_3 are empty.

As A rebuts B, the question is under which condition the rebut is effective such that A is considered an attack against B.

Suppose A attacks B. What could we say about the preferences between defeasible rules in A and B? Let $K' = K + \{a\}$ and suppose $K' \in \mathcal{K}$. It is clear that A' is an argument wrt K'. From the attack monotonicity and context-independence of *att*, it follows that A' attacks B. From the effective rebut property, it follows that $d_2 \not\prec d_3$.[30]

Similarly we could also conclude that $d_1 \not\prec d_3$ (if $K + \{b\}$ belongs to $\mathcal{K}$).

It turns out that the above discussed scenarios of preferences between defeasible rules in A, B are special cases of a general pattern of rebut, referred to as normal-rebut that are employed by stable extensions wrt ordinary attack relation assignments to attack arguments not belonging them. We give a formal definition of normal-rebut below followed by a lemma capturing this insight.

Definition 6.5 *(Normal rebut).* Let K be a knowledge base and $A, B \in AR_K$. We say that A normal-rebuts B (at X) iff A rebuts B (at X) and the following normal condition holds.

(Normal condition). There is no defeasible rule $d \in ldr(A)$ such that $d \prec last(X)$. □

It is not difficult to see that if K is basic, normal-rebuts coincide with rebuts.

The following Lemma 6.2 generalizes the characteristic Lemma 5.1 for ordinary attack relation assignments.

Lemma 6.2 *(General characteristic lemma). Let att be an ordinary attack relation assignment for a sensible class $\mathcal{K}$ of knowledge bases. Further let $K \in \mathcal{K}$ and E be a stable extension of $(AR_K, att(K))$, $A \in E$ and $B \in AR_K \setminus E$ such that A attacks B wrt att(K). Then A undercuts or normal-rebuts B.*

Proof. As any ordinary attack relation assignment satisfies all basic properties, from Lemma 5.1, it follows immediately that A undercuts or rebuts B. The rest of the proof follows from Lemma E.1 in Appendix E. □

Lemma 6.2 shows that stable extensions wrt ordinary attack relations employ normal-rebuts or undercuts to attack arguments not belonging to it. This insight suggests that an attack relation based on undercuts and normal-rebuts could be ordinary. We show below that it is indeed the case.

[29] $A'_2 \downarrow \{A_1\} = \{A'_2, A_2\}$.

[30] Note that the condition $d_2 \not\prec d_3$ does not necessarily hold if K' does not belong to $\mathcal{K}$.

6.2. Introducing normal attack relation assignment

Definition 6.6 *(Normal attack relation assignment).* Let $\mathcal{K}$ be the sensible class of all consistent knowledge bases. The **normal attack relation** att_{nr} is defined for $\mathcal{K}$ as follows:

For any knowledge base $K \in \mathcal{K}$ and any arguments $A, B \in AR_K$, $(A, B) \in att_{nr}(K)$ if and only if A undercuts B or A normal-rebuts B. □

It is obvious that the normal attack relation assignment generalizes the basic attack relation assignment of basic knowledge bases.

Lemma 6.3. *For any consistent basic knowledge base K, $att_{nr}(K) = att_{bs}(K)$.* □

We proceed to show that the normal attack relation assignment att_{nr} represents a normal form among the ordinary attack relation assignments in the following sense:

- *The normal attack relation assignment att_{nr} is ordinary.*
- *Stable extensions wrt any ordinary attack relation assignment are also stable extensions wrt the normal attack relation assignment.*

6.3. Normal attack relation assignment is ordinary

We start with three relevant lemmas.

Lemma 6.4. *att_{nr} satisfies the properties of context-independence, subargument structure, attack closure, effective rebuts and link-orientation for any sensible class of knowledge bases.*

Proof. It is obvious that normal attack relations satisfy the properties of context-independence, attack closure, effective rebuts and subargument structure.

It remains to show that att_{nr} also satisfies the link-oriented property. Let $A, B, C \in AR_K$ for a knowledge base K such that C is a weakening of B by $AS \subseteq AR_K$ (i.e. $C \in B \downarrow AS$) and A does not attack AS wrt $att_{nr}(K)$ and $(A, C) \in att_{nr}(K)$. There is a basic defeasible subargument C' of C such that either $cnl(A) = ab_{last(C')}$ or $cnl(A) = \neg cnl(C')$ and there is no $d \in ldr(A)$ s.t. $d \prec last(C')$. Since A does not attack AS wrt $att_{nr}(K)$, the default $last(C')$ does not occur in any argument belonging to AS. Hence $last(C')$ occurs in B. Therefore $(A, B) \in att_{nr}(K)$. □

Lemma 6.5. *Let $\mathcal{K}$ be a sensible class of knowledge bases. The normal attack relation assignment att_{nr} satisfies the property of attack monotonicity for $\mathcal{K}$.*

Proof. Let $K \in \mathcal{K}$ and A attacks B wrt att_{nr} and $X \in A \uparrow \Omega$. It is not difficult to see that if A undercuts B then X also undercuts B. Suppose now that A rebuts B (at B') and there is no $d \in ldr(A)$ s.t. $d \prec last(B')$. From $ldr(X) \subseteq ldr(A)$ and cnl(X) = cnl(A), it follows obviously that X rebuts B (at B') and there is no $d \in ldr(X)$ s.t. $d \prec last(B')$. We have proved that X also attacks B wrt att_{nr}.

Suppose C attacks X wrt att_{nr} for $X \in A \uparrow \Omega$. It is easy to see that there exists a basic defeasible subargument X' of X such that either $last(C) = ab_{last(X')}$ or C normal-rebuts X' (at X'). From Lemma 4.2, there is a subargument A' of A such that $X' \in A' \uparrow \Omega$. Hence $last(A') = last(X')$ and A' is basic defeasible. It holds obviously that either $last(C) = ab_{last(A')}$ or C normal-rebuts A' (at A'). C thus attacks A wrt att_{nr}. □

Lemma 6.6. *Let $\mathcal{K}$ be a sensible class of knowledge bases that satisfy the property of self-contradiction. Then the normal attack relation assignment att_{nr} satisfies the credulous cumulativity property for $\mathcal{K}$.*

Proof. See Appendix F. □

It is now possible for us to present a key result stating that the normal attack relation assignment is ordinary.

Theorem 6.1. *The normal attack relation assignment att_{nr} is ordinary for any sensible class of knowledge bases that satisfy the self-contradiction property.*

Proof. The properties of context-independence, subargument structure, attack closure, effective rebuts and link-orientation follow immediately from Lemma 6.4. The property of attack monotonicity follows from Lemma 6.5. The property of credulous cumulativity is proved in Lemma 6.6. □

6.4. Normal form of ordinary semantics

Our next result is that the normal attack relation assignment represents a normal form of ordinary semantics in the sense that the stable extensions wrt any ordinary attack relation assignment are also stable wrt the normal attack relation assignment.

Theorem 6.2. *Let $\mathcal{K}$ be a sensible class of knowledge bases. Let att be an ordinary attack relation assignment defined for $\mathcal{K}$. Then for each $K \in \mathcal{K}$, each stable extension of $(AR_K, att(K))$ is also a stable extension of $(AR_K, att_{nr}(K))$.*

Proof. Let E be a stable extension of $(AR_K, att(K))$.

Because att satisfies the credulous cumulativity property, $cnl(E)$ is consistent (Lemma 4.1). Since E is conflict-free wrt att(K) and att satisfies the property of attack closure, E is free of undercut-attacks.[31]

We show that E is conflict-free wrt $att_{nr}(K)$.

Suppose E is not conflict-free wrt $att_{nr}(K)$. Therefore there is $A, B \in E$ s.t. $(A, B) \in att_{nr}(K)$. Since E is free from undercut-attacks, it follows A rebuts B at a basic defeasible subargument B' of B. Therefore $cnl(A) = \neg cnl(B')$. From Lemma 4.3, $B' \in E$. Therefore $cnl(E)$ is contradictory. This is impossible since $cnl(E)$ is consistent. Hence E is conflict-free wrt $att_{nr}(K)$.

It remains to show that E attacks every argument not belonging to it wrt $att_{nr}(K)$.

Let $B \in AR_K \setminus E$. Therefore there is $A \in E$ s.t. $(A, B) \in att(K)$. From Lemma 6.2, it follows that A undercuts or normal-rebuts B. From the definition of att_{nr}, it follows immediately that $(A, X) \in att_{nr}(K)$.

E is hence also a stable extension of K wrt the normal attack relation $att_{nr}(K)$. □

Examples 5.1, 5.2 and Lemmas 5.4, 6.3 show that the reverse of Theorem 6.2 does not hold in general.

7. Relations to other approaches

7.1. Operational interpretation of rule ordering

Preference orders between rules in prioritized default logics or logic programming are viewed in [15,44] as specifying application orders of rules. We show in this section that this operational reading of preferences is sound wrt normal semantics and also complete for the class of stratified knowledge bases.

The operational reading of preferences defines the semantics of a knowledge base $K = (RS, RD, \preceq, BE)$ in two steps: first determining the stable extensions of the basic knowledge base $K_{basic} = (RS, RD, \emptyset, BE)$ underlying K and then applying the preference relation to pick the preferred extensions.

We first adapt the definitions in [15,44] to structured argumentation below.

Definition 7.1. A stable extension E of $(AR_K, att_{bs}(K_{basic}))$[32] is said to be an **enumeration-based extension** of K iff there is an enumeration $(d_i)_{i \geq 1}$ of $\Gamma_E = \{d \in RD \mid \text{d appears in some argument of E}\}$ such that for all i, j, we have:

1. $\{hd(d_k) | k < i\} \cup BE \vdash_K bd(d_i)$;
2. if $d_i \prec d_j$ then $j < i$;
3. if $d_i \prec d$ and $d \in RD \setminus \Gamma_E$ then
   (a) $bd(d) \nsubseteq cnl(E)$, or
   (b) $\{hd(d_k) | k < i\} \cup BE \vdash_K \neg hd(d)$, or
   (c) $\{hd(d_k) | k < i\} \cup BE \vdash_K ab_d$. □

The intuitions of the first two conditions in Definition 7.1 should be clear. The third condition states that if a default d is not applied while a less preferred d_i is, then this is either because some of the premises of d are not satisfied wrt the belief set generated by the extension or because d has been already rebutted or undercut before d_i is applied.

The following theorem shows the soundness of the enumeration-based semantics wrt the normal semantics.

Theorem 7.1. *Every enumeration-based extension of K is a stable extension of $(AR_K, att_{nr}(K))$.*

Proof. See Appendix G. □

When the operational interpretation of rule preferences interferes with the basic control mechanism of "applying a rule when its premises are satisfied", there could be no enumeration-based extension. For example, consider a knowledge base

[31] That means that there are no arguments $X, Y \in E$ s.t. X undercuts Y.

[32] Note that the arguments of K and K_{basic} are identical, and $att_{bs}(K_{basic}) = att_{nr}(K_{basic})$.

consisting of just two defeasible rules $d: \Rightarrow a$, $d': a \Rightarrow b$ with $d \prec d'$. There is no enumeration-based extension in this case as the first two conditions in Definition 7.1 cannot be fulfilled.

When the basic control mechanism and the operational reading of the default preferences do not interfere, the enumeration-based semantics and the normal semantics coincide. We introduce below a class of stratified knowledge bases where the two mechanisms do not interfere by applying the concept of stratification in logic programming and Brewka's idea of ranking function [4,8] to our framework.

7.1.1. Stratified knowledge bases

A preference relation $\preceq$ is said to be **ranked** iff there is a ranking function ρ assigning non-negative integers to defeasible rules in RD such that for all $d, d' \in RD$, $d \preceq d'$ iff $\rho(d') \leq \rho(d)$.

For an illustration, the preference relation in Example 1.3 could be ranked by $\rho(d_2) = 0$, $\rho(d_3) = 1$, $\rho(d_1) = 2$.

Notation 7.1. Abusing the notation for simplicity, for any argument A, we denote the maximum of the ranks of the defeasible rules appearing in A by $\rho(A)$.

We introduce below the notion of stratified knowledge base that is inspired by both the concept of stratification in logic programming and Brewka's ranking function [4,8].

Definition 7.2. A knowledge base K is said to be **stratified** iff its preference relation $\preceq$ is ranked with a ranking function ρ such that following conditions are satisfied:

1. For each basic defeasible argument A, for each defeasible rule d occurring in A and different to $last(A)$, $\rho(d) < \rho(last(A))$.
2. For each argument A such that $cnl(A) = ab_d$ for some defeasible rule d, $\rho(A) \leq \rho(d)$. □

Interpreting the ranking of defaults as their application order, the first condition states the obvious that the last rule in a basic defeasible argument should be applied last while the meaning of the second one is that if a default is undercut then the undercut should happen before the default's turn to be applied.

For an illustration, the preference relation in Example 1.3 could be ranked by $\rho(d_2) = 0$, $\rho(d_3) = 1$, $\rho(d_1) = 2$. The knowledge base is not stratified as in argument A_2, $\rho(d_1) = 2 > 0 = \rho(d_2) = \rho(last(A_2))$.

Lemma 7.1. *Suppose K is stratified and $A, B \in AR_K$ such that $(A, B) \in att_{nr}(K)$. Then $\rho(A) \leq \rho(B)$.*

Proof. See Appendix G. □

Theorem 7.2. *Suppose K is a consistent and stratified knowledge base satisfying the self-contradiction property. Then each stable extension of $(AR_K, att_{nr}(K))$ is an enumeration-based extension of K.*

Proof. See Appendix G. □

7.2. Brewka and Eiter principles

Brewka and Eiter [10] have proposed two principles, referred to as BE-principles in this section, for the evaluation of semantics of prioritized default reasoning. We show in this section that both BE-principles are satisfied by the ordinary attack relation assignments.

The first BE principle concerns the intuition of preferences between defeasible rules while the second one is about relevance of rules. We start with the second one as it is the simpler of them.

The intuition of the second BE-principle is best illustrated by a simple example. Suppose the weather forecast for tomorrow is sunshine, light wind, no rain. You plan to go yachting. Suppose somebody tells you a rule that in stormy weather, yachting is forbidden. Of course knowing this rule will not affect your plan as the weather is fine and the rule is not applicable.

We say that a rule r is applicable wrt set of literals S if $bd(r) \subseteq S$. We adapt the second BE-principle from [10] to our framework below.

Notation 7.2. Let $K = (RS, RD, \preceq, BE)$ be a knowledge base and r be a rule. Define $K + r = (RS', RD', \preceq, BE)$ as follows:

$$(RS', RD') = \begin{cases} (RS \cup \{r\}, RD) & \text{if r is strict} \\ (RS, RD \cup \{r\}) & \text{if r is defeasible} \end{cases}$$

Definition 7.3 *(Second BE-principle).* An attack relation assignment att is said to satisfy the second BE-principle for a sensible class of knowledge bases $\mathcal{K}$ iff for each knowledge base $K \in \mathcal{K}$ if S is a stable belief set of K (wrt att) and r is a rule not applicable wrt S and $K + r$ belongs to $\mathcal{K}$ then S is also a stable belief set of $K + r$ (wrt att). □

Theorem 7.3. *Let att be an attack relation assignment satisfying the properties of subargument structure and context-independence for a sensible class of knowledge bases $\mathcal{K}$. Then att satisfies the second BE-principle for $\mathcal{K}$.*

Proof. See Appendix I. □

The intuition of the first BE-principle is that if two belief sets are generated by the same sets of rules with the exceptions of two defaults d, d' with $d \prec d'$ then the one with d should not be preferred.

For illustration, in Example 1.3, the belief sets $\{D, A, P, T\}$, $\{D, A, P, \neg T\}$ are generated respectively by the sets of rules $\{r, d_1, d_2\}$, $\{r, d_1, d_3\}$. Since $d_3 \prec d_2$, the set $\{D, A, P, \neg T\}$ should not be a stable belief set according to the first BE-principle.

In [10], the first BE-principle is presented for extended logic programming. We adapt it to our framework below.

Let S be a set of literals. *S is said to be generated by a set $\Gamma \subseteq RS \cup RD$* iff the following conditions are satisfied

1. For each literal σ, $\sigma \in S$ iff there is an argument A such that $cnl(A) = \sigma$ and all rules appearing in A belong to Γ.
2. For each strict rule $r \in RS$, $r \in \Gamma$ iff $bd(r) \subseteq S$, $hd(r) \in S$.
3. For each defeasible rule $d \in RD$, $d \in \Gamma$ iff $bd(d) \subseteq S$, $hd(d) \in S$ and $ab_d \notin S$.

Note that the uniqueness of Γ wrt S follows immediately from the second and third conditions. In general, these two conditions are not sufficient to guarantee the first one. To see this simple but relevant point, consider a knowledge base consisting of just a unique strict rule $r : a \to a$. Then $S = \{a\}$ and $\Gamma = \{r\}$ satisfy both the second and third condition but not the first.

Γ, if exists, is often referred to as the generating set of S.

Definition 7.4 *(First BE-principle).* We say that an attack relation assignment att satisfies the first BE-principle for a sensible class $\mathcal{K}$ of knowledge bases iff for each $K \in \mathcal{K}$, the following condition is satisfied:

Suppose S, S' be consistent sets of literals generated respectively by sets of rules $\Gamma \cup \{d\}$ and $\Gamma \cup \{d'\}$ such that $d, d' \in RD \setminus \Gamma$ and $d \prec d'$. Then S is not a stable belief set of K wrt att. □

Theorem 7.4. *Let att be an ordinary attack relation assignment defined for a sensible class $\mathcal{K}$ of knowledge bases. Then att satisfies the first BE-principle for $\mathcal{K}$.*

Proof. See Appendix I. □

An attentive reader may wonder whether the reverses of Theorems 7.4, 7.3 hold. The answer is "No". To see this, let us consider again Example 5.2 and Lemma 5.4. From Lemma 5.4, it follows that the set of stable extensions wrt att and att_{bs} coincide. Therefore these two attack relation assignments have identical stable belief sets. Since att_{bs} satisfies both BE-principles for $\mathcal{K}$, att also satisfies both BE-principles. But att does not satisfy the property of subargument structure for $\mathcal{K}$.

7.3. Credulous cumulativity in prioritized logic programming

In this section we show that credulous cumulativity is indeed embraced, although implicitly, in prioritized logic programming. We first show that the answer set semantics of extended logic programs without rule preferences satisfies credulous cumulativity. In the two following sections we then show that both well-known and well-studied approaches to preference handling in logic programming, the approach advocated by Delgrande, Schaub, Tompits, Wang and others [15,44] and the Brewka and Eiter approach [10], satisfy the credulous cumulativity property.

From now on until the end of this section, we assume a language $\mathcal{C}$ consisting of ground atoms α and their classical negation $\neg\alpha$. Literals in $\mathcal{C}$ are often simply referred to as *classical literals*. Further, for each classical literal λ, we introduce a naf-literal[33] of the form *not*_λ.

A logic program rule (or just lp-rule for short) r is of the form

$$h \leftarrow l_1, \ldots, l_n, not\_l_{n+1}, \ldots, not\_l_{n+k}$$

such that $h, l_1, \ldots, l_n, l_{n+1}, \ldots, l_{n+k}$ are classical literals. For ease of reference, we denote $hd(r) = h$, $bd^+(r) = \{l_1, \ldots, l_n\}$ and $bd^-(r) = \{l_{n+1}, \ldots, l_{n+k}\}$.

[33] naf stands for "negation as failure".

An extended logic program is a finite set of logic program rules. The semantics of extended logic programs is defined by their answer sets [22,23].

Given a set of lp-rules P and a set X of classical literals. The Gelfond–Lifschitz reduction (or just GL-reduction for short) of P, denoted by P_X, is obtained by

- deleting each rule $h \leftarrow l_1, \ldots, l_n, not\_l_{n+1}, \ldots, not\_l_{n+k}$ in P such that some $l_{n+j} \in X$ for $1 \leq j \leq k$, and
- deleting all naf-literals from the remaining rules.

A non-contradictory set[34] S of classical literals is an answer set of P iff S is the smallest set of literals such that S is closed wrt P_S, i.e. for any rule $h \leftarrow l_1, \ldots l_n$ in P_S if $l_1, \ldots l_n$ belong to S, then $h \in S$.[35]

For a logic program P and a finite set Ω of classical literals, define

$$P + \Omega = P \cup \{\omega \leftarrow \ | \ \omega \in \Omega\}$$

The following theorem shows that the credulous cumulativity property is satisfied by the answer set semantics of extended logic programming.

Theorem 7.5. *Let P be an extended logic program and S be an answer set of P. Further let $\Omega \subseteq S$. Then S is also an answer set of $P + \Omega$.*

Proof. Let $P' = P + \Omega$. It is not difficult to see that $P'_S = P_S + \Omega$. Therefore S is closed wrt rules in P'_S. As any set that is closed wrt rules in P'_S is also closed wrt rules in P_S, S is obviously the smallest set that is closed wrt P'_S. S is hence an answer set of P'. □

7.3.1. Delgrande, Schaub, Tompits (DST) – preferred answer sets for prioritized logic programming

Delgrande, Schaub and Tompits [15] view the preferences between rules as constraints on their application orders.

Formally, a prioritized logic program is a pair $(P, \prec)$ where P is an extended logic program and $\prec$ is a strict partial order on P.[36]

Given a set of classical literals X, denote

$$\Gamma_{P,X} = \{r \in P \mid bd^+(r) \subseteq X \text{ and } bd^-(r) \cap X = \emptyset\}$$

Definition 7.5 *(DST-preferred answer sets).* (See [15].) Let $\Pi = (P, \prec)$ be a prioritized logic program. A DST-preferred answer set of Π is an answer set S of P such that there is an enumeration $(r_i)_{i \geq 1}$ of the rules in $\Gamma_{P,S}$ such that for all i, j, we have:

1. $bd^+(r_i) \subseteq \{hd(r_k) | k < i\}$;
2. if $r_i \prec r_j$ then $j < i$;
3. if $r_i \prec r$ and $r \in P \setminus \Gamma_{P,S}$ then
   (a) $bd^+(r) \not\subseteq S$, or
   (b) $bd^-(r) \cap \{hd(r_k) | k < i\} \neq \emptyset$. □

For a finite set Ω of classical literals and a prioritized logic program $\Pi = (P, \prec)$, define

$$\Pi + \Omega = (P + \Omega, \prec)$$

The following theorem shows that the credulous cumulativity property is satisfied by the DST-answer set semantics of prioritized logic programming.

Theorem 7.6. *Let S be a DST-preferred answer set of a prioritized logic program Π and $\Omega \subseteq S$. Then S is also a DST-preferred answer set of $\Pi + \Omega$.*

Proof. See Appendix L. □

[34] Such sets are referred to as consistent in the literature [15,10]. We do not use this notion of consistency here to avoid possible misunderstanding with our notion of consistency in Definition 3.5.

[35] In earlier papers [22], the inventors of answer set semantics allow contradictory answer sets. In later papers [30,32] they suggest to eliminate contradictory answer sets altogether. We follows their suggestion in this paper.

[36] A strict partial order is an irreflexive, antisymmetric and transitive relation.

7.3.2. Brewka, Eiter (BE) – preferred answer sets for prioritized logic programming

Brewka and Eiter [10] consider a prioritized logic program $(P, \prec)$ as representative of the fully prioritized programs $(P, \ll)$ such that $\prec \subseteq \ll$ and $\ll$ is a strict total order.[37] The preferred answer sets of $(P, \prec)$ are defined as preferred answer sets of $(P, \ll)$.

A lp-rule r is said to be *prerequisite-free* iff $bd^+(r) = \emptyset$. A *prerequisite-free* program consists only of prerequisite-free rules. The preferred answer sets of prerequisite-free programs rely on an operator defined below.

Definition 7.6. (See [10].) Let $\Pi = (P, \ll)$ be a prerequisite-free fully prioritized program and $(r_i)_{i \geq 1}$ be the enumeration of P according to $\ll$,[38] and let X be a non-contradictory set of literals. A sequence $X_0, X_1, \ldots, X_n$ where $n = |P|$,[39] is defined by:

1. $X_0 = \emptyset$.
2. For $1 \leq i \leq n$,

$$X_i = \begin{cases} X_{i-1} & \text{if } bd^-(r_i) \cap X_{i-1} \neq \emptyset \\ X_{i-1} & \text{if } hd(r_i) \in X \text{ and } bd^-(r_i) \cap X \neq \emptyset \\ X_{i-1} \cup \{hd(r_i)\} & \text{otherwise} \end{cases}$$

Define $C_\Pi(X) = X_n$. □

Definition 7.7. (See [10].) Let $\Pi = (P, \ll)$ be a prerequisite-free fully prioritized program. A set of classical literals S is a BE-preferred answer set of Π iff S is an answer set of P and $C_\Pi(S) = S$. □

In the next step, we recall the definition of BE-preferred answer sets of fully prioritized (possibly not prerequisite-free) logic programs.

For a lp-rule r of the form $h \leftarrow l_1, \ldots, l_n, not\_l_{n+1}, \ldots, not\_l_{n+k}$, let r^- denote the prerequisite-free rule $h \leftarrow not\_l_{n+1}, \ldots, not\_l_{n+k}$.

Definition 7.8. (See [10,15].) Let $\Pi = (P, \ll)$ be a fully prioritized logic program and X be a set of classical literals. The BE-reduction of Π wrt X is the logic program $\Pi^X = (P^X, \ll^X)$ obtained from $(P, \ll)$ as follows:

1. $P^X = \{r^- \mid r \in P \text{ and } bd^+(r) \subseteq X\}$, and
2. for any $r'_1, r'_2 \in P^X$, $r'_1 \ll^X r'_2$ iff $r_1 \ll r_2$ where $r_i = max_\ll\{r \in P \mid r^- = r'_i\}$. □

In other words, viewing a fully prioritized program as a list of rules where the more preferred rules are listed before the less preferred ones, Π^X is obtained from Π by 1) eliminating rules r whose prerequisites $bd^+(r)$ are not satisfied by X and 2) deleting all prerequisites from the remaining rules. If some rules appear more than once then the repeating copies that appear later in the list are deleted.

Definition 7.9 *(BE-preferred answer sets).* (See [10].) Let P be an extended logic program and S be an answer set of P.

1. S is a BE-preferred answer set of a fully prioritized logic program $\Pi = (P, \ll)$ iff S is a BE-preferred answer set of $\Pi^S = (P^S, \ll^S)$, the BE-reduction of Π wrt S.
2. S is a BE-preferred answer set of prioritized logic program $(P, \prec)$ iff S is a BE-preferred answer set of a fully prioritized logic program $(P, \ll)$ such that $\prec \subseteq \ll$. □

Before proceeding to show that the credulous cumulativity property is satisfied wrt BE-preferred answer set semantics, we present a lemma characterizing the structure of BE-preferred answer sets of prerequisite-free fully prioritized logic programs.

Lemma 7.2. *Let $\Pi = (P, \ll)$ be a fully prioritized prerequisite-free logic program and X be an answer set of P. Then X is a BE-preferred answer set of Π iff for each $r \in P \setminus \Gamma_{P,X}$, if $hd(r) \notin X$ then*

$$bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{P,X} \text{ and } r \ll r'\} \neq \emptyset$$

[37] I.e. $\ll$ is a strict partial order such that for every pair of distinct rules r, r', either $r \ll r'$ or $r' \ll r$.
[38] I.e. $r_i \ll r_j$ iff $j < i$.
[39] Note that a logic program is a finite set of lp-rules.

Proof. See Appendix L. $\square$

Lemma 7.3. *Let $\Pi = (P, \ll)$ be a fully prioritized logic program and S be a BE-preferred answer set of Π and $\Omega \subseteq S$. Then there is strict total order $\ll'$ on $P + \Omega$ such that $\ll \subseteq \ll'$ and S is also a BE-preferred answer set of $(P + \Omega, \ll')$.*

Proof. See Appendix L. $\square$

The following theorem shows that the credulous cumulativity property is satisfied wrt the BE-answer set semantics for prioritized extended logic programs.

Theorem 7.7. *Let $\Pi = (P, \prec)$ be a prioritized logic program and S be a BE-preferred answer set of Π. Further let $\Omega \subseteq S$. Then S is also a BE-preferred answer set of $\Pi + \Omega$.*

Proof. Let $\ll$ be a strict total order over P such that $\prec \subseteq \ll$ and S be a BE-preferred answer set of $(P, \ll)$. From Lemma 7.3, there is a total order $\ll'$ such that $\ll \subseteq \ll'$ and S is a BE-preferred answer set of $(P + \Omega, \ll')$. Therefore $\prec \subseteq \ll'$. Therefore S is also a BE-preferred answer set of $\Pi + \Omega = (P + \Omega, \prec)$. $\square$

7.4. Relationship to ASPIC+

ASPIC+ [35,34] is an influential and complex approach to structured argumentation incorporating many key concepts from distinct approaches to structured argumentation [38,40,7,15,10]. The semantics of ASPIC+ is based on the intuitive idea of defining attack relations based on preferences between arguments when a less preferred argument cannot attack a more preferred one as advocated in [3,6]. A recent work of Hunter and Williams [27] demonstrates the practicality of preference-based abstract argumentation by applying it in aggregating evidence about treatments in medicine.

We first recall below several key definition in ASPIC+ [34,35] adapted to our framework.

For simplification, a basic knowledge base $K = (RS, RD, \emptyset, BE)$ is often abbreviated as a triple (RS, RD, BE).

Definition 7.10. An ASPIC+-structured argumentation framework is a pair $(K, \sqsubseteq)$ where K is a basic knowledge base and $\sqsubseteq$ is a binary relation over the set of arguments in AR_K.[40]

We write $A \sqsubset B$ iff $A \sqsubseteq B$ and $B \not\sqsubseteq A$. $\square$

Definition 7.11. Given an ASPIC+-structured argumentation framework $SA = (K, \sqsubseteq)$, the ASPIC+-attack relation (or just AP-attack relation for short) $att_{AP}(SA) \subseteq AR_K \times AR_K$ is defined as follows.

$(A, B) \in att_{AP}(SA)$ iff one of the following conditions is satisfied:

1. A undercuts B.
2. A rebuts B (at B') and $A \not\sqsubset B'$. $\square$

It is easy to see that the following lemma holds.

Lemma 7.4. *Let $SA = (K, \sqsubseteq)$ be an ASPIC+-structured argumentation framework. It holds that*

1. *$att_{AP}(SA) \subseteq att_{bs}(K)$, and*
2. *the ASPIC+-attack relation $att_{AP}(K)$ satisfies the properties of subargument structure and attack closure.* $\square$

Modgil and Prakken apply ASPIC+ to reasoning with rule preferences by presenting four ways to derive argument preferences from the preferences between rules. Based on these derived argument preference relations, four attack relations[41] are derived. We recall them in the following Definitions 7.12, 7.13, 7.14.

Remark 7.1. In [34,35], the Definitions 7.12, 7.13, 7.14 are given only for knowledge bases with preorder-preference relations. We adapt them to general knowledge bases directly.

Definition 7.12. Let K be a knowledge base and $\preceq$ be the preference relation over defeasible rules of K and Γ, Γ' be two finite sets of defeasible rules of K and $y \in \{E, D\}$,[42] define:

---

[40] In [34], defeasible rules of the form "$\delta : b_1, \ldots, b_n \Rightarrow ab_d$" are also allowed. Such rules could be captured in our framework by two rules "$\delta : b_1, \ldots, b_n \Rightarrow new_\delta$" and "$new_\delta \rightarrow ab_d$" where new_δ is a new atom not appearing in any other rule.

[41] Called defeats in [35,34].

[42] E, D stand for Elitist and Democratic respectively.

$$\Gamma \trianglelefteq_y \Gamma' \quad \text{iff} \quad \Gamma \neq \emptyset \text{ and one of the following conditions holds:}$$

1. $\Gamma' = \emptyset$.
2. $y = E$ and $\exists d \in \Gamma$ s.t. $\forall d' \in \Gamma' : d \preceq d'$.
3. $y = D$ and $\forall d \in \Gamma \exists d' \in \Gamma' : d \preceq d'$.

We write $\Gamma \triangleleft_y \Gamma'$ iff $\Gamma \trianglelefteq_y \Gamma'$ and $\Gamma' \not\trianglelefteq_y \Gamma$. □

Definition 7.13. Let K be a knowledge base, A, B be two arguments in AR_K and $y \in \{E, D\}$.

1. B is **preferred to** A according to the **last link principle** and the **y-ordering** (or **y-principle**), denoted by $A \sqsubseteq_{ly} B$ if and only if $ldr(A) \trianglelefteq_y ldr(B)$.
2. B is **preferred to** A according to the **weakest link principle** and the **y-ordering** (or **y-principle**), denoted by $A \sqsubseteq_{wy} B$ if and only if $dr(A) \trianglelefteq_y dr(B)$. □

Remark 7.2. Note that for ease of reference, we often talk interchangeably of elitist (resp. democratic) ordering or elitist (resp. democratic) principle.

According to Definition 7.11 there are four different attack relations which are recalled in the definition below.

Definition 7.14. Let K be a knowledge base, A, B be two arguments in AR_K. For $x \in \{l, w\}$ and $y \in \{E, D\}$, define

$$(A, B) \in att_{xy}(K) \text{ if and only if}$$

- A rebuts B (at B') such that $A \not\sqsubset_{xy} B'$, or
- A undercuts B. □

We start our analysis of the attack relation assignments att_{xy}, $x \in \{l, w\}$ and $y \in \{E, D\}$, with a simple and easy lemma.

Lemma 7.5. *For each basic knowledge base K, for $x \in \{l, w\}$ and $y \in \{E, D\}$, $att_{xy}(K) = att_{bs}(K)$.*

Proof. It follows immediately from Definition 7.12 that for any basic knowledge base, for any two finite sets of defeasible rules Γ, Γ', for $y \in \{E, D\}$, $\Gamma \trianglelefteq_y \Gamma'$ iff $\Gamma \neq \emptyset$ and $\Gamma' = \emptyset$. It holds immediately that if A rebuts B then $(A, B) \in att_{x,y}(K)$ for any $x \in \{l, w\}$ and $y \in \{E, D\}$. □

From Lemma 7.4, it is easy to see that the following lemma holds.

Lemma 7.6. *Every attack relation assignment att_{xy}, for $x \in \{l, w\}$ and $y \in \{E, D\}$, satisfies the properties of context-independence, subargument structure, attack closure and effective rebuts.* □

It turns out that attack relation assignments based on the elitist ordering att_{xE}, $x \in \{w, l\}$, satisfy the property of attack monotonicity but not the property of credulous cumulativity, while the situation is reverse for attack relation assignments based on the democratic ordering.

Theorem 7.8. *Let $\mathcal{K}$ be a sensible class of knowledge bases. Both attack relations assignments att_{lE} and att_{wE} satisfy the property of attack monotonicity for $\mathcal{K}$.*

Proof. See Appendix H. □

Theorem 7.9. *Let $\mathcal{K}$ be a sensible class of knowledge bases satisfying the self-contradiction property. Both attack relation assignments att_{lD} and att_{wD} satisfy the property of credulous cumulativity for $\mathcal{K}$.*

Proof. See Appendix H. □

The following lemma reveals further relationships between the attack relation assignments att_{xy}.

Lemma 7.7. $att_{lE} \sqsubseteq att_{nr} \sqsubseteq att_{lD} \sqsubseteq att_{bs}$.[43]

[43] Note that for attack relation assignments att, att', $att \sqsubseteq att'$ iff for each knowledge base K, $att(K) \subseteq att'(K)$.

Proof. See Appendix H. □

The following example shows that both semantics based on the elitist-ordering satisfy neither the consistency postulate nor the credulous cumulativity property in general.

Example 7.1. Consider a knowledge base K consisting of

1. an empty base of evidence, and
2. four strict rules

$$r_1 : a_2, a_3, a_4 \to \neg a_1 \quad \ldots\ldots\ldots \quad r_4 : a_1, a_2, a_3 \to \neg a_4$$

together with four defeasible rules

$$d_i : \Rightarrow a_i, \quad 1 \le i \le 4$$

and
3.

$$\preceq \ = \ \{d_1, d_2\} \times \{d_1, d_2\} \cup \{d_3, d_4\} \times \{d_3, d_4\}$$

It is clear that $\preceq$ is a preorder and the knowledge base is consistent and closed under transposition.
There are in total 8 arguments:

$$A_i \equiv \ [\Rightarrow a_i], \quad 1 \le i \le 4$$

and

$$B_1 \equiv \ [A_2, A_3, A_4 \to \neg a_1], \ \ldots\ldots, \ B_4 \equiv \ [A_1, A_2, A_3 \to \neg a_4]$$

We first show

$$\{d_1, d_3, d_4\} \lhd_E \{d_2\}$$

From $d_1 \preceq d_2$, it is clear that $\{d_1, d_3, d_4\} \unlhd_E \{d_2\}$. From $d_2 \not\preceq d_3$, it is obvious that $\{d_2\} \not\unlhd_E \{d_1, d_3, d_4\}$.
Similarly, it holds:

$$\{d_2, d_3, d_4\} \lhd_E \{d_1\} \quad \{d_1, d_2, d_3\} \lhd_E \{d_4\} \quad \{d_1, d_2, d_4\} \lhd_E \{d_3\}$$

Therefore B_i does not attack A_i for $1 \le i \le 4$ according to the attack relation $att_{xE}(K)$ for $x = l, w$. Therefore $att_{xE}(K) = \emptyset$.
All arguments belong to the unique stable (complete) extension whose set of conclusions is $S = \{a_1, \neg a_1, \ldots, a_4, \neg a_4\}$, which is obviously inconsistent.
Since $\Omega = \{a_1, \ldots, a_4\} \subseteq S$ is inconsistent wrt set of strict rules of K, $K + \Omega$ is not consistent. Hence the credulous cumulativity property is violated for the sensible class of knowledge bases closed under transposition. □

Note that Example 1.2 shows that attack relation assignments based on the democratic ordering principle do not satisfy the attack monotonicity property.

Theorem 7.10.

1. *Attack relations assignments based on the elitist ordering satisfy in general neither the consistency postulate nor the credulous cumulativity property for sensible classes of knowledge bases that are closed under transposition or contraposition.*
2. *Attack relations assignments based on democratic ordering satisfy in general neither the attack monotonicity property nor the property of irrelevance of redundant defaults for sensible classes of knowledge bases.*

Proof. Example 1.2 shows the second assertion.
Example 7.1 shows the first assertion for the case of closure under transposition. To show this assertion for the case of closure under contraposition, just add the absurd rules $\alpha, \neg\alpha \to l$ for each domain atom α and each domain literal l. The resulting knowledge bases are closed under contraposition (Lemma B.1 in Appendix B).
In the new knowledge base, apart from the 8 previous arguments, it should be clear that each new argument contains at least one occurrence of some absurd rule. Therefore for each new argument X, $dr(X) = ldr(X) = \{d_1, \ldots, d_4\}$.[44]

[44] To see this point, consider arguments containing exactly one occurrence of absurd rules and also as their last rule. Therefore, there are two subarguments containing no absurd rules with conclusions $a_i, \neg a_i$ for $i = 1, .., 4$. Therefore these two subarguments must be A_i, B_i for some i.

From $\{d_1, \ldots, d_4\} \lhd_E \{d_i\}$, it follows $X \sqsubset_{xE} A_i$, $i = 1, \ldots, 4$, $x = l, w$. Hence X does not attack any other argument. The unique stable belief set is still the same like before. □

There could be more than one way to interpret the elitist and democratic principles leading to different orderings between arguments. One could for example strengthen slightly the elitist principle as in the following example.

Definition 7.15 *(Strengthened elitist principle).*

1. Let Γ, Γ' be sets of defeasible rules. We say that Γ is less preferred than Γ' wrt the strengthened elitist principle, denoted by $\Gamma \unlhd_{sE} \Gamma'$, iff
   - $\Gamma \unlhd_E \Gamma'$ and
   - if $\Gamma' \neq \emptyset$ then $\exists d \in \Gamma\ \exists d' \in \Gamma'$: $d \prec d'$.
2. We say that an argument A is less preferred than an argument B wrt the last link principle and the strengthened elitist principle, denoted by $A \sqsubseteq_{lsE} B$, iff $ldr(A) \unlhd_{sE} ldr(B)$.
3. For each knowledge base K, $att_{lsE}(K) \subseteq AR_K \times AR_K$ is an attack relation defined by: $(A, B) \in att_{lsE}(K)$ iff one of the following conditions holds:
   - A rebuts B (at B') such that $A \not\sqsubseteq_{lsE} B'$.
   - A undercuts B. □

Let us consider again the Example 7.1. It is not difficult to see that the following assertions hold:

$$\{d_1, d_3, d_4\} \not\unlhd_{sE} \{d_2\}, \{d_2, d_3, d_4\} \not\unlhd_{sE} \{d_1\}, \{d_1, d_2, d_3\} \not\unlhd_{sE} \{d_4\}, \{d_1, d_2, d_4\} \not\unlhd_{sE} \{d_3\}.$$

Therefore B_i attacks A_i for each $i = 1, \ldots, 4$. Therefore there are three stable extensions that are also consistent.

Lemma 7.8. $\unlhd_{sE} \subset \unlhd_E$.

Proof. It is obvious that $\unlhd_{sE} \subseteq \unlhd_E$. In Example 7.1, it is clear that $\{d_1, d_3, d_4\} \unlhd_E \{d_2\}$. As elaborated above, it holds that $\{d_1, d_3, d_4\} \not\unlhd_{sE} \{d_2\}$. Therefore $\unlhd_{sE} \subset \unlhd_E$. □

Lemma 7.9. *Let $\mathcal{K}$ be a sensible class of knowledge bases satisfying the self-contradiction property. Further let $K \in \mathcal{K}$ and E be a stable extension wrt the attack relation assignment att_{lsE}. Then cnl(E) is consistent and closed.*

Proof. See Appendix H. □

Prakken in [39] has studied an even stronger version of elitist principle where in Definition 7.12 the preorder relation $\preceq$ is replaced by its strict order $\prec$, i.e. Γ is said to be less preferred than Γ' wrt the strict elitist principle, denoted by $\Gamma \unlhd_{\sigma E} \Gamma'$, iff $\Gamma' = \emptyset$ or $\Gamma' \neq \emptyset$ and $\exists d \in \Gamma\ \forall d' \in \Gamma'$: $d \prec d'$.

It is not difficult to see that $\unlhd_{\sigma E}$ is a subset of $\unlhd_{sE}$. To see that there is a proper-subset-relationship, consider the example where $\{d_0\} \unlhd_{sE} \{d_0, d_1\}$ given $d_0 \prec d_1$, but $\{d_0\} \not\unlhd_{\sigma E} \{d_0, d_1\}$.

It is proved in [39] that strict elitism satisfies the consistency postulate when combined with both the last link and weakest link principles. We leave for future works the questions concerning other properties of the attack relation assignments for both strengthened and strict elitist principles.

The discussion in this chapter shows that ASPIC+ is a fertile framework for studying the semantics of structured argumentation in which a rich and diverse set of argument orderings could be defined. The ordinary properties proposed in this paper could be used for evaluating and classifying the attack relations obtained from these argument orderings.

8. Discussion and conclusion

We have presented in this paper an axiomatic analysis of semantics for structured argumentation both with and without preferences between defeasible rules by giving a set of simple and intuitive properties that could be used to analyze the attack relations underlying the semantics of structured argumentation. We have shown that the normal attack relation assignment could be viewed as a normal form for ordinary attack relation assignments in the sense that the normal attack relation assignment is ordinary and the stable extensions wrt any ordinary attack relation assignments are also stable extensions wrt the normal attack relation assignment. This insight suggests that one can study the stable semantics of ordinary attack relation assignment by looking at the stable semantics of the normal attack relation assignment.

Many distinct interpretations of preferences between defaults have been proposed and explored intensively in the literature [8,6,10,15,34,35]. In this paper, we follow a minimalist approach and embrace a view of the preference between defaults that is shared by the other interpretations. This suggests that the semantics for any more elaborated interpretations of default preferences could be characterized by a subset of the set of stable extensions wrt normal attack relation assignments. Theorem 6.2 suggests that an axiomatic characterization for the new interpretation could be obtained by adding new properties to the axioms of the ordinary semantics presented in this paper.

Gelfond and Son [24] proposed to represent prioritized default reasoning in answer set logic programming. Similarly to our view, Gelfond and Son [24] consider the role of preferences of defaults as a tool for conflict resolution where the knowledge base should specify explicitly which rule is to be picked when two rules are applicable but where accepting both leads to contradiction. Gelfond and Son's approach could be viewed as a kind of "off-the-shelf" programming solution that could be easily implemented in logic programming. Gorogiannis and Hunter, Amgoud and Besnard as well as Hunter and Woltran [25,2,1,28,45] studied the structure of arguments and their semantics based on classical or abstract Tarski logics. The relationship between these approaches and the ASPIC+ framework has also been studied in [34]. Garcia and Simari's defeasible logic programming [21] studied attacks based on the specificity principle while Bondarenko, Dung, Kowalski and Toni [7] considered assumption-based attacks.

An interesting challenge is to study the possibilities to integrate our approach with these approaches. A combination of our approach with Gelfond and Son's one could provide a methodology for developing prioritized default reasoning where guidelines for the conflict relations should be provided to capture the normal semantics. Such combinations would be both ordinary and suitable for quick deployment in applications. Some interesting results have been presented in Lemma 7.4 showing that the ASPIC+-attack relations already embody in them self key features of ordinary semantics. The results in section 7.4 suggests that there is much in common between ASPIC+ semantics and the ordinary semantics. Though the formal interpretation of the elitist and democratic principles in [34,35] are appealing, a deeper analysis of these principles may be helpful to an integrated framework for structured argumentation.

In this paper, we focus our attention on stable semantics. We believe many of the results in this paper also hold for other semantics like complete extension semantics. For example, the credulous cumulativity property should hold wrt complete extension semantics for the same reasons it holds wrt stable semantics. We show below that it is indeed the case. We first adapt the definition of credulous cumulativity property for complete extension semantics.

Definition 8.1 *(Credulous cumulativity wrt complete extensions).* Let $\mathcal{K}$ be a sensible class of knowledge bases. An attack relation assignment att is said to satisfy the property of credulous cumulativity for $\mathcal{K}$ wrt complete extension semantics if and only if for each $K \in \mathcal{K}$, for each complete belief set S of K wrt *att* and for each finite subset $\Omega \subseteq S$ of literals,

1. $K + \Omega$ belongs to $\mathcal{K}$ and
2. S is a complete belief set of $K + \Omega$ wrt *att*. □

An identical proof of Lemma 4.1 where references to stable extensions are simply replaced by complete extensions shows that credulous cumulativity implies both the consistency and closure postulates for complete extensions.

The following theorem shows that the credulous cumulativity property wrt complete extension semantics is also satisfied by the normal attack relation assignment.

Theorem 8.1. *Let $\mathcal{K}$ be a sensible class of knowledge bases that are closed under transposition or contraposition. Then the normal attack relation assignment att_{nr} satisfies the credulous cumulativity property for $\mathcal{K}$ wrt complete extensions semantics.*

Proof. See Appendix J. □

We conclude this section with a discussion of the relationship between our notion of credulous cumulativity and Gabbay's "skeptical" version of cumulativity [20]. Skeptical cumulativity has been studied extensively by Kraus, Lehman and Magidor [29], Brewka [9], Geffner and Pearl [26], Dung and Son [18]. Skeptical cumulativity intuitively states that adding a skeptical conclusion to the knowledge base does not change the other conclusions. Brewka [9] generalized and adapted the idea of skeptical cumulativity to default logic. We adopt Brewka's concept as the semantics of both default logic and our knowledge bases are based on the notion of extensions.

A stable belief set could be viewed as representing a possible world given a background knowledge base K. In this spirit, we could say that a set Ω of domain literals is *predictable* wrt K iff Ω is a subset of some stable belief set of K.

An expansion $K + \Omega$ of K by a predictable Ω is then said to be a *predictable expansion* of K. Skeptical cumulativity could be viewed as stating that the stable belief sets of predictable expansions of a knowledge base K coincide with the stable belief sets of K. This intuition is formalized in the following definition adapted from Brewka's formulation of skeptical cumulativity for default logics [9].

Definition 8.2. We say that an attack relation assignment *att* satisfies the skeptical cumulativity property for a sensible class of knowledge bases $\mathcal{K}$ iff

1. all predictable expansions of knowledge bases in $\mathcal{K}$ belong to $\mathcal{K}$, and

2. for each predictable expansion $K + \Omega$ of $K \in \mathcal{K}$, the set of stable belief sets of $K + \Omega$ coincides with the set of stable belief set S of K where $\Omega \subseteq S$. □

It is obvious that skeptical cumulativity implies credulous cumulativity. But the reverse does not hold as the example below shows.

Example 8.1. Consider a basic knowledge base $K = (RS, RD, \preceq, BE)$ where $RS = \emptyset$, $\preceq = \emptyset$, $BE = \emptyset$ and RD consists of three defeasible rules

$$d_1 : \Rightarrow a \quad d_2 : \Rightarrow \neg a \quad d_3 : \ a \Rightarrow b$$

It is clear that $S = \{a, b\}$ is a stable belief set of K wrt attack relation $att_{bs}(K)$. It is also clear that S is a stable belief set of $K' = K + \{b\}$. But K' has another stable belief set $\{\neg a, b\}$ (wrt $att_{bs}(K')$) that is not a stable belief set of K. Therefore the skeptical cumulativity property does not hold for the basic attack relation assignment att_{bs}. □

Appendix A. Recall proof of Lemma 3.1

For a strict argument A over a set of domain literals $X \subseteq \mathcal{L}$, the *set of premises of* A, denoted by $Prem(A)$, is the set of conclusions of subarguments of A of the form $[\alpha]$, $\alpha \in X$.

Lemma A.1. *Let K be a knowledge base closed under contraposition or transposition and A be a strict argument (wrt $K + X$) with conclusion σ and $Prem(A) \subseteq X$. Then for each $\alpha \in Prem(A)$, there is a strict argument B (wrt $K + (X \cup \{\neg\sigma\})$) with $Prem(B) \subseteq Prem(A) \cup \{\neg\sigma\}$ and conclusion $\neg\alpha$.*

Proof. If K is closed under contraposition, the lemma is obvious. We prove the lemma for the case of closure under transposition by induction on the structure of A.

Base Case: $A = [\alpha]$, $\alpha \in X$. Obvious.

Inductive Case: Suppose A is of the form $[A_1, \ldots, A_n \to \sigma]$ where $Cnl(A_i) = \alpha_i$. Let $\alpha \in Prem(A)$. Without loss of generality, let $\alpha \in Prem(A_n)$. From the closure under transposition, the rule $\alpha_1, \ldots, \alpha_{n-1}, \neg\sigma \to \neg\alpha_n$ also belongs to RS. Let B be the argument $A_1, \ldots, A_{n-1}, \neg\sigma \to \neg\alpha_n$.

From the induction hypothesis, there is an argument T whose premises are in $Prem(A_n) \cup \{\neg\alpha_n\}$ and whose conclusion is $\neg\alpha$.

Let T' be the argument obtained from T by replacing each occurrence of premise $\neg\alpha_n$ by the argument B. It is clear that $Prem(T') \subseteq Prem(A) \cup \{\neg\sigma\}$ and $Cnl(T') = \neg\alpha$. □

Lemma 3.1. *If K is closed under transposition or contraposition then K satisfies the self-contradiction property.*

Proof. Let X be a minimal inconsistent set of domain literals. Since X is inconsistent, there is a domain literal λ such that $X \vdash_K \lambda$ and $X \vdash_K \neg\lambda$.

- Let K be closed under contraposition. Let $x \in X$. It is clear $\{x, \lambda\} \vdash_K \lambda$. Since K is closed under contraposition, it follows obviously $\{\lambda, \neg\lambda\} \vdash_K \neg x$. Therefore $X \vdash_K \neg x$.
- Let K be closed under transposition. There are two arguments A_0, A_1 with premises in X and conclusions $\lambda, \neg\lambda$ respectively. From the minimality of X, it holds: $X = Prem(A_0) \cup Prem(A_1)$. Let $x \in X$. Without loss of generality, suppose $x \in Prem(A_0)$. From the Lemma A.1, it follows that there exists an argument B with conclusion $\neg x$ and $Prem(B) \subseteq Prem(A_0) \cup \{\neg\lambda\}$. Let A be the argument obtained by replacing each subargument of the form $[\neg\lambda]$ in B by argument A_1. It is clear that $Prem(A) \subseteq X$ and the conclusion of A is $\neg x$. □

Appendix B

Lemma B.1. *Let RS consist of rules of the forms*

$$a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \to \neg a_i$$

and

$$a_i, \neg a_i \to l$$

where $1 \le i \le n$ and l is a literal over $a_1, \ldots, a_n$.

Then RS is closed under contraposition.

Proof. Let S be a set of literals and $S \vdash \lambda$ and $\sigma \in S$. We show $S \setminus \{\sigma\} \cup \{\neg\lambda\} \vdash \neg\sigma$. There are two cases:

1. S is consistent.
   (a) $\lambda \in S$.
   If $\sigma = \lambda$ then we are done. Suppose $\sigma \neq \lambda$. Then $\{\lambda, \neg\lambda\} \subseteq S \setminus \{\sigma\} \cup \{\neg\lambda\}$. We are done.
   (b) $\lambda \notin S$.
   Because S is consistent and $\lambda \notin S$, λ must be derived from S using a rule of the form $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \to \neg a_i$. Without loss of generality, we could assume $\lambda = \neg a_1$. Therefore $S = \{a_2, \ldots, a_n\}$. We are done.
2. S is inconsistent.
   (a) There is i: $\{a_i, \neg a_i\} \subseteq S$.
   - $\sigma \in \{a_i, \neg a_i\}$.
     Obviously $\neg\sigma \in S \setminus \{\sigma\} \cup \{\neg\lambda\}$. We are done.
   - $\sigma \notin \{a_i, \neg a_i\}$.
     Obviously $\{a_i, \neg a_i\} \subseteq S \setminus \{\sigma\} \cup \{\neg\lambda\}$. We are done as absurd rules derive any thing.

   (b) There is no i: $\{a_i, \neg a_i\} \subseteq S$.
   Since S is inconsistent, there is i such that there are strict arguments over S supporting $a_i, \neg a_i$ without using absurd rules. The only such arguments are $[a_i]$ and $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \to \neg a_i$. Therefore $S = \{a_1, a_2, \ldots, a_n\}$.
   - $\lambda = a_i$.
     If $\sigma \neq a_i$, obviously $\{a_i, \neg a_i\} \subseteq S \setminus \{\sigma\} \cup \{\neg\lambda\}$. We are done. Suppose $\sigma = a_i$, we are also done obviously.
   - $\lambda = \neg a_i$.
     If $\sigma = a_i$, we are done. If $\sigma = a_j \neq a_i$, we have $S \setminus \{\sigma\} \cup \{\neg\lambda\} = \{a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n\} \vdash \neg a_j$. We are done. □

Appendix C

Lemma 4.2. *Let K be a knowledge base, $A \in AR_K$ and $\Omega \subseteq BE$ be finite. It holds that X' is a subargument of an argument $X \in A \uparrow \Omega$ iff there exists a subargument A' of A such that $X' \in A' \uparrow \Omega$.*

Proof. We prove by induction on the structure of A.

1. $A = [\alpha]$, $\alpha \in BE$. The lemma holds obviously.
2. $A = [A_1, \ldots, A_n, r]$.
   - "Only IF". There are two cases:
     Case 1: $X = [\alpha]$, $\alpha \in BE$. Hence $X' = X$. The lemma holds obviously for $A' = A$.
     Case 2: $X = [X_1, \ldots, X_n, r]$ for $X_i \in A_i \uparrow \Omega$. If $X' = X$ then just set $A' = A$. We are done.
     Suppose $X' \neq X$. Hence X' is a subargument of some X_i, say X_1. From induction hypothesis, there is some subargument A'_1 of A_1 such that $X' \in A'_1 \uparrow \Omega$. Set $A' = A'_1$, we are done.
   - "IF" If $A' = A$, we are done.
     Suppose $A' \neq A$. Hence A' is a subargument of some A_i, say A_1. From induction hypothesis, X' is subargument of some $X \in A_1 \uparrow \Omega$. Hence X' is a subargument of $Y = [X, A_2, \ldots, A_n, r]$. It is clear $Y \in A \uparrow \Omega$. □

Theorem 4.1. *Let $\mathcal{K}$ be a sensible class of knowledge bases such that for each $K = (RSB, BE) \in \mathcal{K}$, for each evidence $\omega \in BE$, $K + d_\omega$ belongs to $\mathcal{K}$. Further let att be an attack relation assignment satisfying the properties of attack monotonicity and context-independence for $\mathcal{K}$. Then att also satisfies the property of irrelevance of redundant defaults for $\mathcal{K}$.*

Proof. Let K be a knowledge base and $\omega \in BE$ and $K' = K + d_\omega$. For each argument $X \in AR_{K'}$, let $st(X)$ be the argument obtained from X by replacing each occurrence of defeasible rule d_ω in X by ω. It is clear that $st(X) \in X \uparrow \omega$. For any set of arguments $AS \subseteq AR_{K'}$, let $st(AS) = \{st(X) \mid X \in AS\}$.

1. We show that the stable belief sets of K, K' coincide.
   (a) Let E be a stable extension of K and $S = cnl(E)$. Due to the context-independence, E is conflict-free wrt $att(K')$.
   We show that S is a stable belief set of K'.
   Let $E' = \{X \in AR_{K'}$ s.t. $st(X) \in E$ and $E \cup \{X\}$ is conflict-free wrt att(K')$\}$. It is obvious that $E \subseteq E'$ and $S = cnl(E')$.
   We show that E' is stable extension of $(AR'_K, att(K'))$.
   We first show E' is conflict-free wrt att(K'). Suppose there are $X, Y \in E'$ s.t. X attacks Y wrt $att(K')$. From the attack monotonicity, $st(X)$ attacks Y wrt att(K') implying that E attacks Y wrt $att(K')$. Contradiction since $E \cup \{Y\}$ is conflict-free wrt $att(K')$.
   We show that E' attacks each argument in $AR_{K'}$ not belonging to it. Let $X \in AR_{K'} \setminus E'$. From the definition of E', it follows immediately that $st(X) \notin E$ or $E \cup \{X\}$ is not conflict-free wrt att(K').
   If $st(X) \notin E$ then there is $A \in E$ s.t. A attacks st(X) wrt att(K). Due to the context independence, A attacks st(X) wrt att(K'). Due to attack monotonicity, A attacks X wrt att(K'). Hence E' attacks X wrt att(K').

Suppose now $st(X) \in E$ and $E \cup \{X\}$ is not conflict-free wrt att(K'). If X does not attack E wrt $att(K')$ then E attacks X wrt $att(K')$ since $E \cup \{X\}$ is not conflict-free wrt att(K'). We are done.
Suppose X attacks E wrt att(K'). Therefore st(X) attacks E wrt att(K'). Thus st(X) attacks E wrt att(K) due to the context-independence. Contradiction to the fact that $st(X) \in E$ and E is conflict-free. This case hence does not occur.

(b) Let E' be a stable extension of $(AR'_K, att(K'))$. Therefore $st(E') \subseteq E'$ (otherwise E' attacks $st(E')$ wrt att(K'). Due to attack-monotonicity, E' attacks itself wrt att(K'), a contradiction). Let $E = st(E')$. E is hence conflict-free wrt att(K'). Due to the context-independence, E is hence conflict-free wrt att(K).
Let $X \in AR_K \setminus E$. We show that E attacks X wrt att(K). From $st(X) = X$, it follows $X \notin E'$. Therefore E' attacks X wrt att(K'). From attack monotonicity and $E = st(E')$, E attacks X wrt att(K'). From the context-independence, E attacks X wrt att(K).
E hence attacks every argument in AR_K not belonging to E wrt att(K). E is stable wrt att(K).

2. We show that the complete belief sets of K, K' (wrt att) coincide.

(a) Let E be a complete extension of $(AR_K, att(K))$. Due to the context-independence, E is conflict-free wrt $att(K')$.

Property 1. *For each $A \in AR_K$, A is defended by E wrt att(K) iff A is defended by E wrt att(K').*

Proof. "⇐" Let A be defended by E (wrt $att(K')$). Let $X \in AR_K$ attack A wrt $att(K)$. Due to the context-independence, X also attacks A wrt $att(K')$. Therefore E attacks X wrt $att(K')$. Due to the context independence, E attacks X wrt $att(K)$. We proved that A is defended by E wrt att(K).
"⇒" Let A be defended by E (wrt $att(K)$). Let $X \in AR_{K'}$ attack A wrt $att(K')$. Due to the attack monotonicity, $st(X)$ also attacks A wrt $att(K')$. Due to the context independence, $st(X)$ also attacks A wrt att(K). Hence E attacks st(X) wrt att(K). Due to the context independence, E attacks st(X) wrt $att(K')$. Therefore E attacks X wrt $att(K')$ (due to attack monotonicity). We proved that A is defended by E wrt att(K'). □

Property 2. *Let $A \in AR_{K'}$ such that A is defended by E wrt att(K'). Then st(A) is defended by E wrt att(K').*

Proof. Let $X \in AR_{K'}$ attack st(A) wrt att(K'). Due to the attack monotonicity, $st(X)$ also attacks st(A) wrt $att(K')$. Hence st(X) attacks A wrt att(K') due to the attack monotonicity. Therefore E attacks st(X) wrt att(K'). Due to attack monotonicity, E attacks X wrt att(K'). □

It follows immediately from Properties 1, 2: For each $A \in AR_{K'}$, if A is defended by E wrt att(K') then st(A) is defended by E wrt att(K) implying that $st(A) \in E$. Let $E' = E \cup \{A \in AR_{K'}$ s.t. A is defended by E wrt att(K')}. From Property 1, each argument in E' is defended by E wrt att(K'). It is also clear $E' \subseteq E \cup \{A \in AR_{K'} \mid st(A) \in E\}$. Hence $st(E') \subseteq E \subseteq E'$.
We show that E' is conflict-free wrt att(K'). Suppose $\exists X, Y \in E' : (X, Y) \in att(K')$. Since Y is defended by E wrt att(K'), E attacks X wrt att(K'). Since X is also defended by E wrt att(K'), E attacks E wrt att(K'). From context-independence, E attacks E wrt att(K). Contradiction.
From $st(E') \subseteq E$ and the attack monotonicity, it follows immediately that any argument in $AR_{K'}$ attacked by E' wrt $att(K')$ is also attacked by E wrt $att(K')$. Hence any argument in $AR_{K'}$ defended by E' wrt $att(K')$ is also defended by E wrt $att(K')$ and hence belongs to E'. E' is a complete extension of $(AR_{K'}, att(K'))$. From $st(E') \subseteq E \subseteq E'$, it follows immediately $Cnl(E') = Cnl(E)$.

(b) Let E' be a complete extension of $(AR'_K, att(K'))$. As each attack against $st(E')$ wrt att(K') is an attack against E' wrt att(K'), all attacks against $st(E')$ (wrt att(K')) are counter-attacked (wrt att(K')) by E'. Therefore $st(E') \subseteq E'$.
Let $E = st(E') \subseteq AR_K$.
We show that E is a complete extension of $(AR_K, att(K))$.
Due to the context independence and $E \subseteq E'$, it is clear that E is conflict-free.
Let $A \in AR_K$ be defended by E wrt att(K).
We first show A is defended by E' wrt att(K').
Let $X \in AR_{K'}$ attack A wrt att(K'). Therefore st(X) attacks A wrt att(K') due to the attack monotonicity. Due to the context-independence, st(X) attacks A wrt att(K). Since A is defended by E wrt att(K), st(X) is attacked by E wrt att(K). From $E = st(E')$, st(X) is attacked by $st(E')$ wrt att(K). From the context-independence, st(X) is attacked by $st(E')$ wrt att(K'). From $st(E') \subseteq E'$, E' attacks st(X) wt att(K'). From attack monotonicity, E' attacks X wrt att(K').
Since E' is complete, $A \in E'$. Hence $A \in E$. We proved that E is complete. □

Appendix D

Lemma D.1. *Let $\mathcal{K}$ be a sensible class of knowledge bases and att be an attack relation assignment satisfying the basic axioms for $\mathcal{K}$. Further let $K \in \mathcal{K}$ and E be a stable extension of $(AR_K, att(K))$ and $B \in AR_K \setminus E$ such that B is not undercut by any argument in E.*

Then there is a subargument X of B such that

1. $X \notin E$ *and all proper subarguments of X belong to E, and*
2. *for each argument* $A \in E$*, if* $(A, X) \in att(K)$ *then A directly rebuts X.*

Proof. Since E is stable extension, E attacks B. From Definition 4.1, it is obvious that B is defeasible. Let SU be the set of subarguments of B not belonging to E. Since $B \notin E$, SU is not empty. As there exists no infinite sequence of arguments $(A_i)_i$ such that A_{i+1} is a proper subargument of A_i, there is no infinite sequence of arguments $(B_i)_i$ such that for each i, $B_i \in SU$ and B_{i+1} is a proper subargument of B_i. Therefore there exists $X \in SU$ such that no proper subarguments of X belong to SU. Hence all proper subarguments of X belong to E. From $X \notin E$, and E is a stable extension, E attacks X wrt att(K).

Let $A \in E$ such that $(A, X) \in att(K)$. Since X is a subargument of B, and B is not undercut by any argument in E, X is not undercut by any argument in E. Hence X is not undercut by A.

We first prove that X is basic defeasible. Suppose the contrary. Therefore $X = [X_1, \dots, X_n, r]$ and r is a strict rule. Since all proper subarguments of X belong to E, $\forall i : cnl(X_i) \in cnl(E)$. From Lemma 4.1, $cnl(E)$ is closed. Therefore $cnl(X) \in cnl(E)$. Because $(A, X) \in att(K)$, $X = [X_1, \dots, X_n, r]$ and r is a strict rule and for every i, X_i and all subarguments of X_i belong to E, A contradicts X (at X) (Definition 4.10). Hence $cnl(E)$ is contradictory. Contradiction. We have proved that X is basic defeasible.

Because $(A, X) \in att(K)$, all proper subarguments of X belong to E and A does not undercut X, it is clear that A directly rebuts X. $\square$

Appendix E

Lemma E.1 *(Reduced general characteristic lemma). Let att be an ordinary attack relation assignment defined for a sensible class* $\mathcal{K}$ *of knowledge bases. Further let* $K \in \mathcal{K}$ *and E be a stable extension of* $(AR_K, att(K))$ *and* $B \in AR_K \setminus E$ *such that B is not undercut by any argument in E.*

Then there is a subargument X of B such that

1. $X \notin E$ *and all proper subarguments of X belong to E, and*
2. *for each argument* $A \in E$ *such that* $(A, X) \in att(K)$*, following properties hold:*
   (a) *A directly rebuts X.*
   (b) *(**Normal condition**) There is no defeasible rule* $d \in ldr(A)$ *such that* $d \prec last(X)$*.*

Proof. The proof proceeds in two steps. In the first step, we introduce a helpful notation and prove a lemma. The main proof is given in the second step.

1. For simplicity, for any arbitrary defeasible rule d of the form $b_1, \dots, b_n \Rightarrow h$ with $bd(d) \subseteq BE$, define $[|d|]$ to be of the form $[[b_1], \dots, [b_n], d]$.
   It is clear that for any basic defeasible argument $B = [B_1, \dots, B_n, d]$, B is a weakening of $[|d|]$ (if $[|d|]$ is defined) by $\{B_1, \dots, B_n\}$. The following lemma follows immediately from the link-oriented property.

   Lemma E.2. *Let K be knowledge base and* $A, B \in AR_K$ *such that*
   (a) *B is basic defeasible, and*
   (b) $[|last(B)|] \in AR_K$*, and*
   (c) *A does not attack any proper subargument of B wrt att(K).*
   It holds that if $(A, B) \in att(K)$ then $(A, [|last(B)|]) \in att(K)$. $\square$

2. From Lemma D.1, it follows immediately that there is a subargument X of B s.t. conditions 1 and 2a hold. We only need to show the NC condition.
   Let $(A, X) \in att(K)$ and $A \in E$. Hence A directly rebuts X. Let $d_X = last(X)$.
   Suppose there is $d \in ldr(A)$ s.t. $d \prec d_X$.
   Let $\Delta = ldr(A) \setminus \{d\}$ and $\Omega = bd(d) \cup hd(\Delta) \cup bd(d_X)$. From Lemma 4.3, $mbd(A) \subseteq E$. From the definition of X, $bd(d_X) \subseteq E$. Therefore $\Omega \subseteq cnl(E)$. Let $K' = K + \Omega$.
   Let D be the argument obtained from A by replacing each argument $Z \in mbd(A)$ whose last default belongs to Δ (i.e. $last(Z) \in \Delta$) by its conclusion $cnl(Z)$ and replacing each argument $Y \in mbd(A)$ whose last default is d (i.e. $last(Y) = d$) by $[|d|]$.
   It is clear that $D \in AR_{K'}$ and $cnl(D) = cnl(A)$. It is also clear that $X \in AR_{K'}$. From $(A, X) \in att(K)$ and the context-independence of att, it follows $(A, X) \in att(K')$. It is not difficult to see that $D \in A \uparrow \Omega$. Therefore $(D, X) \in att(K')$ following the property of attack monotonicity.
   From $cnl(D) = cnl(A)$ and the fact that all proper arguments of X belong to E and the conflict-freeness and consistency of E and the property of attack closure, it is clear that D does not attack any proper subargument of X.

Since $bd(d_X) \subseteq \Omega$, it holds $[|d_X|] \in AR_{K'}$. From the effective rebut property, $(D, [|d_X|]) \notin att(K')$. It is clear that X is a weakening of $[|d_X|]$. Since D does not attack any proper subargument of X wrt $att(K')$ and $(D, X) \in att(K')$, $(D, [|d_X|]) \in att(K')$ (following Lemma E.2). Impossible since $(D, [|d_X|]) \notin att(K')$. □

Appendix F. Credulous cumulativity of att_{nr}

Lemma F.1. *Let $X \in A \downarrow AS$ and X' be a subargument of X such that X' is not a subargument of any argument in AS. Then there exists a subargument A' of A s.t. $X' \in A' \downarrow AS$.*

Proof. We prove by induction on the structure of A.

1. $A = [\alpha]$, $\alpha \in BE$. Therefore $A \downarrow AS = \{[\alpha]\} \cup \{Z \in AS \mid cnl(Z) = \alpha\}$. Since X' is not a subargument of any argument in AS, $X' = [\alpha]$ and $[\alpha] \notin AS$. Let $A' = A$. It is clear that $X' \in A' \downarrow AS$.
2. $A = [A_1, \ldots, A_n, r]$. Therefore $X = [X_1, \ldots, X_n, r]$ for $X_i \in A_i \downarrow \Omega$. If $X' = X$ then just set $A' = A$. We are done. Suppose $X' \neq X$. Hence X' is a subargument of some X_i, say X_1. From induction hypothesis, there is some subargument A'_1 of A_1 such that $X' \in A'_1 \downarrow AS$. Set $A' = A'_1$, we are done. □

From the definition of normal attack relation assignment and Lemmas 4.3, 6.4 it is easy to see that the following lemma holds.

Lemma F.2. *For each knowledge base K, for each complete extension E of $(AR_K, att_{nr}(K))$ and each argument $A \in AR_K$, it holds that $A \in E$ iff each maximal basic defeasible subargument of A belongs to E (i.e. $mbd(A) \subseteq E$).* □

Lemma F.3. *Let $\mathcal{K}$ be a sensible class of knowledge bases that satisfy the property of self-contradiction. Then for each knowledge base $K \in \mathcal{K}$, each complete extension E of $(AR_K, att_{nr}(K))$, $cnl(E)$ is closed.*

Proof. Let E be a complete extension of $(AR_K, att_{nr}(K))$ and $S = cnl(E)$.

Let $S \vdash \alpha$. If $\alpha \in S$, we are done.

Suppose $\alpha \notin S$. Therefore there is a finite $\Omega \subseteq S$ of domain literals s.t. $\Omega \vdash \alpha$. Let A_0 be a strict argument over Ω wrt the set of strict rule RS of K. It is clear $A_0 \in AR_{K+\Omega}$.

Let A be a weakening of A_0 by replacing each subargument of the form $[\alpha]$, $\alpha \in \Omega$, in A_0 by an argument $X_\alpha \in E$ s.t. $cnl(X_\alpha) = \alpha$. Therefore $A \in AR_K$. It is clear each argument in $mbd(A)$ is a subargument of some argument in $\{X_\alpha \mid \alpha \in \Omega\}$. From Lemma 4.3, $mbd(A) \subseteq E$. From Lemma F.2, $A \in E$. Therefore $\alpha \in S$. Impossible since we assume that $\alpha \notin S$. This case hence cannot occur.

We have proved that S is closed. □

Lemma F.4. *Let $\mathcal{K}$ be a sensible class of knowledge bases that satisfy the property of self-contradiction.*
Then for each knowledge base $K \in \mathcal{K}$, each complete extension E of $(AR_K, att_{nr}(K))$, $cnl(E)$ is consistent.

Proof. Let E be a complete extension of $(AR_K, att_{nr}(K))$ and $S = cnl(E)$. From Lemma F.3, it is obvious that S is closed.

We next show the consistency of S. Suppose the contrary. Since S is closed, S is contradictory. Thus there are two arguments $A, B \in E$ such that $cnl(B) = \neg cnl(A)$. Let $AS = mbd(A) \cup mbd(B) \subseteq E$. Therefore $cnl(AS) \cup BE$ is inconsistent (wrt strict rules RS of K).

Let M be a minimal inconsistent subset of $cnl(AS) \cup BE$. Since BE is consistent, $M \cap cnl(AS) \neq \emptyset$. Let $AS_0 \subseteq AS$ such that $cnl(AS_0) = M \cap cnl(AS)$. Hence $AS_0 \neq \emptyset$.

Let $\Delta_0 = \{last(X) \mid X \in AS_0\}$ and $C \in AS_0$ such that $last(C)$ be minimal wrt $\prec$ in Δ_0. From self-contradiction property, it follows $M \vdash \neg hd(last(C))$. Thus there is an argument Y s.t. $mbd(Y) \subseteq AS_0$ and $cnl(Y) = \neg hd(last(C))$. From Lemma F.2, it follows $Y \in E$. It is clear that $ldr(Y) \subseteq \Delta_0$. Therefore, there is no default $d' \in ldr(Y)$ s.t. $d' \prec last(C)$. Therefore $(Y, C) \in att_{nr}(K)$.

From $C \in AS_0 \subseteq AS \subseteq E$, it follows E is not conflict-free wrt $att_{nr}(K)$. Impossible. □

Lemma 6.6. *Let $\mathcal{K}$ be a sensible class of knowledge bases that satisfy the property of self-contradiction. Then the normal attack relation assignment att_{nr} satisfies the credulous cumulativity property for $\mathcal{K}$.*

Proof. Let $K \in \mathcal{K}$ and E be a stable extension of $(AR_K, att_{nr}(K))$, $S = cnl(E)$ and $\Omega \subseteq S$ be a finite set of domain literals. Further let $E' = E \uparrow \Omega = \bigcup_{A \in E} A \uparrow \Omega$ and $K' = K + \Omega$. It is clear that $E \subseteq E' \subseteq AR_{K'}$, $cnl(E) = cnl(E')$ and $BE \cup \Omega \subseteq S$.

We show that E' is a stable extension of $(AR_{K'}, att_{nr}(K'))$.

From Lemma F.4, it follows S is consistent. S is hence not contradictory.

As E is conflict-free wrt $att_{nr}(K)$, E' is also free from the undercut-attacks. From Lemma 4.2 and Lemma 4.3, it follows that E' contains all subarguments of its arguments. Since $S = cnl(E')$ is not contradictory and E' contains all subarguments of its arguments, E' is free from rebuts. Therefore E' is conflict-free wrt $att_{nr}(K')$.

Property. *Let $A = [X_1, \ldots, X_n, r] \in AR_{K'}$ and r be a strict rule. It holds: A belongs to E' iff $\forall i : X_i \in E'$.*

Proof. Since E' contains all subarguments of its arguments, we only need to prove the "if"-direction.

Suppose $\forall i : X_i \in E'$. Therefore $\forall i \exists Y_i \in E : X_i \in Y_i \uparrow \Omega$. Let $B = [Y_1, \ldots, Y_n, r]$. From Lemma F.2, $mbd(B) \subseteq E$. Therefore $B \in E$ (also from Lemma F.2). From $A \in B \uparrow \Omega$, it follows immediately $A \in E'$. □

Let $Z \in AR_{K+\Omega} \setminus E'$. We show that E' attacks Z wrt $att_{nr}(K')$.

It is clear that there exists a subargument X of Z such that all proper subarguments of X belong to E' and $X \notin E'$. From $S = cnl(E) = cnl(E')$ and $BE \cup \Omega \subseteq S$ and the above property, it follows $X = [X_1, \ldots, X_n, d]$ and d is defeasible and $\forall i : X_i \in E'$.

We show that E' attacks X wrt $att_{nr}(K')$.

There are two cases:

Case 1: $X \in AR_K$. Then $X \notin E$. Hence E attacks X wrt $att_{nr}(K)$. From Lemma 6.4, att_{nr} satisfies the context-independence property. Hence E' attacks X wrt $att_{nr}(K')$.

Case 2: $X \notin AR_K$. Therefore $X \in AR_{K+\Omega} \setminus AR_K$. Let Ω' be a minimal subset of Ω such that $X \in AR_{K+\Omega'}$. Further let AS be a minimal subset of E such that $cnl(AS) = \Omega'$.

Let Y be obtained from X by replacing each subargument $[\alpha]$, $\alpha \in \Omega'$, of X by an argument $A_\alpha \in AS$ where $cnl(A_\alpha) = \alpha$. It is clear $Y \in AR_K$ and $Y \in X \downarrow AS$. Therefore $X \in Y \uparrow \Omega'$ (Lemma 6.1). Because $\Omega' \subseteq \Omega$, it is clear $X \in Y \uparrow \Omega$.

It is easy to see that $Y = [Y_1, \ldots, Y_n, d]$ where $\forall i : Y_i \in X_i \downarrow AS$.

We show that $\forall i : Y_i \in E$. Suppose $\exists i : Y_i \notin E$. From $Y_i \in AR_K$, E attacks Y_i wrt $att_{nr}(K)$. There are two cases:

- E undercuts Y_i. Since E does not undercut any argument in AS, E undercuts X_i. From $X_i \in E'$, E undercuts itself. Impossible. This case hence does not happen.
- E rebuts Y_i at Y'_i. Therefore $Y'_i \notin E$. From $AS \subseteq E$ and E contains all subarguments of its arguments, it follows that Y'_i is not a subargument of any argument in AS. From $Y_i \in X_i \downarrow AS$ and Lemma F.1, it follows that there exists a subargument X'_i of X_i s.t. $Y'_i \in X'_i \downarrow AS$. Therefore E contradicts X'_i at X'_i. From $X_i \in E'$ and E' contains all subarguments of its arguments, it follows that $X'_i \in E'$. Thus E' is contradictory. Contradiction to the consistency of $cnl(E')$ (Note that $cnl(E') = cnl(E)$).

We show $Y \notin E$. Suppose $Y \in E$. From $X \in Y \uparrow \Omega$, $X \in E'$. Contradiction.

Therefore $Y \in AR_K \setminus E$ and $Y = [Y_1, \ldots, Y_n, d]$ and $\forall i : Y_i \in E$. Therefore $\exists C \in E$ s.t. $(C, Y) \in att_{nr}(K)$. From context-independence (Lemma 6.4), $(C, Y) \in att_{nr}(K')$. From $AS \subseteq E$ and $C \in E$, it follows that C does not attack AS wrt $att_{nr}(K)$. From the context-independence property, C does not attack AS wrt $att_{nr}(K')$. Since Y is a weakening of X wrt AS and att_{nr} satisfies the axiom of link-orientation (Lemma 6.4), $(C, X) \in att_{nr}(K')$. From $E \subseteq E'$, we have $C \in E'$. Hence E' attacks X wrt $att_{nr}(K')$. □

Appendix G

Theorem 7.1. *Every enumeration-based extension of K is a stable extension of $(AR_K, att_{nr}(K))$.*

Proof. Let E be an enumeration-based extension of K. We show that E is a stable extension of $(AR_K, att_{nr}(K))$.

As E is a stable extension of $(AR_K, att_{bs}(K_{basic}))$, and $att_{nr}(K) \subseteq att_{bs}(K_{basic})$, E is conflict-free wrt $att_{nr}(K)$.

Let X be an argument not belonging to E. We show E attacks X wrt $att_{nr}(K)$. Let SU_X be the set of subarguments of X not belonging to E. Hence there exists one such argument such that no proper subarguments of it belong to SU_X. Let A be such an argument. Thus all proper subarguments of A belong to E. Therefore A is basic defeasible (otherwise $A \in E$ following Lemma F.2 as E is a stable extension of $(AR_K, att_{bs}(K_{basic}))$ and $att_{bs}(K_{basic}) = att_{nr}(K_{basic})$).

Since E is a stable extension of $(AR_K, att_{bs}(K_{basic}))$, there is an argument $B \in E$ that either undercuts or rebuts A. If B undercuts A then it is obvious B undercuts X and hence $(B, X) \in att_{nr}(K)$.

Suppose B rebuts A. It is clear that B rebuts A at A. Let $last(A) = d$. Therefore $bd(d) \subseteq cnl(E)$. From $cnl(B) = \neg hd(A)$, it follows $d \notin \Gamma_E$.

There are two cases:

1. $\nexists d' \in ldr(B)$ s.t. $d' \prec d$. Therefore $(B, X) \in att_{nr}(K)$.
2. There is $d' \in ldr(B) \subseteq \Gamma_E$ such that $d' \prec d$. Let $(d_i)_{i \geq 1}$ be an enumeration of Γ_E as described in Definition 7.1. From $d' \in \Gamma_E$, $\exists i$ s.t. $d' = d_i$. Let $n = min\{ j \mid d_j \prec d\}$. Hence $d_n \prec d$. From Definition 7.1, it follows $\{hd(d_k) | k < n\} \cup BE \vdash_K \neg hd(d)$. Therefore, there exists a strict argument Z over $\{hd(d_k) | k < n\} \cup BE$ such that $cnl(Z) = \neg hd(d)$. For each $d_k, k < n$, there

exists a basic defeasible argument $X_k \in E$ such that $last(X_k) = d_k$. Let Y be obtained by replacing each premise $hd(d_k)$ in Z by X_k. Therefore $ldr(Y) \subseteq \{d_k \mid k < n\}$ and $cnl(Y) = \neg hd(d)$. From $X_k \in E$, Z is strict, and E is a stable extension of $(AR_K, att_{bs}(K_{basic}))$ and $att_{bs}(K_{basic}) = att_{nr}(K_{basic})$ and Lemma F.2, we conclude $Y \in E$. From the definition of n, it follows that there is no rule in $ldr(Y)$ that is strictly less preferred than d. Therefore, Y rebuts X (at A) and there is no rule in $ldr(Y)$ that is strictly less preferred than $last(A)$. We have proved $(Y, X) \in att_{nr}(K)$. □

Lemma 7.1. *Suppose K is stratified and $A, B \in AR_K$ such that $(A, B) \in att_{nr}(K)$. Then $\rho(A) \leq \rho(B)$.*

Proof. Suppose A undercuts or rebuts B at B'. Therefore $\rho(B') \leq \rho(B)$. If A undercuts B' (at B'), it follows directly from Definition 7.2, $\rho(A) \leq \rho(last(B')) = \rho(B') \leq \rho(B)$.

Suppose A rebuts B at B'. Therefore there is no $d \in ldr(A)$ s.t. $d \prec last(B')$ implying that there is no $d \in ldr(A)$ s.t. $\rho(d) > \rho(last(B'))$. Hence for each $d \in ldr(A)$, it holds: $\rho(d) \leq \rho(last(B'))$ implying that $\rho(A) \leq \rho(B') \leq \rho(B)$. □

Theorem 7.2. *Suppose K is a consistent and stratified knowledge base satisfying the self-contradiction property. Then each stable extension of $(AR_K, att_{nr}(K))$ is an enumeration-based extension of K.*

Proof. Let E be a stable extension of $(AR_K, att_{nr}(K))$. From Lemmas 6.6, 4.1, it follows immediately that $cnl(E)$ is consistent. As E is conflict-free wrt $att_{nr}(K)$, arguments in E do not undercut each other. Thus from the consistency of $cnl(E)$, and Lemma 4.3, arguments in E do not rebut each other. Therefore, E is conflict-free wrt $att_{bs}(K_{basic})$. Hence E is a stable extension of $(AR_K, att_{bs}(K_{basic}))$. Let $\Gamma_i = \{d \in \Gamma_E \mid \rho(d) = i\}$. Define an enumeration of Γ_E as follows:

1. List arbitrarily all rules in Γ_0 resulting in $(d_i)_{i \leq n_0}$.
2. Suppose the list $(d_i)_{i \leq n_i}$ of rules in $\Gamma_0 \cup \ldots \cup \Gamma_i$ has been constructed. $(d_i)_{i \leq n_{i+1}}$ is obtained by from $(d_i)_{i \leq n_i}$ by appending to it an arbitrary list of rules in Γ_{i+1}.

We show that $(d_i)_{i \geq 1}$ is an enumeration of E as defined in Definition 7.1.

1. It is obvious that if $d_i \prec d_j$ then $j < i$.
2. We show $\{hd(d_k) | k < i\} \cup BE \vdash_K bd(d_i)$. Since $d_i \in \Gamma_E$, there is an argument $A \in E$ s.t. $last(A) = d_i$. Since all defeasible rules in $dr(A) \setminus \{d_i\}$ have a rank less than $\rho(d_i)$, they all belong to $(d_j)_{j<i}$. From $(hd(dr(A)) \setminus \{d_i\}) \cup BE \vdash_K bd(d_i)$, it follows $\{hd(d_k) | k < i\} \cup BE \vdash_K bd(d_i)$.
3. Let $d \in RD \setminus \Gamma_E$ such that $bd(d) \subseteq cnl(E)$ and there is j such that $d_j \prec d$ (i.e. $\rho(d) < \rho(d_j)$).
There exists hence a basic defeasible argument B with $last(B) = d$ and whose proper subarguments all belong to E. Since E is stable extension of $(AR_K, att_{nr}(K))$, there is an argument $A \in E$ s.t. $(A, B) \in att_{nr}(K)$. From $\rho(A) \leq \rho(d)$ (Lemma 7.1) and $\rho(d) < \rho(d_j)$, it follows immediately that all defeasible rules in A are listed before d_j in $(d_i)_{i \geq 1}$. Therefore $\{hd(d_k) | k < j\} \cup BE \vdash_K \neg hd(d)$ or $\{hd(d_k) | k < j\} \cup BE \vdash_K ab_d$. □

Appendix H

Theorem 7.8. *Let $\mathcal{K}$ be a sensible class of knowledge bases. Both attack relations assignments att_{lE} and att_{wE} satisfy the property of attack monotonicity for $\mathcal{K}$.*

Proof. It is not difficult to see that for $X' \subseteq X$, if $X' \lhd_E Y$ then $X \lhd_E Y$, and if $Y \lhd_E X$ then $Y \lhd_E X'$.

Let $S \subseteq BE$ and $A \in AR_K$ and $A' \in A \uparrow S$. Let $f_w = dr$ and $f_l = ldr$.

- Suppose $(A, B) \in att_{xE}$. If A undercuts B then A' also undercuts B. Hence $(A', B) \in att_{xE}$.
Let A rebuts B (at B'). Therefore $f_x(A) \not\lhd_E f_x(B')$, From $f_x(A') \subseteq f_x(A)$ it follows $f_x(A') \not\lhd_E f_x(B')$. Hence $(A', B) \in att_{xE}$.
- Suppose $(C, A') \in att_{xE}$. If C undercuts A' then C undercuts A. Let C rebuts A' (at B'). Therefore $C \not\lhd_{xE} B'$, i.e. $f_x(C) \not\lhd_E f_x(B')$. From $A' \in A \uparrow S$ and Lemma 4.2, there is a basic defeasible subargument B of A such that $B' \in B \uparrow S$. From $f_x(B') \subseteq f_x(B)$, it follows $f_x(C) \not\lhd_E f_x(B)$. Hence $(C, B) \in att_{xE}$. Since att_{xE} satisfies the property of subargument structure (Lemma 7.6), $(C, A) \in att_{xE}$. □

Theorem 7.9. *Let $\mathcal{K}$ be a sensible class of knowledge bases satisfying the self-contradiction property. Both attack relation assignments att_{lD}, att_{wD} satisfy the property of credulous cumulativity for $\mathcal{K}$.*

Proof. Let $K \in \mathcal{K}$ and E be a stable extension of $(AR_K, att_{xD}(K))$ for $x \in \{l, w\}$, $S = cnl(E)$ and $\Omega \subseteq S$ be a finite set of domain literals. Further let $E' = E \uparrow \Omega = \bigcup_{A \in E} A \uparrow \Omega$ and $K' = K + \Omega$. It is clear that $E \subseteq E' \subseteq AR_{K'}$, $cnl(E) = cnl(E')$ and $BE \cup \Omega \subseteq S$.

We show that E' is a stable extension of $(AR_{K'}, att_{xD}(K'))$.

From Lemma H.1 below, it follows S is consistent. S is hence not contradictory.

As E is conflict-free wrt $att_{xD}(K)$, E' is also free from the undercut-attacks. From Lemma 4.2 and Lemma 4.3, it follows that E' contains all subarguments of its arguments. Since $S = cnl(E')$ is not contradictory and E' contains all subarguments of its arguments, E' is free from rebuts. Therefore E' is conflict-free wrt $att_{xD}(K')$.

Property 1. *Let $A = [X_1, \ldots, X_n, r] \in AR_{K'}$ and r be a strict rule. It holds: A belongs to E' iff $\forall i: X_i \in E'$.*

Proof. Since E' contains all subarguments of its arguments, we only need to prove the "if"-direction.

Suppose $\forall i: X_i \in E'$. Therefore $\forall i \exists Y_i \in E: X_i \in Y_i \uparrow \Omega$. Let $B = [Y_1, \ldots, Y_n, r]$. Therefore $mbd(B) \subseteq E$. It follows that $B \in E$ (Lemma H.2). From $A \in B \uparrow \Omega$, it follows immediately $A \in E'$. $\square$

Property 2. *Let $A \in E'$ and AS be a subset of E such that $cnl(AS) = \Omega$. Further let B be obtained from A by replacing each subargument $[\alpha]$, $\alpha \in \Omega$, of A by an argument $A_\alpha \in AS$ where $cnl(A_\alpha) = \alpha$. It holds that $B \in E$.*

Proof. Suppose $B \notin E$. It is clear that $B \in AR_K$. Therefore E attacks B wrt $att_{xD}(K)$. There are two cases:

- E undercuts B. Since E does not undercut any argument in AS, E undercuts A. From $A \in E'$, E undercuts itself. Contradiction. This case hence does not happen.
- E rebuts B at B'. Therefore $B' \notin E$. From $AS \subseteq E$ and E contains all subarguments of its arguments, it follows that B' is not a subargument of any argument in AS. From $B \in A \downarrow AS$ and Lemma F.1, it follows that there exists a subargument A' of A s.t. $B' \in A' \downarrow AS$. Therefore E contradicts A' at A'. From $A \in E'$ and E' contains all subarguments of its arguments, it follows that $A' \in E'$. Thus E' is contradictory. Contradiction to the consistency of $cnl(E')$ (note that $cnl(E') = cnl(E)$). $\square$

Let $Z \in AR_{K+\Omega} \setminus E'$. We show that E' attacks Z wrt $att_{lD}(K')$.

It is clear that there exists a subargument X of Z such that all proper subarguments of X belong to E' and $X \notin E'$. From $S = cnl(E) = cnl(E')$ and $BE \cup \Omega \subseteq S$ and the above Property 1, it follows $X = [X_1, \ldots, X_n, d]$ and d is defeasible and $\forall i: X_i \in E'$.

We show that E' attacks X wrt $att_{xD}(K')$.

There are two cases:

Case 1: $X \in AR_K$. Then $X \notin E$. Hence E attacks X wrt $att_{xD}(K)$. Hence E' attacks X wrt $att_{xD}(K')$.

Case 2: $X \notin AR_K$. Therefore $X \in AR_{K+\Omega} \setminus AR_K$. Let Ω' be a minimal subset of Ω such that $X \in AR_{K+\Omega'}$. Further let AS be a minimal subset of E such that $cnl(AS) = \Omega'$.

Let Y be obtained from X by replacing each subargument $[\alpha]$, $\alpha \in \Omega'$, of X by an argument $A_\alpha \in AS$ where $cnl(A_\alpha) = \alpha$.

It is clear $Y \in AR_K$ and $Y = [Y_1, \ldots, Y_n, d]$ where each Y_i is obtained from X_i by replacing each subargument $[\alpha]$, $\alpha \in \Omega'$, of X_i by an argument $A_\alpha \in AS$ with $cnl(A_\alpha) = \alpha$.

Therefore from the above Property 2, it holds that $\forall i: Y_i \in E$.

We show $Y \notin E$. Suppose $Y \in E$. It is clear that $Y \in X \downarrow AS$. Therefore $X \in Y \uparrow \Omega'$ (Lemma 6.1). Because $\Omega' \subseteq \Omega$, it is clear $X \in Y \uparrow \Omega$. From $X \in Y \uparrow \Omega$, $X \in E'$. Contradiction.

Therefore it holds that $Y \in AR_K \setminus E$ and $\exists C \in E$ s.t. $(C, Y) \in att_{xD}(K)$. From context-independence (Lemma 7.6), $(C, Y) \in att_{xD}(K')$.

There are two cases:

- C undercuts Y. Since $AS \subseteq E$, it follows C undercuts X. Therefore E' attacks X wrt att_{xD}.
- C rebuts Y. Since $Y = [Y_1, \ldots, Y_n, d]$ and $Y_i \in E$, it follows that C rebuts Y at Y. Therefore C also rebuts X at X.
  - Let x = w.
    We show that $(C, X) \in att_{wD}(K')$. Suppose the contrary that $(C, X) \notin att_{wD}(K')$. Therefore $dr(C) \lhd_D dr(X)$. From $dr(X) \subseteq dr(Y)$, it follows $dr(C) \lhd_D dr(Y)$. Hence $(C, Y) \notin att_{wD}(K)$. Contradiction. We have proved that $(C, X) \in att_{wD}(K')$. Therefore E' attacks X wrt att_{wD}.
  - Let x = l.
    We show that $(C, X) \in att_{lD}(K')$. Suppose the contrary that $(C, X) \notin att_{lD}(K')$. Therefore $ldr(C) \lhd_D \{d\}$. Hence $(C, Y) \notin att_{lD}(K)$. Contradiction. We have proved that $(C, X) \in att_{lD}(K')$. Therefore E' attacks X wrt att_{lD}. $\square$

Lemma H.1. *Let $\mathcal{K}$ be a sensible class of knowledge bases that satisfy the property of self-contradiction.*

Then for each knowledge base $K \in \mathcal{K}$, each complete extension E of $(AR_K, att_{xD}(K))$, $x \in \{l, w\}$, $cnl(E)$ is consistent and closed.

Proof. Let E be a complete extension of $(AR_K, att_{xD}(K))$ and $S = cnl(E)$. From Lemma H.3, it is obvious that S is closed.

We next show the consistency of S. Suppose the contrary. Since S is closed, S is contradictory. Thus there are two arguments $A, B \in E$ such that $cnl(B) = \neg cnl(A)$. Let $AS = mbd(A) \cup mbd(B) \subseteq E$. Therefore $cnl(AS) \cup BE$ is inconsistent (wrt strict rules RS of K).

Let M be a minimal inconsistent subset of $cnl(AS) \cup BE$. Since BE is consistent, $M \cap cnl(AS) \neq \emptyset$. Let AS_0 be a minimal subset of AS such that $cnl(AS_0) = M \cap cnl(AS)$. Hence $AS_0 \neq \emptyset$.

1. Let x = l.
   Let $\Delta_0 = \{last(X) \mid X \in AS_0\}$ and $C \in AS_0$ such that $last(C)$ be minimal wrt $\prec$ in Δ_0. From self-contradiction property, it follows $M \vdash \neg hd(last(C))$. Thus there is an argument Y s.t. $mbd(Y) \subseteq AS_0$ and $cnl(Y) = \neg hd(last(C))$. From Lemma H.2, it follows $Y \in E$. It is clear that $ldr(Y) \subseteq \Delta_0$. Therefore, there is no default $d \in ldr(Y)$ s.t. $d \prec last(C)$.
   We show $(Y, C) \in att_{lD}(K)$. Suppose $(Y, C) \notin att_{lD}(K)$. Therefore $ldr(Y) \lhd_D \{last(C)\}$, i.e. $ldr(Y) \unlhd_D \{last(C)\}$ and $\{last(C)\} \not\unlhd_D ldr(Y)$. Therefore $\forall d \in ldr(Y) : d \preceq last(C)$ and $\neg(\exists d \in ldr(Y) : last(C) \preceq d)$ implying that $\forall d \in ldr(Y) : d \prec last(C)$. Since we have proved before there is no default $d \in ldr(Y)$ s.t. $d \prec last(C)$, it follows that $ldr(Y) = \emptyset$. Therefore $(Y, C) \in att_{lD}(K)$. Contradiction to the assumption that $(Y, C) \notin att_{lD}(K)$. We have proved that $(Y, C) \in att_{lD}(K)$.
   From $C \in AS_0 \subseteq AS \subseteq E$ and $Y \in E$, it follows E is not conflict-free wrt $att_{lD}(K)$. Contradiction. S is thus consistent.
2. Let x = w.
   Let $\Delta_0 = \{last(X) \mid X \in AS_0\}$ and $C \in AS_0$. From self-contradiction property, it follows $M \vdash \neg hd(last(C))$. Thus there is an argument Y s.t. $mbd(Y) \subseteq AS_0$ and $cnl(Y) = \neg hd(last(C))$. From Lemma H.2, it follows $Y \in E$.
   Since E is conflict-free, $(Y, C) \notin att_{wD}$. Hence Y is defeasible and $dr(Y) \lhd_D dr(C)$.
   From $dr(C) \not\unlhd_D dr(Y)$, it follows that there is $d \in dr(C)$ s.t. for each $d' \in dr(Y) : d \not\preceq d'$. Since Y is defeasible, $mbd(Y) \neq \emptyset$. Let $C' \in mbd(Y)$. Therefore $C' \neq C$ and $dr(C') \subseteq dr(Y)$. From the self-contradiction property, it follows $M \vdash \neg hd(last(C'))$. Thus there is an argument Y' s.t. $mbd(Y') \subseteq AS_0$ and $cnl(Y') = \neg hd(last(C'))$. From Lemma H.2, it follows $Y' \in E$. From the minimality of AS_0, it follows that $C' \cup mbd(Y') = AS_0$. From $C \neq C'$, it follows $C \in mbd(Y')$. From the conflict-freeness of E, it follows $dr(Y') \lhd_D dr(C')$. From $dr(C) \subseteq dr(Y')$, it follows $dr(C) \lhd_D dr(C')$. From $dr(C') \subseteq dr(Y)$, it follows $dr(C) \lhd_D dr(Y)$. Contradiction. S is thus consistent. □

From the definition of att_{xD}, $x \in \{l, w\}$, it follows immediately

Lemma H.2. *For each knowledge base K, for each complete extension E of $(AR_K, att_{xD}(K))$, $x \in \{l, w\}$, and for each argument $A \in AR_K$, it holds that $A \in E$ iff each maximal basic defeasible subargument of A belongs to E (i.e. $mbd(A) \subseteq E$).* □

Lemma H.3. *Let $\mathcal{K}$ be a sensible class of knowledge bases that satisfy the property of self-contradiction. Then for each knowledge base $K \in \mathcal{K}$, each complete extension E of $(AR_K, att_{xD}(K))$, $x \in \{l, w\}$, $cnl(E)$ is closed.*

Proof. Let E be a complete extension of $(AR_K, att_{xD}(K))$ and $S = cnl(E)$.

Let $S \vdash \alpha$. If $\alpha \in S$, we are done.

Suppose $\alpha \notin S$. Therefore there is a finite $\Omega \subseteq S$ of domain literals s.t. $\Omega \vdash \alpha$. Let A_0 be a strict argument over Ω wrt the set of strict rule RS of K. It is clear $A_0 \in AR_{K+\Omega}$.

Let A be a weakening of A_0 by replacing each subargument of the form $[\alpha]$, $\alpha \in \Omega$, in A_0 by an argument $X_\alpha \in E$ s.t. $cnl(X_\alpha) = \alpha$. Therefore $A \in AR_K$. It is clear each argument in $mbd(A)$ is a subargument of some argument in $\{X_\alpha \mid \alpha \in \Omega\}$. From Lemma 4.3, $mbd(A) \subseteq E$. From Lemma H.2, $A \in E$. Therefore $\alpha \in S$. Impossible since we assume that $\alpha \notin S$. This case hence cannot occur.

We have proved that S is closed. □

Lemma 7.7. $att_{lE} \sqsubseteq att_{nr} \sqsubseteq att_{lD} \sqsubseteq att_{bs}$.[45]

Proof. Let K be a knowledge base. It is obvious that $att_{lE}(K), att_{nr}(K), att_{lD}(K)$ are all subsets of $att_{bs}(K)$.

1. We show $att_{lE}(K) \subseteq att_{nr}(K)$ by showing $att_{bs}(K) \setminus att_{nr}(K) \subseteq att_{bs}(K) \setminus att_{lE}(K)$.
   Let $(A, B) \in att_{bs}(K) \setminus att_{nr}(K)$. Hence A rebuts B and for each basic defeasible subargument X of B s.t. $cnl(A) = \neg cnl(X)$, there is $d_X \in ldr(A) : d_X \prec last(X)$. Therefore $ldr(A) \unlhd_E \{last(X)\}$ and $\{last(X)\} \not\unlhd_E ldr(A)$. Therefore $ldr(A) \lhd_E \{last(X)\}$. Hence A does not attack B by rebut at X wrt $att_{lE}(K)$. Hence $(A, B) \notin att_{lE}(K)$.
2. We show $att_{nr}(K) \subseteq att_{lD}(K)$ by showing $att_{bs}(K) \setminus att_{nr}(K) \supseteq att_{bs}(K) \setminus att_{lD}(K)$.
   Let $(A, B) \in att_{bs}(K) \setminus att_{lD}(K)$. Hence A rebuts B and for each basic defeasible subargument X of B s.t. $cnl(A) = \neg cnl(X)$, $A \sqsubset_{lD} X$.
   Therefore $\emptyset \neq ldr(A) \unlhd_D \{last(X)\}$ and $\{last(X)\} \not\unlhd_D ldr(A)$, i.e. $\forall d \in ldr(A) : d \preceq last(X)$ and $\forall d \in ldr(A) : last(X) \not\preceq d$. Hence $ldr(A) \neq \emptyset$ and $\forall d \in ldr(A) : d \prec last(X)$.
   It follows immediately that for all basic defeasible subargument X of B s.t. $cnl(A) = \neg cnl(X)$, $\exists d \in ldr(A) : d \prec last(X)$. As A does not undercut B, it follows obviously that $(A, B) \notin att_{nr}(K)$. □

Lemma 7.9. *Let $\mathcal{K}$ be a sensible class of knowledge bases satisfying the self-contradiction property. Further let $K \in \mathcal{K}$ and E be a stable extension wrt the attack relation assignment att_{lsE}. Then cnl(E) is consistent and closed.*

[45] Note that for attack relation assignments att, att', $att \sqsubseteq att'$ iff for each knowledge base K, $att(K) \subseteq att'(K)$.

Proof. Let $S = cnl(E)$. From Lemma H.4, it is obvious that S is closed.

We next show the consistency of S. Suppose the contrary. Since S is closed, S is contradictory. Thus there are two arguments $A, B \in E$ such that $cnl(B) = \neg cnl(A)$. Let $AS = mbd(A) \cup mbd(B) \subseteq E$. Therefore $cnl(AS) \cup BE$ is inconsistent (wrt strict rules RS of K).

Let M be a minimal inconsistent subset of $cnl(AS) \cup BE$. Since BE is consistent, $M \cap cnl(AS) \neq \emptyset$. Let $AS_0 \subseteq AS$ such that $cnl(AS_0) = M \cap cnl(AS)$. Hence $AS_0 \neq \emptyset$.

Let $\Delta_0 = \{last(X) \mid X \in AS_0\}$ and $C \in AS_0$ such that $last(C)$ be minimal wrt $\prec$ in Δ_0. From self-contradiction property, it follows $M \vdash \neg hd(last(C))$. Thus there is an argument Y s.t. $mbd(Y) \subseteq AS_0$ and $cnl(Y) = \neg hd(last(C))$. From Lemma H.5, it follows $Y \in E$. It is clear that $ldr(Y) \subseteq \Delta_0$.

We show that $Y \not\sqsubset_{lsE} C$. Suppose the contrary that $Y \sqsubset_{lsE} C$. Therefore there is $y \in ldr(Y)$ s.t. $y \prec last(C)$. This is impossible since $last(C)$ is minimal wrt $\prec$ in Δ_0 and $ldr(Y) \subseteq \Delta_0$.

Therefore $(Y, C) \in att_{lsE}(K)$. From $C \in AS_0 \subseteq AS \subseteq E$, it follows E is not conflict-free wrt $att_{lsE}(K)$. This is impossible since E is a stable extension wrt att_{lsE}. □

Lemma H.4. *Let $\mathcal{K}$ be a sensible class of knowledge bases that satisfy the property of self-contradiction. Then for each knowledge base $K \in \mathcal{K}$, for each complete extension E of $(AR_K, att_{lsE}(K))$, $cnl(E)$ is closed.*

Proof. Let E be a complete extension of $(AR_K, att_{lsE}(K))$ and $S = cnl(E)$.

Let $S \vdash \alpha$. If $\alpha \in S$, we are done.

Suppose $\alpha \notin S$. Therefore there is a finite $\Omega \subseteq S$ of domain literals s.t. $\Omega \vdash \alpha$. Let A_0 be a strict argument over Ω wrt the set of strict rule RS of K. It is clear $A_0 \in AR_{K+\Omega}$.

Let A be a weakening of A_0 by replacing each subargument of the form $[\alpha]$, $\alpha \in \Omega$, in A_0 by an argument $X_\alpha \in E$ s.t. $cnl(X_\alpha) = \alpha$. Therefore $A \in AR_K$. It is clear each argument in $mbd(A)$ is a subargument of some argument in $\{X_\alpha \mid \alpha \in \Omega\}$. From Lemma 4.3, $mbd(A) \subseteq E$. From Lemma H.5, $A \in E$. Therefore $\alpha \in S$. Impossible since we assume that $\alpha \notin S$. This case hence cannot occur.

We have proved that S is closed. □

From the definition of attack relation assignments att_{lsE}, and Lemmas 4.3, 7.4 it is easy to see that the following lemma holds.

Lemma H.5. *For each knowledge base K, for each complete extension E of $(AR_K, att_{lsE}(K))$ and each argument $A \in AR_K$, it holds that $A \in E$ iff each maximal basic defeasible subargument of A belongs to E (i.e. $mbd(A) \subseteq E$).* □

Appendix I

Theorem 7.3. *Let att be an attack relation assignment satisfying the properties of subargument structure and context-independence for a sensible class of knowledge bases $\mathcal{K}$. Then att satisfies the second BE-principle for $\mathcal{K}$.*

Proof. Let $K \in \mathcal{K}$, E be a stable extension of $(AR_K, att(K))$, $S = cnl(E)$ and r be a rule not applicable wrt S. Let $K' = K + r$.

We show that E is also a stable extension of K'.

From the property of context-independence, E is conflict-free wrt $att(K')$. We only need to show that E attacks (wrt $att(K')$) each argument in $AR_{K'}$ not belonging to E.

Let A be an argument in $AR_{K'}$ not belonging to E. If A is also an argument in AR_K then E attacks A wrt $att(K)$ and hence also wrt $att(K')$ (due to the property of context-independence).

Suppose that A is not an argument in AR_K. Hence r appears in A. As r is not applicable in S, there is a subargument B of A without containing r and whose conclusion does not belong to S. Hence E attacks B wrt $att(K)$ (and therefore also wrt $att(K')$ due to the property of context-independence). B is hence defeasible. Therefore from the property of subargument structure, E attacks A (wrt $att(K')$). E is hence a stable extension of K'. □

Lemma I.1. *Let $\mathcal{K}$ be a sensible class of knowledge bases and att be an attack relation assignment satisfying the properties of subargument structure and attack closure for $\mathcal{K}$. Further let $K \in \mathcal{K}$ and E be a stable extension of $(AR_K, att(K))$ such that $cnl(E)$ is consistent. Then $cnl(E)$ is generated by the set of rules appearing in arguments in E.*

Proof. Let Γ be the set of rules appearing in E. We show that $S = cnl(E)$ is generated by Γ.

1. It is obvious that for each literal σ, $\sigma \in S$ iff there is an argument A such that $cnl(A) = \sigma$ and all rules appearing in A belong to Γ.
2. To show that for each rule $\gamma \in RS \cup RD$, $\gamma \in \Gamma$ iff (γ is strict and $bd(\gamma) \subseteq S$ and $hd(\gamma) \in S$) or (γ is defeasible and $bd(\gamma) \subseteq S$ and $hd(\gamma) \in S$ and $ab_\gamma \notin S$), we only need to show the "if"-direction.

Suppose $\gamma \notin \Gamma$. Let $bd(\gamma) = \{\alpha_1, \ldots, \alpha_n\}$. From $bd(\gamma) \subseteq S$ and $S = cnl(E)$, there are arguments $A_i \in E$ s.t. $cnl(A_i) = \alpha_i$. Since $\gamma \notin \Gamma$, argument $A = [A_1, \ldots, A_n, \gamma]$ does not belong to E. Hence there is $B \in E$ such that $(B, A) \in att(K)$. From the property of attack closure, B either undercuts or contradicts A. Since $\forall i : A_i \in E$ and (γ is strict or $ab_\gamma \notin S$), B does not undercut A. Therefore B contradicts A. From Lemma 4.3, all subarguments of A_i belong to E. From the consistency of $cnl(E)$, B does not contradicts any subargument of any A_i. Therefore B contradicts A at A. Hence $\neg hd(\gamma) \in cnl(E)$. $cnl(E)$ is thus inconsistent. Contradiction. □

Theorem 7.4. *Let att be an ordinary attack relation assignment defined for a sensible class $\mathcal{K}$ of knowledge bases. Then att satisfies the first BE-principle for $\mathcal{K}$.*

Proof. Let $K \in \mathcal{K}$. Suppose S, S' are sets of literals as specified in Definition 7.4. We show that S is not a stable belief set of K wrt att.

Suppose S is a stable belief set of K wrt att. Let E be a stable extension of $(AR_K, att(K))$ such that $cnl(E) = S$. From Lemma 4.1, S is consistent and from Lemma I.1, $\Gamma \cup \{d\}$ is the set of rules appearing in arguments in E.

Let d' be of the form $\alpha_1, \ldots, \alpha_n \Rightarrow \gamma$. As $\Gamma \cup \{d'\}$ is the generating set of S', there are arguments $A_1, \ldots, A_n$ with conclusions $\alpha_1, \ldots, \alpha_n$ respectively and all rules of A_i belong to Γ.

We show $\forall i : A_i \in E$. Suppose $A_i \notin E$ for some i. Hence E either undercuts or rebuts A_i. Hence there is a rule δ appearing in A_i s.t. $ab_\delta \in S$ or $\neg hd(\delta) \in S$. Contradiction since $\delta \in \Gamma$ and $\Gamma \cup \{d\}$ is the generating set of S.

Therefore $\{\alpha_1, \ldots, \alpha_n\} \subseteq S$. Since S is generated by $\Gamma \cup \{d\}$, it follows $A = [A_1, \ldots, A_n, d'] \notin E$. Since E is stable, E attacks A (wrt att(K)). Since $\forall i : A_i \in E$, it follows that $ab_{d'} \in S$ or $\neg\gamma \in S$.

Suppose $ab_{d'} \in S$. Hence there is a strict rule $r \in \Gamma$ with head $ab_{d'}$. Therefore $ab_{d'} \in S'$. Hence d' does not belong to the set of rules generating S', i.e. $d' \notin \Gamma \cup \{d'\}$. Contradiction. We have hence proved $\neg\gamma \in S$.

From the reduced characteristic Lemma E.1, there is $B \in E$ such that B attacks A (wrt $att(K)$) by directly rebutting A and there is no defeasible rule $\delta \in ldr(B)$ s.t. $\delta \prec d'$.

From B directly rebutting A, it follows $cnl(B) = \neg cnl(A) = \neg\gamma$. As S' is consistent, it follows $\neg\gamma \notin S'$. Hence the last rule of B belongs to $\Gamma \cup \{d\}$ but not to $\Gamma \cup \{d'\}$ implying that the last rule of B is d. B is hence basic defeasible. From $\nexists \delta \in ldr(B)$ s.t. $\delta \prec d'$ and $ldr(B) = \{d\}$, it follows $d \not\prec d'$. Contradiction to the assumption that $d \prec d'$. Hence the assumption that S is a stable belief set is wrong. □

Appendix J. Proof of Theorem 8.1

Theorem 8.1. *Let $\mathcal{K}$ be a sensible class of knowledge bases that are closed under transposition or contraposition. Then att_{nr} satisfies the credulous cumulativity axiom for $\mathcal{K}$ wrt complete extensions semantics.*

Proof. Let $K \in \mathcal{K}$. K is hence consistent and closed under transposition or contraposition. Let E be a complete extension of $(AR_K, att_{nr}(K))$, $S = cnl(E)$ and $\Omega \subseteq S$ a finite set of domain literals. Further let $AS \subseteq E$ s.t. $\Omega = cnl(AS)$.

We show that $E' = E \uparrow \Omega = \bigcup\{X \uparrow \Omega \mid X \in E\}$ is a complete extension of $(AR_{K'}, att(K'))$ for $K' = K + \Omega$.

Since E is free from undercut-attacks, it is easy to see that E' is also free from undercut-attacks. From Lemmas 4.2, 4.3, E' contains all subarguments of its arguments. Since S is not contradictory (Lemma F.4), E' is free from rebuts. E' is hence conflict-free wrt $att_{nr}(K')$.

Property 1. *E' is admissible.*

Proof. Suppose $A' \in AR_{K'}$ attacks $B' \in E'$ wrt $att_{nr}(K')$.

We need to show that E' attacks A' wrt $att_{nr}(K')$. From $E' = E \uparrow \Omega$, there is $B \in E$ s.t. $B' \in B \uparrow \Omega$.

From the attack monotonicity of att_{nr}, it follows $(A', B) \in att_{nr}(K')$.

Let A be a weakening of A' by replacing each subargument $[\alpha]$, $\alpha \in \Omega \setminus BE$ by an $X_\alpha \in AS$ s.t. $cnl(X_\alpha) = \alpha$. It is clear $A \in AR_K$.

We show that there is $C \in E$ attacking A wrt $att_{nr}(K)$. There are two cases:

- If $(A, B) \in att_{nr}(K)$ then there is $C \in E$ attacking A wrt $att_{nr}(K)$.
- Suppose $(A, B) \notin att_{nr}(K)$. Therefore $(A, B) \notin att_{nr}(K')$. From Lemma J.1, E attacks A or B wrt $att_{nr}(K)$. Since $B \in E$, E attacks A wrt $att_{nr}(K)$. Hence there is $C \in E$ attacking A wrt $att_{nr}(K)$.

Due to the context-independence property, C also attacks A wrt $att_{nr}(K')$.

As A is a weakening of A' by AS and C does not attack AS (wrt $att_{nr}(K)$ and hce also wrt $att_{nr}(K')$), C attacks A' wrt $att_{nr}(K')$ (due to the property of link-orientation). From $C \in E \subseteq E'$, it follows E' attacks A' wrt $att_{nr}(K')$. □

Property 2. *E' is complete.*

Proof. Suppose E' is not complete. Hence there is $A' \in AR_{K'}$ s.t. A' is defensible wrt E' (wrt $att_{nr}(K')$) but $A' \notin E'$. Let A be obtained from A' by replacing each subargument $[\alpha]$, $\alpha \in \Omega \setminus BE$ by an $X_\alpha \in AS$ s.t. $cnl(X_\alpha) = \alpha$. It is clear $A \in AR_K$ and $A \in A' \downarrow AS$.

Therefore $A \notin E$ (otherwise $A' \in E'$ from Lemma 6.1), A is hence not defensible wrt E, i.e. there is $B \in AR_K$ attacking A, but not counter-attacked by E (wrt $att_{nr}(K)$).

Therefore B does not attack those subarguments in $AS \subseteq E$. Since att_{nr} satisfies the link-oriented property, B attacks A' wrt $att_{nr}(K')$. Since A' is defensible by E', there is $X' \in E'$ attacking B wrt $att_{nr}(K')$. Since E does not attack B wrt $att_{nr}(K)$, X' attacks B wrt $att_{nr}(K')$ by rebut at B' (if X' undercuts B, E would also undercut B, a contradiction).

From $X' \in E'$, it follows there is $X \in E$ s.t. $X' \in X \uparrow \Omega$. Hence from Lemma J.3, there is $AS_0 \subseteq E$ s.t. $X \in X' \downarrow AS_0$.

From $X \in E$, X does not attack B wrt $att_{nr}(K)$. We hence have $(X', B) \in att_{nr}(K')$, $(X, B) \notin att_{nr}(K')$ and X is a weakening of X' by $AS_0 \subseteq E$. From Lemma J.1, it follows that E attacks X or B wrt $att_{nr}(K')$. Since $X \in E$, E attacks B wrt $att_{nr}(K')$. Due to the context-independence of att_{nr}, E attacks B wrt $att_{nr}(K)$. Contradiction. □

Lemma J.1. *Let K be a knowledge base closed under transposition or contraposition. Let $E \subseteq AR_K$ be admissible in $(AR_K, att_{nr}(K))$, $\Omega \subseteq cnl(E)$ be a finite set of domain literals and $AS \subseteq E$ s.t. $\Omega = cnl(AS)$.*

Further let $K' = K + \Omega$, $A, B \in AR_K$ and $A' \in AR_{K'}$ such that $A \in A' \downarrow AS$ and $(A', B) \in att_{nr}(K')$ and $(A, B) \notin att_{nr}(K')$. Then E attacks A or B wrt $att_{nr}(K)$.

Proof. From $(A', B) \in att_{nr}(K')$ and $(A, B) \notin att(K')$, it follows that A' does not undercut B. Therefore both A, A' rebut B (at B_1).

Therefore there is no $d \in ldr(A')$ s.t. $d \prec last(B_1)$ and there is $d \in ldr(A) \setminus ldr(A')$ s.t. $d \prec last(B_1)$.

Let $d' \in min_\prec\{d \in ldr(A) \setminus ldr(A') \mid d \prec last(B_1)\}$. From that fact that there is no $d \in ldr(A')$ s.t. $d \prec last(B_1)$, it follows clearly $d' \in min_\prec(ldr(A) \cup \{last(B_1)\})$.

Let $H \in mbd(A)$ such that $H \in E$ and $last(H) = d'$. From Lemma J.2, it follows that there is an argument Y in AR_K contradicting H at H and $mbd(Y) \subseteq mbd(A) \cup \{B_1\}$. From the definition of d', it is clear that Y rebuts H (at H) s.t. $\nexists d \in ldr(Y) : d \prec d'$. Hence Y attacks H wrt $att_{nr}(K)$.

Due to the admissibility of E wrt $att_{nr}(K)$, E attacks Y at some basic defeasible subargument Y'. From $mbd(Y) \subseteq mbd(A) \cup \{B_1\}$, it follows Y' is a subargument of some argument in $mbd(A) \cup \{B_1\}$. Thus E attacks A or B wrt $att_{nr}(K)$. □

Lemma J.2. *Let K be a knowledge base closed under transposition or contraposition. Suppose an argument A directly rebuts a basic defeasible argument B and $d \in ldr(A)$. Then there is an argument C with conclusion $\neg hd(d)$ such that $mbd(C) \subseteq mbd(A) \cup \{B\}$.*

Proof. Let $\Omega = hd(mbd(A))$ and T be the strict argument from $AR_{K+\Omega}$ obtained by replacing in A all subarguments from mbd(A) by their conclusions. Thus $cnl(T) = cnl(A)$ and $hd(mbd(A)) \subseteq Prem(T)$.[46] From Lemma A.1, there exists a strict argument T' (wrt $K + \Omega + \{\neg cnl(T)\}$) with conclusion $\neg hd(d)$ and $Prem(T') \subseteq Prem(T) \cup \{\neg cnl(T)\} = Prem(T) \cup \{cnl(B)\}$. Let $C \in AR_K$ be obtained from T' by replacing each subargument $[\alpha]$, $\alpha \in Prem(T') \cap hd(mbd(A) \cup \{B\})$ by an argument $X_\alpha \in mbd(A) \cup \{B\}$. Since T' is strict, it is obvious that $mbd(C) \subseteq mbd(A) \cup \{B\}$. □

Lemma J.3. *Let att be an attack relation assignment satisfying the axiom of subargument structure and K be a knowledge base. Further let E be a complete extension of $(AR_K, att(K))$, $\Omega \subseteq cnl(E)$ be a finite set of domain literals and $X \in E$. Then for each $X' \in X \uparrow \Omega$, there is $AS \subseteq E$ s.t. $X \in X' \downarrow AS$.*

Proof. We prove the lemma by induction on the structure of X.

1. $X = [\alpha]$ for $\alpha \in BE$. Hence $X' = X$ and $AS = \{X\}$.
2. $X = [X_1, \ldots, X_n, r]$. Therefore from Lemma 4.3, $X_i \in E$ for all i.
   If $X' = [cnl(r)]$ then $AS = \{X\}$.
   Let $X' = [X'_1, \ldots, X'_n, r]$. Therefore $X'_i \in X_i \uparrow \Omega$. From induction hypothesis, there is $AS_i \subseteq E$ s.t. $X_i \in X'_i \downarrow AS_i$. Let $AS = AS_1 \cup \ldots \cup AS_n$. Therefore $X \in X' \downarrow AS$. □

Appendix K. Closure operator

It is clear that for any $X \subseteq \mathcal{L}$, $CN_{RS}(X) \subseteq CN_{RS}(CN_{RS}(X))$. Suppose $CN_{RS}(CN_{RS}(X)) \setminus CN_{RS}(X) \neq \emptyset$. Let $\alpha \in CN_{RS}(CN_{RS}(X)) \setminus CN_{RS}(X)$. Therefore there is a strict argument A wrt RS over the domain literals in $CN_{RS}(X)$ s.t. $cnl(A) = \alpha$. Therefore for each premise λ of A, there is a strict argument A_λ wrt RS over X s.t. $cnl(A_\lambda) = \lambda$. Expand A by expanding each premise λ of A by A_λ. The obtained argument is a strict argument over X. Therefore $\alpha \in CN_{RS}(X)$. Contradiction.

---

[46] $Prem(A)$ is the set of premises of A consisting of conclusions of subarguments of A of the form $[\alpha]$.

Appendix L. Extended logic programming

Theorem 7.6. *Let S be a DST-preferred answer set of a prioritized logic program Π and $\Omega \subseteq S$. Then S is also a DST-preferred answer set of $\Pi + \Omega$.*

Proof. Let $(r_i)_{i \geq 1}$ be an enumeration of rules in $\Gamma_{P,S}$ as defined in Definition 7.5. Further let $P' = P + \Omega$. It is obvious that $\Gamma_{P',S} = \Gamma_{P,S} + \Omega$.

For ease of reference, from now until the end of this proof, we identify Ω and the set of rules $\{\omega \leftarrow | \omega \in \Omega\}$.

Further let n_0 be the number of elements in $\Omega \setminus P$. Define an enumeration $(r'_i)_{i \geq 1}$ of rules in $\Gamma_{P',S}$ as follows:

- $(r'_i)_{1 \leq i \leq n_0}$ is an arbitrary enumeration of the rules in $\Omega \setminus P$.
- For each $k \geq 1$: $r'_{n_0+k} = r_k$.

To show that S is a preferred answer set of $\Pi + \Omega$, we show for all i, j:

1. $bd^+(r'_i) \subseteq \{hd(r'_k) | k < i\}$;
2. if $r'_i \prec r'_j$ then $j < i$;
3. if $r'_i \prec r$ and $r \in P' \setminus \Gamma_{P',S}$ then
   (a) $bd^+(r) \nsubseteq S$, or
   (b) $bd^-(r) \cap \{hd(r'_k) | k < i\} \neq \emptyset$.

1. If $i \leq n_0$, it is clear that $bd^+(r'_i) = \emptyset \subseteq \{hd(r'_k) | k < i\}$.
   Let $i = n_0 + j$ and $k > 1$. It holds immediately:
   $bd^+(r'_i) = bd^+(r'_{n_0+j}) = bd^+(r_j) \subseteq \{hd(r_k) | k < j\} = \{hd(r'_k) | n_0 < k < n_0 + j\} \subseteq \{hd(r'_k) | k < n_0 + j\} = \{hd(r'_k) | k < i\}$.
2. Let $r'_i \prec r'_j$. From $r'_i \prec r'_j$, it immediately follows that $r'_i, r'_j \in P$. Therefore $i = n_0 + k$ and $j = n_0 + t$ and $r_k \prec r_t$. Therefore $t < k$. Hence $j < i$.
3. Let $r'_i \prec r$ and $r \in P' \setminus \Gamma_{P',S}$. It follows $r'_i, r \in P$. Therefore $i = n_0 + j$ and $j > 1$. Suppose $bd^+(r) \subseteq S$. Therefore $bd^-(r) \cap \{hd(r_k) | k < j\} \neq \emptyset$. From $\{hd(r_k) | k < j\} = \{hd(r'_k) \,|\, n_0 < k < n_0 + j\} \subseteq \{hd(r'_k) \,|\, k < n_0 + j\}$, it follows $bd^-(r) \cap \{hd(r'_k) | k < i\} \neq \emptyset$. □

Lemma 7.2. *Let $\Pi = (P, \ll)$ be a fully prioritized prerequisite-free logic program and X be an answer set of P. Then X is an BE-preferred answer set of Π iff for each $r \in P \setminus \Gamma_{P,X}$, if $hd(r) \notin X$ then*

$$bd^-(r) \cap \{\, hd(r') \mid r' \in \Gamma_{P,X} \text{ and } r \ll r' \,\} \neq \emptyset$$

Proof. Let $(r_i)_{i \geq 1}$ be the enumeration of P according to $\ll$ (i.e. $r_i \ll r_j$ iff $j < i$) and $X_0, X_1, \ldots, X_n$ be the sequence defined as in Definition 7.6, i.e. $X_0 = \emptyset$ and for $1 \leq i \leq n$,

$$X_i = \begin{cases} X_{i-1} & \text{if } bd^-(r_i) \cap X_{i-1} \neq \emptyset \\ X_{i-1} & \text{if } hd(r_i) \in X \text{ and } bd^-(r_i) \cap X \neq \emptyset \\ X_{i-1} \cup \{hd(r_i)\} & \text{otherwise} \end{cases}$$

It is not difficult to see that for $1 \leq i \leq n$, $X_i = X_{i-1} \cup \{hd(r_i)\}$ iff $bd^-(r_i) \cap X_{i-1} = \emptyset$ and $(hd(r_i) \notin X$ or $bd^-(r_i) \cap X = \emptyset)$.

1. "Only-If-Part".
   Suppose X is a BE-preferred answer set of Π and $r \in P \setminus \Gamma_{P,X}$ such that $hd(r) \notin X$.
   We first show by induction that for each $0 \leq j \leq n$,

   $$X_j = \{\, hd(r') \,|\, r' \in \Gamma_{P,X} \text{ and } (r_j \ll r' \text{ or } r' = r_j) \,\}$$

   - Base Step: j = 0. Obvious since both sides of the equation are empty.
   - Inductive Step. There are two cases:
     - $X_j = X_{j-1}$.
       Therefore $bd^-(r_j) \cap X_{j-1} \neq \emptyset$ or $(hd(r_j) \in X$ and $bd^-(r_j) \cap X \neq \emptyset)$. Since X is BE-preferred, it follows that $X_j \subseteq X$. Therefore it holds that $bd^-(r_j) \cap X \neq \emptyset$. Since X is an answer set of P, it follows from $bd^-(r_j) \cap X \neq \emptyset$ that $r_j \notin \Gamma_{P,X}$. From induction hypothesis, we can conclude: $X_j = X_{j-1} = \{\, hd(r') \,|\, r' \in \Gamma_{P,X}$ and $(r_{j-1} \ll r'$ or $r' = r_{j-1}) \,\} = \{\, hd(r') \,|\, r' \in \Gamma_{P,X}$ and $(r_j \ll r'$ or $r' = r_j) \,\}$.
     - $X_j \neq X_{j-1}$.
       Therefore $hd(r_j) \in X_j$. Since X is a BE-preferred answer set, $X_j \subseteq X$. It follows $hd(r_j) \in X$. Therefore $bd^-(r_j) \cap X = \emptyset$ (otherwise, it would hold that $X_i = X_{i-1}$. Impossible since $X_j \neq X_{j-1}$ in this case). Hence $r_j \in \Gamma_{P,X}$. From induction

hypothesis, it is easy to see $X_j = \{hd(r_j)\} \cup X_{j-1} = \{hd(r_j)\} \cup \{hd(r') \mid r' \in \Gamma_{P,X}$ and $(r_{j-1} \ll r'$ or $r' = r_{j-1})\} = \{hd(r') \mid r' \in \Gamma_{P,X}$ and $(r_j \ll r'$ or $r' = r_j)\}$.

We show that $bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{P,X}$ and $r \ll r'\} \neq \emptyset$. As $r_1, r_2, \ldots, r_n$ is an enumeration of P according to $\ll$, there is a unique i s.t. $r = r_i$. Hence $hd(r_i) \notin X_i$. Therefore $X_i = X_{i-1}$ implying that $bd^-(r_i) \cap X_{i-1} \neq \emptyset$ or $hd(r_i) \in X$ and $bd^-(r_i) \cap X \neq \emptyset$.

Since $hd(r_i) \notin X$, it follows $bd^-(r_i) \cap X_{i-1} \neq \emptyset$. Since X is BE-preferred, $X_i \subseteq X$. Therefore $r_i \notin \Gamma_{P,X}$.

From $X_i = \{hd(r') \mid r' \in \Gamma_{P,X}$ and $(r_i \ll r'$ or $r' = r_i)\}$ and $r_i \notin \Gamma_{P,X}$, it follows $X_i = \{hd(r') \mid r' \in \Gamma_{P,X}$ and $r_i \ll r'\} = \{hd(r') \mid r' \in \Gamma_{P,X}$ and $r \ll r'\}$.

From $X_i = X_{i-1}$ and $bd^-(r_i) \cap X_{i-1} \neq \emptyset$, it follows $bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{P,X}$ and $r \ll r'\} \neq \emptyset$.

1. "If-Part".
   Suppose for each $r \in P \setminus \Gamma_{P,X}$, if $hd(r) \notin X$ then $bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{P,X}$ and $r \ll r'\} \neq \emptyset$. Further let $P_i = \{r_1, \ldots, r_i\}$, $0 \leq i \leq n$.
   We first show by induction that for each i,

$$X_i = \{hd(r) \mid r \in P_i \cap \Gamma_{P,X}\} \subseteq X$$

- Base step: i = 0. Obvious.
- Inductive step. Suppose $X_{i-1} = \{hd(r) \mid r \in P_{i-1} \cap \Gamma_{P,X}\} \subseteq X$. We show $X_i = \{hd(r) \mid r \in P_i \cap \Gamma_{P,X}\} \subseteq X$.
  There are two cases:
  - $X_i = X_{i-1}$.
    Therefore $bd^-(r_i) \cap X_{i-1} \neq \emptyset$ or $hd(r_i) \in X$ and $bd^-(r_i) \cap X \neq \emptyset$. From $X_{i-1} \subseteq X$, it follows that $bd^-(r_i) \cap X \neq \emptyset$. Hence $r_i \notin \Gamma_{P,X}$. Therefore $P_{i-1} \cap \Gamma_{P,X} = P_i \cap \Gamma_{P,X}$. From induction hypothesis, we can state: $X_i = X_{i-1} = \{hd(r) \mid r \in P_{i-1} \cap \Gamma_{P,X}\} = \{hd(r) \mid r \in P_i \cap \Gamma_{P,X}\} \subseteq X$.
  - $X_i \neq X_{i-1}$.
    Therefore $bd^-(r_i) \cap X_{i-1} = \emptyset$ and $(hd(r_i) \notin X$ or $bd^-(r_i) \cap X = \emptyset)$.
    There are two cases:
    * $hd(r_i) \notin X$.
      Since X is an answer set of P, it follows that the head of each rule in $\Gamma_{P,X}$ belongs to X. Therefore $r_i \notin \Gamma_{P,X}$. Therefore, $bd^-(r_i) \cap \{hd(r') \mid r' \in \Gamma_{P,X}$ and $r_i \ll r'\} \neq \emptyset$. Since $r_1, \ldots, r_n$ is an enumeration of P according to $\ll$, it holds: $\{r' \in \Gamma_{P,X} \mid r_i \ll r'\} = P_{i-1} \cap \Gamma_{P,X}$. Therefore from the inductive hypothesis: $\{hd(r') \mid r' \in \Gamma_{P,X}$ and $r_i \ll r'\} = \{hd(r') \mid r' \in P_{i-1} \cap \Gamma_{P,X}\} = X_{i-1}$. Therefore $bd^-(r_i) \cap X_{i-1} \neq \emptyset$ implying that $X_i = X_{i-1}$. Contradiction. This case hence cannot occur.
    * $hd(r_i) \in X$.
      Therefore $bd^-(r_i) \cap X = \emptyset$. Thus $r_i \in \Gamma_{P,X}$. It holds obviously $r_i \in P_i \cap \Gamma_{P,X}$. From induction hypothesis:
      $X_i = \{hd(r_i)\} \cup X_{i-1} = \{hd(r_i)\} \cup \{hd(r) \mid r \in P_{i-1} \cap \Gamma_{P,X}\} = \{hd(r) \mid r \in P_i \cap \Gamma_{P,X}\} \subseteq X$.

From $X_n = \{hd(r) \mid r \in P_n \cap \Gamma_{P,X}\} = \{hd(r) \mid r \in \Gamma_{P,X}\} = X$, it follows that X is a BE-preferred answer set of $(P, \ll)$. □

Lemma 7.3. *Let $\Pi = (P, \ll)$ be a fully prioritized logic program and S be a BE-preferred answer set of Π and $\Omega \subseteq S$. Then there is strict total order $\ll'$ on $P + \Omega$ such that $\ll \subseteq \ll'$ and S is also a BE-preferred answer set of $(P + \Omega, \ll')$.*

Proof. From Theorem 7.5, S is also an answer set of $Q = P + \Omega$. For each $\omega \in \Omega$, let C_ω denote the clause $\omega \leftarrow$. Let $\ll'$ be a strict total order on $P + \Omega$ such that $\ll \subseteq \ll'$ and for each $\omega \in \Omega$, each $C \in P$, it holds that $C_\omega \ll' C$ if $C_\omega \neq C$.

From Lemma 7.2, it holds:

- S is a BE-preferred answer set of $(P, \ll)$ iff
  S is a BE-preferred answer set of $(P^S, \ll^S)$ iff
  for each $r \in P^S \setminus \Gamma_{P^S,S}$, if $hd(r) \notin S$ then

$$bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{P^S,S} \text{ and } r \ll^S r'\} \neq \emptyset$$

- S is a BE-preferred answer set of $(Q, \ll')$ iff
  S is a BE-preferred answer set of $(Q^S, \ll'^S)$ iff
  for each $r \in Q^S \setminus \Gamma_{Q^S,S}$, if $hd(r) \notin S$ then

$$bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{Q^S,S} \text{ and } r \ll'^S r'\} \neq \emptyset$$

It is not difficult to see that $Q^S = P^S + \Omega$ and $\Gamma_{Q^S,S} = \Gamma_{P^S,S} + \Omega$. Therefore it follows that $Q^S \setminus \Gamma_{Q^S,S} = P^S \setminus \Gamma_{P^S,S}$. It follows immediately that for each $r \in Q^S \setminus \Gamma_{Q^S,S} = P^S \setminus \Gamma_{P^S,S}$,

$$\{r' \mid r' \in \Gamma_{P^S,S} \text{ and } r \ll^S r'\} = \{r' \mid r' \in \Gamma_{Q^S,S} \text{ and } r \ll'^S r'\}$$

Since S is a BE-preferred answer set of $(P, \ll)$, it holds:

For each $r \in P^S \setminus \Gamma_{P^S,S}$, if $hd(r) \notin S$ then

$$bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{P^S,S} \text{ and } r \ll^S r'\} \neq \emptyset$$

Because $Q^S \setminus \Gamma_{Q^S,S} = P^S \setminus \Gamma_{P^S,S}$ and for each $r \in Q^S \setminus \Gamma_{Q^S,S} = P^S \setminus \Gamma_{P^S,S}$, $\{r' \mid r' \in \Gamma_{P^S,S} \text{ and } r \ll^S r'\} = \{r' \mid r' \in \Gamma_{Q^S,S} \text{ and } r \ll'^S r'\}$, it follows:

For each $r \in Q^S \setminus \Gamma_{Q^S,S}$, if $hd(r) \notin S$ then

$$bd^-(r) \cap \{hd(r') \mid r' \in \Gamma_{Q^S,S} \text{ and } r \ll'^S r'\} \neq \emptyset$$

S is hence a BE-preferred answer set of $(P + \Omega, \ll')$. □

Appendix M. Lemma 5.4

Let $\mathcal{K}$ be the sensible class of all consistent basic knowledge bases closed under transposition. Let *att* be the attack relation assignment defined as in Example 5.2. Following assertions hold:

1. For each $K' \in \mathcal{K}$, the stable extensions of $(AR_{K'}, att(K'))$ and $(AR_{K'}, att_{bs}(K'))$ coincide.
2. The reverse of Theorem 5.1 does not hold for $\mathcal{K}$.

Proof. Let K be the knowledge base defined in Example 5.1. It is clear that $K \in \mathcal{K}$. Let $A_0, \dots, A_3$ be arguments defined as in Fig. 6.

For ease of reference, we recall the attack relation assignment att defined for $\mathcal{K}$ below:

For each $K' \in \mathcal{K}$:

$$att(K') = \begin{cases} att_{bs}(K') & \text{if } \{A_0, A_1\} \not\subseteq AR_{K'} \\ att_{bs}(K') \cup \{(A_0, A_1), (A_1, A_0)\} & \text{otherwise} \end{cases}$$

As $K \in \mathcal{K}$, it follows from the elaboration in Example 5.1 that att does not satisfy the property of subargument structure for $\mathcal{K}$.

1. Let $K' \in \mathcal{K}$. We show that the stable extensions of $(AR_{K'}, att(K'))$ and $(AR_{K'}, att_{bs}(K'))$ coincide.
   If $att(K') = att_{bs}(K')$, there is nothing to prove here.
   Let $att(K') = att_{bs}(K') \cup \{(A_0, A_1), (A_1, A_0)\}$. Since K' is closed under transposition, $A_2, A_3 \in AR_{K'}$.
   (a) Let E be a stable extension of $(AR_{K'}, att(K'))$. It is obvious that E is conflict-free wrt $att_{bs}(K')$.
   We show that E attacks every argument not belonging to E wrt $att_{bs}(K')$.
   Suppose the contrary. Hence there is $X \in AR_{K'} \setminus E$ s.t. E does not attack X wrt $att_{bs}(K')$ but there is $C \in E$ s.t. $(C, X) \in att(K') \setminus att_{bs}(K')$. Therefore $(C, X) = (A_0, A_1)$ or $(C, X) = (A_1, A_0)$. Let us consider each case.
   - Let $(C, X) = (A_0, A_1)$. Hence $C = A_0$ and $X = A_1$.
     We first show that $[d_0] \in E$. Suppose the contrary. Therefore E attacks $[d_0]$ wrt $att(K')$. From $att(K') = att_{bs}(K') \cup \{(A_0, A_1), (A_1, A_0)\}$, it follows that E attacks $[d_0]$ wrt $att_{bs}(K')$. Therefore E attacks A_0 wrt $att_{bs}(K')$. Hence E attacks A_0 wrt $att(K')$. Impossible as $A_0 \in E$ and E is conflict-free wrt $att(K')$.
     We show that $A_2 \in E$. Suppose the contrary. There is hence $\exists Y \in E$ s.t. $(Y, A_2) \in att(K')$. From $att(K') = att_{bs}(K') \cup \{(A_0, A_1), (A_1, A_0)\}$, it follows that $(Y, A_2) \in att_{bs}(K')$. Therefore $(Y, [d_0]) \in att_{bs}(K')$. From $att(K') = att_{bs}(K') \cup \{(A_0, A_1), (A_1, A_0)\}$, it follows that E is not conflict-free wrt $att(K')$. Impossible as $[d_0] \in E$ and E is conflict-free wrt $att(K')$. Therefore $A_2 \in E$.
     It is clear that $(A_2, A_1) \in att_{bs}(K')$. Contradiction to the assumption that E does not attack X wrt $att_{bs}(K')$.
   - Let $(C, X) = (A_1, A_0)$. The proof is similar to the previous case.

   We have proved that E is also a stable extension of $(AR_{K'}, att_{bs}(K'))$.
   (b) Let E be a stable extension of $(AR_{K'}, att_{bs}(K'))$. It is easy to see that for each argument X, $X \in E$ iff $mbd(X) \subseteq E$.
   We only need to show that E is conflict-free wrt $att(K')$.
   Suppose the contrary that E is not conflict-free wrt $att(K')$. Therefore there are arguments $X, Y \in E$ s.t. $(X, Y) \in att(K')$. Since E is conflict-free wrt $att_{bs}(K')$, it follows that $(X, Y) = (A_0, A_1)$ or $(X, Y) = (A_1, A_0)$. Therefore $[d_i] \in E$ for both $i = 0, 1$. Therefore argument $A_0, \dots, A_3$ all belong to E. Since $(A_2, A_1), (A_3, A_0) \in att_{bs}(K')$, E is not conflict-free wrt $att_{bs}(K')$. Contradiction.
   Thus E is conflict-free wrt $att(K')$.
2. As $K \in \mathcal{K}$, it follows from the elaboration in Example 5.1 that att does not satisfy the property of subargument structure for $\mathcal{K}$. Therefore the reverse of Theorem 5.1 does not hold for $\mathcal{K}$. □

References

[1] L. Amgoud, Postulates for logic-based argumentation systems, Int. J. Approx. Reason. 55 (9) (2014) 2028–2048.

[2] L. Amgoud, P. Besnard, Bridging the gap between abstract argumentation systems and logic, in: Proc. 3rd International Conf. on Scalable Uncertainty, SUM'09, 2009, pp. 12–27.
[3] L. Amgoud, C. Cayrol, Inferring from inconsistency in preference-based argumentation framework, J. Autom. Reason. 29 (2) (2002) 197–215.
[4] K. Apt, H. Blair, A. Walker, Towards a theory of declarative knowledge, in: J. Minker (Ed.), Foundation of Deductive Data Bases and Logic Programming, Morgan Kaufmann, 1988, pp. 89–148.
[5] C. Baral, M. Gelfond, Logic programming and knowledge representation, J. Log. Program. 19/20 (1994) 73–148.
[6] T.J.M. Bench-Capon, Persuasion in practical argument using value-based argumentation frameworks, J. Log. Comput. 13 (3) (2003) 429–448.
[7] A. Bondarenko, P. Dung, R. Kowalski, F. Toni, An abstract, argumentation-theoretic approach to default reasoning, Artif. Intell. 93 (1997) 63–101.
[8] G. Brewka, Preferred subtheories: an extended logical framework for default reasoning, in: Proc. of IJCAI'89, Morgan Kaufmann, 1989, pp. 1043–1048.
[9] G. Brewka, Cumulative default logic: in defence of nonmonotonic inference rules, Artif. Intell. 50 (1991) 183–205.
[10] G. Brewka, T. Eiter, Preferred answer sets for extended logic programs, Artif. Intell. 109 (1999) 297–356.
[11] G. Brewka, I. Niemelä, M. Truszczynski, Preferences and nonmonotonic reasoning, AI Mag. 29 (4) (2008) 69–78.
[12] M. Caminada, L. Amgoud, On the evaluation of argumentation formalisms, Artif. Intell. 171 (2007) 286–310.
[13] B.A. Davey, H.A. Priestley, Introduction to Lattices and Order, Cambridge University Press, 2002.
[14] J. Delgrande, A first order conditional logic for prototypical properties, Artif. Intell. 33 (1987) 105–130.
[15] J. Delgrande, T. Schaub, H. Tompits, A framework for compiling preferences in logic programs, Theory Pract. Log. Program. 3 (2) (2003) 129–187.
[16] P. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, Artif. Intell. 77 (2) (1995) 321–358.
[17] P. Dung, An axiomatic analysis of structured argumentation for prioritized default reasoning, in: Proc. of ECAI 2014, 2014, pp. 267–272.
[18] P. Dung, T. Son, An argument-based approach to reasoning with specificity, Artif. Intell. 133 (2001) 35–85.
[19] P.M. Dung, P.M. Thang, Closure and consistency and logic-associated argumentation, J. Artif. Intell. Res. 49 (2014) 79–109.
[20] D. Gabbay, Theoretical foundation for nonmonotonic reasoning in expert systems, in: K.R. Apt (Ed.), Proc. of the NATO Advanced Study Institute on Logics and Models of Concurrent Systems, Springer Verlag, 1985, pp. 439–457.
[21] A. Garcia, G. Simari, Defeasible logic programming: an argumentative approach, Theory Pract. Log. Program. 4 (1–2) (2004) 95–138.
[22] M. Gelfond, V. Lifschitz, The stable model semantics for logic programming, in: Proc. of 5th ICLP, MIT Press, 1988, pp. 1070–1080.
[23] M. Gelfond, V. Lifschitz, Classical negation in logic programs and disjunctive databases, New Gener. Comput. 9 (1991) 365–385.
[24] M. Gelfond, T.C. Son, Reasoning with prioritized defaults, in: Proc. of LPKR, 1997, pp. 164–223.
[25] N. Gorogiannis, A. Hunter, Instantiating abstract argumentation with classical logic arguments: postulates and properties, Artif. Intell. 175 (2011).
[26] H. Geffner, J. Pearl, Conditional entailment: bridging two approaches to default reasoning, Artif. Intell. 53 (1992) 209–244.
[27] A. Hunter, M. Williams, Aggregating evidence about positive and negative effects of treatments, Artif. Intell. Med. 56 (2012) 173–190.
[28] A. Hunter, S. Woltran, Structural properties for deductive argument systems, in: L.C. van der Gaag (Ed.), Proc. of 12th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, in: Lecture Notes in Computer Science, vol. 7598, Springer Verlag, 2013, pp. 278–289.
[29] S. Kraus, D. Lehmann, M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artif. Intell. 44 (1990) 167–207.
[30] V. Lifschitz, Answer set planning, in: Proc. of the ICLP-99, 1999, pp. 23–37.
[31] R. Loui, Defeat among arguments: a system of defeasible inference, Comput. Intell. 3 (1987) 100–106.
[32] M. Gelfond, N. Leone, Logic programming and knowledge representation – the A-Prolog perspective, Artif. Intell. 138 (2002) 3–38.
[33] D.C. Martinez, A.J. Garcia, G.R. Simari, On acceptability in abstract argumentation frameworks with an extended defeat relation, in: P.E. Dunne, T.J.M. Bench-Capon (Eds.), Proc. of Int. Conference on "Computational Models of Arguments", IOS Press, 2006.
[34] S. Modgil, H. Prakken, A general account of argumentation with preferences, Artif. Intell. 197 (2013) 361–397.
[35] S. Modgil, H. Prakken, The ASPIC+ framework for structured argumentation: a tutorial, J. Argument Comput. 5 (2014) 31–62.
[36] D. Nute, Defeasible logic, in: Handbook of Logic in Artificial Intelligence and Logic Programming, University Press, 1994.
[37] D. Nute, M. Lewis, d-Prolog: a user's manual, Tech. rep., University of Georgia, Athens, GA, 1986.
[38] J. Pollock, Defeasible reasoning, Cogn. Sci. 11 (4) (1987) 481–518.
[39] H. Prakken, An abstract framework for argumentation with structured arguments, J. Argument Comput. 1 (2) (2010) 93–124.
[40] H. Prakken, G. Sartor, Argument-based extended logic programming with defeasible priorities, J. Appl. Non-Class. Log. 7 (1) (1997) 25–75.
[41] R. Reiter, A logic for default reasoning, Artif. Intell. 13 (1980) 81–132.
[42] R. Reiter, G. Criscuolo, Some representational issues in default reasoning, Comput. Math. Appl. 9 (1983) 15–27.
[43] J. Rintanen, Lexicographic priorities in default logics, Artif. Intell. 106 (1998) 221–265.
[44] T. Schaub, K. Wang, A comparative study of logic programs with preferences, in: Proc. of IJCAI 2001, Morgan Kaufmann Publishers, 2001, pp. 597–602.
[45] L.C. van der Gaag (Ed.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Springer Verlag, 2013.
[46] G. Vreeswijk, Abstract argumentation systems, Artif. Intell. 90 (1997) 225–279.