The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)

On Structured Argumentation with Conditional Preferences

Phan Minh Dung*
Department of Computer Science
Asian Institute of Technology
Thailand

Phan Minh Thang
BUUIC College of Burapha University
Thailand

Tran Cao Son
Department of Computer Science
New Mexico State University
Las Cruces, NM, USA

Abstract

We study defeasible knowledge bases with conditional preferences (DKB). A DKB consists of a set of undisputed facts and a rule-based system that contains different types of rules: strict, defeasible, and preference. A major challenge in defining the semantics of DKB lies in determining how conditional preferences interact with the attack relations represented by rebuts and undercuts, between arguments.

We introduce the notions of preference attack relations as sets of attacks between preference arguments and the rebuts or undercuts among arguments as well as of preference attack relation assignments which map knowledge bases to preference attack relations. We present five rational properties (referred to as regular properties), the inconsistency-resolving, effective rebuts, context-independence, attack monotonicity and link-orientation properties generalizing the properties of the same names for the case of unconditional preferences.

Preference attack relation assignment are defined as *regular* if they satisfy all regular properties. We show that the set of regular assignments forms a complete lower semilattice whose least element is referred to as the canonical preference attack relation assignment. Canonical attack relation assignment represents the semantics of preferences in defeasible knowledge bases as intuitively, it could be viewed as being uniquely identified by the regular properties together with the principle of minimal removal of undesired attacks. We also present the *normal* preference attack relation assignment as an approximation of the canonical attack relation assignment.

Keywords: Conditional preferences, regular properties, minimal removal principle, preference attack relation assignments, canonical and normal preference attack relation assignments.

Introduction

There are extensive research on rule-based systems with prioritized rules (see, e.g., Modgil and Prakken (2014); Amgoud and Cayrol (2002); Brewka (1989); Delgrande, Schaub, and Tompits (2003); Schaub and Wang (2001); Brewka and Eiter (1999)). Prakken and Sartor (1997) is arguably the first attempt to study the application of priorities of defeasible rules to define a preference order between arguments and then using this preference order to remove undesired attacks. Amgoud and Cayrol (2002) has studied different ways to define preference order between arguments. Prakken (2010) and Modgil and Prakken (2013; 2014) have proposed ASPIC+, a rich framework for structured argumentation with prioritized rules with several distinct systems of preference orders between arguments.

The rich diversity of proposed attack relations poses a serious challenge for any potential user of structured argumentation as such a user would have to decide which attack relation should be selected and implemented for her/his domain. To address this problem, general principles for the characterization and evaluation of alternative attack relations for rule-based systems are needed. Caminada and Amgoud (2007) have introduced the postulates of consistency and closure for argument-based systems. A subargument closure postulate stating that any extension should contain all subarguments of its arguments has been studied by Martinez, Garcia, and Simari (2006), Amgoud (2014), Modgil and Prakken (2013).

The three proposed postulates give important insights into the characteristics of attack relations in structured argumentation. But they are not sufficient to guarantee intuitive semantics, as they do not take into account the preferences among defeasible rules. Dung (2016) has proposed a set of simple and intuitive properties, referred to as regular properties, to study the attack relations. Dung and Thang (2018) have showed that these properties coupled with the minimal removal principle stating that the removal of attacks should be kept to a minimum, uniquely identify the canonical attack relations that could be viewed as representing the intended attack relations for structured argumentation with prioritized rules.

Until now, axiomatic analysis of attack relations has been carried out mostly for systems with unconditional preferences. In the next example, an extended version of the Sherlock Holmes example from Dung (2016); Dung and Thang (2018), we can see that conditional preferences could be a natural part of many defeasible knowledge bases.

Example 1 (Adapted from Dung (2016))**.** *Sherlock Holmes is investigating a case involving three persons P_1, P_2 and S together with the dead body of a big man. Furthermore S is a small child who cannot kill a big man and P_1 is a beneficiary from the dead of the big man.*

The case could be represented by the following knowledge

*The lead and corresponding author

base.

1. *The knowledge that one of the persons is the murderer is represented by three strict rules:*
   $r_1: \ Inno(P_1), Inno(S) \rightarrow \neg Inno(P_2)$[1]
   $r_2: \ Inno(P_2), Inno(S) \rightarrow \neg Inno(P_1)$
   $r_3: \ Inno(P_1), Inno(P_2) \rightarrow \neg Inno(S)$
2. *The legal principle that people are considered innocent until proven otherwise could be represented by three defeasible rules* $d_i: \ \Rightarrow Inno(P_i)$ *for* $i = 1, 2$ *together with* $d: \ \Rightarrow Inno(S)$
3. *A "rule-of-thumb" for the investigation is to find out whether the possible suspects have any motives and to focus the investigation on the one with strong motive to commit the crime. Such "rule-of-thumb" can be represented by two conditional preferences:*
   - $\pi_1: Have\_Motive(P_1), \neg Have\_Motive(P_2) \rightarrow d_1 \prec d_2$.
   - $\pi_2: \neg Have\_Motive(P_1), Have\_Motive(P_2) \rightarrow d_2 \prec d_1$.

   The rules state that if P_i *has a motive and* P_j $(i \neq j)$ *does not have a motive then the default* d_j *is more preferred than* d_i.
4. *A good reason for having a motive to kill is to be a beneficiary from the dead of the deceased.*
   - $r_4: \ Beneficiary(P_1) \rightarrow Have\_Motive(P_1)$
   - $r_5: \ Beneficiary(P_2) \rightarrow Have\_Motive(P_2)$
5. *Peoples are normally assumed not to have motives to kill.*
   $d_3: \Rightarrow \neg Have\_Motive(P_1) \quad d_4: \Rightarrow \neg Have\_Motive(P_2)$.
6. *The facts that S is a small child and* P_1 *is a beneficiary from the dead of the big man are represented by atoms* $Inno(S), Beneficiary(P_1)$.

The considered knowledge base is represented by $K_{sh} = (RS_{sh}, RD_{sh}, RP_{sh}, BE_{sh})$ *with* $RS_{sh} = \{r_1, r_2, r_3, r_4, r_5\}$ *containing of the strict rules,* $RD_{sh} = \{d_1, d_2, d, d_3, d_4\}$ *consisting of the defeasible rules and* $RP_{sh} = \{\pi_1, \pi_2\}$ *consisting of the preference rules while* $BE_{sh} = \{Inno(S), Beneficiary(P_1)\}$ *represents the set of facts.*

Relevant arguments are given in figures 1. Among these, X_1 *is a preference argument (formal definition a bit later) whose conclusion is* $d_1 \prec d_2$. *It is easy to see that* X_1 *is defeasible since* d_4 *is defeasible.*

In this paper, we develop a framework for dealing with conditional preferences in structured argumentation. We define defeasible knowledge bases with conditional preferences (DKB) over a rule-based system. The semantics of a DKB is defined by a preference-based argumentation framework whose set of arguments consists of the arguments of the DKB and its direct undercuts and rebuts and whose set of attacks consists of the set of attacks created from the rebuts and undercuts of the DKB and a preference attack relation—a set of attacks between preference arguments and the rebuts or undercuts among arguments. We prove that this semantics is a conservative generalization of the semantics of basic knowledge bases when restricted to knowledge bases without preferences.

[1] 'Inno' stands for 'Innocent'.

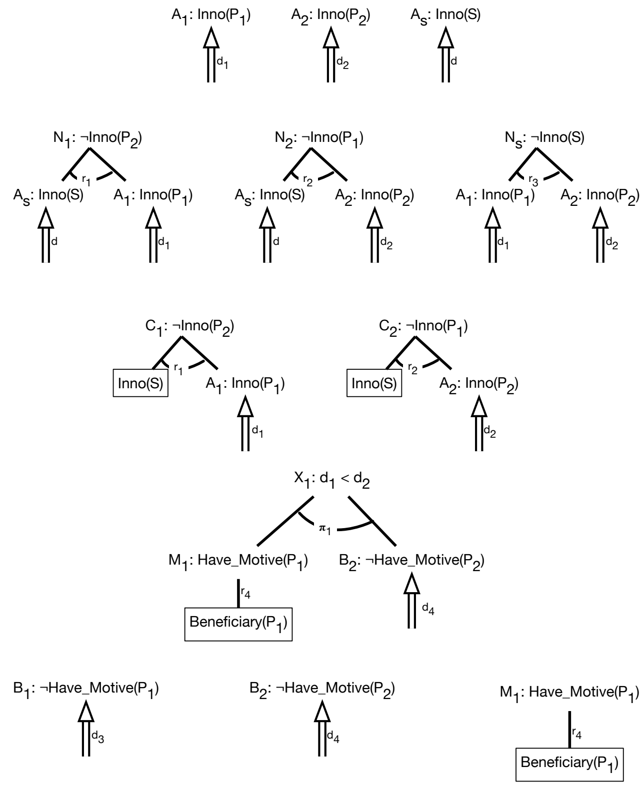

Figure 1: Sherlock Holmes Arguments

To analyze attack relations, we present the notion of a preference attack relation assignment which maps knowledge bases to preference attack relations and discuss five regular properties, the context-independence, effective rebuts, inconsistency-resolving, attack monotonicity, and link-orientation properties of preference attack relations generalizing properties of the same names in the unconditional cases in Dung (2016) and Dung and Thang (2018). We say an assignment is regular if it satisfies all regular properties and show that the set of regular assignments forms a complete lower semilattice whose least element is referred to as the canonical preference attack relation assignment. Canonical attack relation assignment could be viewed as representing the semantics of preferences in defeasible knowledge bases as intuitively, it is uniquely identified by the regular properties together with the principle of minimal removal of undesired attacks.

We also define a notion of a normal preference attack relation assignment that satisfies all regular properties and discuss sufficient conditions under which canonical and normal preference attack relation assignments coincide.

Preliminary: Abstract Argumentation

An abstract argumentation framework Dung (1995) is a pair $AF = (AR, At)$ where AR is a set of arguments and $att \subseteq AR \times AR$. $(A, B) \in At$ means that A *attacks* B. A set of arguments 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. A set of arguments S *defends* an argument A if S attacks every argument that attacks A. S is *admissible* if S is conflict-free and defends each argument in it. A *preferred extension* is a maximal admissible set of arguments. 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.

We introduce the concept of semilattice. A partial order[2] $\leq$ on a set S is a *complete lower-semilattice* (Davey and Priestley (2002)) iff each non-empty subset of S has a infimum w.r.t. $\leq$. It follows immediately that each non-empty complete lower semilattice S has an unique least element.

Defeasible KBs with Conditional Preferences

We assume a non-empty set $\mathcal{L}$ of ground atoms. We distinguish between

- *domain atoms* representing propositions about the concerned domains;
- *non-applicability atoms* (or *undercut atoms*) of the form ab_d representing the non-applicability of a defeasible rule d even if its premises hold; and
- *preference atoms (or p-atom for short)* of the form $d \prec d'$ (also $d' \succ d$) where d, d' are distinct defeasible rules, representing the preference of d' over d.

A domain atom a (resp. its negation $\neg a$) is called a positive (resp. negative) *domain literal*. a and $\neg a$ are said to be the complementary of each other. $\prec$ is a strict partial order. We write $\neg(d \prec d')$ for $d' \prec d$ and say that $d \prec d'$ and $d' \prec d$ are complementary of each other. A set of literals is said to be *contradictory* iff it contains an atom and its complement.

We distinguish between strict and defeasible rules as often done in the literature (e.g., Modgil and Prakken (2013, 2014); Garcia and Simari (2004)).

Definition 1.
- *A* defeasible *(resp.* strict*)* domain rule *r is of the form* $b_1, \dots, b_n \Rightarrow h$ *(resp.* $b_1, \dots, b_n \rightarrow h$*) where $b_1, \dots, b_n$ are domain literals, and h is*
  - *a domain literal, or*
  - *a non-applicability atom ab_d (and r is also often called an* undercut rule*).*
- *A* preference rule (or p-rule for short) *r is a strict rule of the form $b_1, \dots, b_n \rightarrow d_0 \prec d_1$ where $b_1, \dots, b_n$ are domain literals and d_0, d_1 are distinct defeasible domain rules.*
- *The set $\{b_1, \dots, b_n\}$ (resp. the literal h or the preference atom $d_0 \prec d_1$) is referred to as the* body *(resp.* head*) of r and denoted by $bd(r)$ (resp. $hd(r)$).*

When the body of a preference rule r is empty, we say that r represents an unconditional preference.

Definition 2.
- *A* rule-based system with conditional preferences *(or simply a* rule-based system*) is a quadruple $\mathcal{R} = (RS, RD, RP, RT)$ where*
  - *RS is a set of strict domain rules, and*
  - *RD is a set of defeasible domain rules, and*
  - *RP is a set of p-rules, and*
  - *RT is the set of all* transitive preference rules *that are ground instances of the transitive rule* $x \prec y,\ y \prec z\ \rightarrow\ x \prec z$ *where x, y, z range over RD and x, y, z are pairwise distinct.*
- *A* defeasible knowledge base with conditional preferences *(DKB) (or a* knowledge base*) over $\mathcal{R}$ is a pair $K = (\mathcal{R}, BE)$ consisting of a rule-based system $\mathcal{R}$, and a set of ground domain literals BE, the* base of evidence *of K, representing unchallenged facts, observations.*
- *A DKB K is* ***basic*** *if it contains no precedence rules.*

A DKB K is often written directly as a tuple (RS, RD, RP, RT, BE). For simplicity, *we sometimes omit the set RT in the description of a rule-based system.*

Arguments of a knowledge base are defined as follows.

Definition 3. *Let $K = (RS, RD, RP, RT, BE)$ be a KB. An* **argument** *w.r.t. K is a proof tree defined as follows:*

1. *For each $\alpha \in BE$, $[\alpha]$ is an argument with conclusion α.*
2. *Let r be a rule of the forms $\alpha_1, \dots, \alpha_n \Rightarrow (\rightarrow)\, \alpha,\quad n \geq 0$ from $RS \cup RD \cup RP \cup RT$ and $A_1, \dots, A_n$ be arguments with conclusions α_i, $1 \leq i \leq n$, respectively. Then $A = [A_1, \dots, A_n, r]$ is an argument.*
   α and r are called the ***conclusion*** *and the* **last rule** *of A and are denoted by* ***cnl(A)*** *and* ***last(A)****, respectively. If α is a domain (undercut) atom then A is called a* domain (undercut) argument.
3. *Let $A_1, \dots, A_n$ be arguments with p-atoms α_i as conclusions, $1 \leq i \leq n$, respectively. Then $A = [A_1, \dots, A_n]$ is a* preference argument *with the* ***conclusion*** $cnl(A) = \{\alpha_1, \dots, \alpha_n\}$.
4. *Each argument w.r.t. K is obtained by applying the above steps finitely many times.*

In the following, a preference argument $[A_1, \dots, A_n]$ is identified with any argument obtained from reordering of its components or removal of duplicates among its arguments. The set of all arguments (resp. preference arguments) w.r.t. K is denoted by AR_K (resp. $AR_{p,K}$). For $S \subseteq AR_K$, $cnl(S) = \{l \mid A \in S, l$ appears in $cnl(A)\}$.

An argument B is a *subargument* of an argument A iff (*i*) $B = A$; or (*ii*) $A = [A_1, \dots, A_n, r]$ or $A = [A_1, \dots, A_n]$ and B is a subargument of some A_i. B is a *proper subargument* of A if B is a subargument of A and $B \neq A$.

A rule r is said to appear in an argument A if $A = [A_1, \dots, A_n, r]$ or r appears in some proper subargument A. The set of defeasible rules appearing in an argument A is denoted by $dr(A)$. A is *strict* if no defeasible rule appears in it. It is *defeasible* otherwise. A is called *basic defeasible* iff $last(A)$ is defeasible. The set of last defeasible rules in A, denoted by $ldr(A)$, is $\{last(A)\}$ if A is basic defeasible, otherwise it is equal $ldr(A_1) \cup \ldots \cup ldr(A_n)$ where $A = [A_1, \dots, A_n, r]$ or $A = [A_1, \dots, A_n]$.

A undercuts B (at B') if B' is basic defeasible subargument of B and $cnl(A) = ab_{last(B')}$. *A rebuts B* (at B') if B' is a basic defeasible subargument of B, both A, B' are domain arguments and the conclusions of A and B' are contradictory. *A directly rebuts or undercuts B* iff A rebuts or undercuts B (at B) respectively. We often simply say A re-

[2] A reflexive, transitive and antisymmetric relation

buts or undercuts B if it is not relevant to specify the exact place at which A rebuts or undercuts B.

For a knowledge base K, let $rbut_K = \{(A, B) \mid A \text{ rebuts } B\}$ and $ucut_K = \{(A, B) \mid A \text{ undercuts } B\}$. Furthermore, $dbut_K = \{(A, B) \mid A$ directly rebuts $B\}$ and $dcut_K = \{(A, B) \mid A$ directly undercuts $B\}$.

Example 2. *Let us revisit Ex. 1.*[3] *Most relevant arguments w.r.t. K_{sh} are given in Figure 1:* (i) *all arguments except M_1 are defeasible;* (ii) *A_1, A_2, A_s, B_1, and B_2 are basic defeasible; and* (iii) *X_1 is a preference argument.*

As we interpret the strict rules and the facts in the base of evidences of a knowledge base as representing the unchallenged knowledge and observation of the concerned domain, it is therefore natural to expect that they are consistent as defined in the following definition.[4]

Definition 4. *Let $\mathcal{R}$ be a rule-based system.*

1. *The* **closure** *of a set of literals X over $\mathcal{L}$ w.r.t. a rule-based system $\mathcal{R}$, denoted by $CN_{\mathcal{R}}(X)$, is the union of X and the set of conclusions of all strict arguments w.r.t. the knowledge base $(\mathcal{R}, X_{dom})$ where X_{dom} is the set of domain literals in X. X is said to be* closed *w.r.t. $\mathcal{R}$ iff $X{=}CN_{\mathcal{R}}(X)$. We write $X \vdash_{\mathcal{R}} l$ iff $l \in CN_{\mathcal{R}}(X)$.*
   X is said to be inconsistent *w.r.t. $\mathcal{R}$ iff its closure $CN_{\mathcal{R}}(X)$ is contradictory. X is* consistent *w.r.t. $\mathcal{R}$ iff it is not inconsistent.*
2. *A knowledge base K is said to be* consistent *iff its base of evidence BE is consistent w.r.t. $\mathcal{R}$.*
3. *$\mathcal{R}$ is said to be consistent iff the knowledge base $(\mathcal{R}, \emptyset)$ is consistent.*
4. *$\mathcal{C}_{\mathcal{R}}$ denotes the set of all consistent KBs over $\mathcal{R}$.*

Assumption: From now on, whenever we refer to rule-based systems, we always mean consistent ones.

Preference-based Argumentation Framework of Defeasible KB with Conditional Preferences

The previous section defines different types of arguments and different types of attacks (like rebuts or undercuts) among them. However, the discussion thus far does not yet specify the semantics of a knowledge base. This is because it does not yet take into consideration preference arguments. In this section, we will address this problem.

Example 3. *Consider K_{sh} from Ex. 2. Let us consider the arguments C_1, A_2, and X_1 (Fig. 2).*

By definition, C_1 directly rebuts A_2 (i.e. $(C_1, A_2) \in dbut_{K_{sh}}$). The presence of X_1 indicates that d_2 is preferred to d_1. This means that the direct rebut (C_1, A_2) could be removed from consideration whenever X_1 is acceptable. In other words, we can say that X_1 attacks the rebuttal of A_2 by C_1 and represent this attack *in the form $(X_1, (C_1, A_2))$.*

[3]Note that transitive rules are not mentioned explicitly in the representation of the knowledge base.

[4]Note that if the evidences together with the strict rules are not consistent then they in fact represent defeasible knowledge. As such they should be represented by the defeasible rules, not by strict rules and facts.

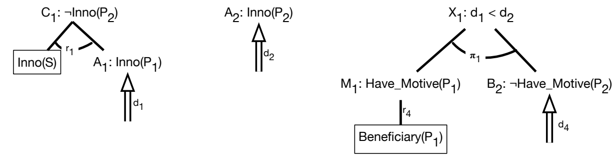

Figure 2: C_1 and A_2 and the preference argument X_1

The above example shows that preference arguments introduce a new type of attacks, i.e., attacks against direct rebuttals among arguments. We will refer to an attack of this new type as an *attack by preference arguments*. Formally, a set of attacks by preference arguments could be represented by a binary relation in $AR_{p,K} \times dbut_K$. Clearly, the choice of the set defining the attacks by preference arguments will determine which arguments will be accepted. The next example illustrates this problem.

Example 4. *Consider again K_{sh} from Ex. 1. It is not difficult to see: $dbut_{K_{sh}} = \{(N_1, A_2), (N_2, A_1), (C_1, A_2), (C_2, A_1), (M_1, B_1)\}$. Furthermore, $dcut_{K_{sh}} = \emptyset$ and the only preference argument w.r.t. K_{sh} is X_1, i.e., $AR_{p,K_{sh}} = \{X_1\}$.*

We consider two situations with respect to the choice of the set of attacks by preference arguments:

- *Assume that we select $\{(X_1, (N_1, A_2)), (X_1, (C_1, A_2))\}$ as the set of attacks by preference arguments. Then, because X_1 is not attacked by any argument w.r.t. K_{sh}, we can conclude that X_1 is always acceptable, and hence, we should remove the attacks (N_1, A_2) and (C_1, A_2) from consideration when determining the acceptability of arguments w.r.t. K_{sh}. Hence there is no attack against A_2 which shows that P_2 is innocent and P_1 is not.*
- *Assume that we select the empty set as the set of attacks by preference arguments. Then argument X_1 plays no role in the attack relations between $A_1, A_2, N_1, N_2, C_1, C_2$. There are then two acceptable scenarios: One in which P_1 is innocent and the other in which P_2 is innocent.*

Ex. 4 shows that the key question in dealing with preference arguments is "*what is the intuitively expected set of attacks by preference arguments?*" More formally, given a knowledge base K, which subset of $AR_{p,K} \times dbut_K$ should be selected as the set of attacks by preference arguments. In the rest of this paper, we will propose a solution for this question. Observe that we only consider attacks by preference arguments against rebuttals but not against undercuts. This is because, in our view, undercuts are preference independent, which is also in agreement with the view expressed by many others (e.g., by Modgil and Prakken (2013)).

At this point, it is instructive to recall that we are dealing with conditional preferences. Consequently, an attack against a rebuttal by a preference argument does not yet mean that the rebuttal should be eliminated from consideration. Let us consider again the argument X_1 in Ex. 4. As we have argued, X_1 attacks the rebuttal of A_2 by C_1. But this attack is effective only if X_1 itself is accepted. In other words, whether the rebut (C_1, A_2) is accepted or not depends on the acceptance of X_1. This insight suggests that rebuts also need to be defended like arguments. We formalize this insight by viewing direct rebuts like (C_1, A_2) as arguments. A question, that arises immediately, is that *how are the "rebut-arguments" attacked?* Clearly, a rebut argument like (C_1, A_2) could be attacked in two ways: Attacking the "rebutting connection" like the way X_1 attacks (C_1, A_2) or attacking the source C_1.

As direct rebuttals are considered as arguments, it is convenient to also consider direct undercuts as arguments. But in contrast to "rebut arguments," there is only one way to attack "undercut arguments," namely attacking their source.

We can now introduce a novel notion of preference argumentation framework capturing the key conceptual ideas discussed above.

Definition 5. (Preference-Based AF) *Given a knowledge base K, a* preference-based argumentation framework (PAF) *of K is an abstract argumentation framework $\mathcal{F}_K = (ARP_K, att_{\mathcal{F}})$ where*

- $ARP_K = AR_K \cup dbut_K \cup dcut_K$, *and*
- $att_{\mathcal{F}} = datt_K \cup patt$ *where*

(a) $datt_K \subseteq (dbut_K \cup dcut_K) \times ARP_K$ *such that* $((A, B), X) \in datt_K$ *iff*

- $X \in AR_K$ *and B is a subargument of X, or*
- $X{=}(C, D){\in}(dbut_K{\cup}dcut_K)$ *and B is a subargument of C.*

(b) $patt \subseteq AR_{p,K} \times dbut_K$ *where $patt$ is often referred to as the* preference attack relation *(or just* p-attack relation*) of $\mathcal{F}$.*

For ease of reference, we often refer to a PAF of K by the triple $(ARP_K, datt_K, patt)$.

As $\mathcal{F}_K$ is an abstract argumentation framework, its semantics is fully defined. As such, we can define the semantics of K by its PAFs. Intuitively, $\mathcal{F}_K$ contains arguments w.r.t. K together with the direct attacks and direct rebuts w.r.t K and the attacks in $\mathcal{F}_K$ are of two types: attacks agains arguments (type (a)) and attacks against attacks (type (b)). In this sense, PAF is similar to extended argumentation frameworks (see, e.g., Gabbay (2009); Baroni et al. (2011); Modgil (2009); Hanh et al. (2011); Young et al. (2018)) in which attacks against attacks are allowed or metalevel arguments are considered. The key difference between PAFs and extended argumentation frameworks is that since PAF are abstract argumentation frameworks, we do not need to define new semantics for PAFs.

The next theorem shows that the concept of PAF is a conservative generalization of the semantics of basic knowledge bases, i.e., when restricted to basic knowledge bases, our new concept of PAF delivers the same semantics captured in the traditional view of undercuts and rebuts as attacks.

Theorem 1. *Let $\mathcal{R}$ be a basic rule-based system and $K \in \mathcal{C}_{\mathcal{R}}$. Let $\mathcal{F}_0$ be the argumentation framework $(AR_K, rbut_K \cup ucut_K)$ and $\mathcal{F} = (ARP_K, datt_K)$ be an PAF of K.*[5] *Then, (*i*) if E is a complete extension of $\mathcal{F}$ then $E \cap AR_K$ is a complete extension of $\mathcal{F}_0$; and (*ii*) if G is a complete extension of $\mathcal{F}_0$ then $G \cup \{ (X, Y) \mid X \in G$ and $(X, Y) \in dbut_K \cup dcut_K \}$ is a complete extension of $\mathcal{F}$.*

[5]Since there is no preference argument, the preference attack relation component is empty.

Definition 5 precisely defines $datt_K$ but leaves the set $patt$ unspecified. As we have alluded to above, the choice of $patt$ will affect the semantics of K as it is specified by $\mathcal{F}_K = (ARP_K, datt_K, patt)$. The next example discusses this issue in more detail.

Example 5. *Reconsider Ex. 4, we can easily check: (***i***) $\mathcal{F}_1{=}(ARP_{K_{sh}}, datt_{K_{sh}}, patt_1)$, where $patt_1{=}\{(X_1, (N_1, A_2)), (X_1, (C_1, A_2))\}$, has a unique stable (complete, preferred) extension that delivers the conclusion that P_2 is innocent in accordance with the 1st choice in Ex. 4; (***ii***) $\mathcal{F}_2 = (ARP_{K_{sh}}, datt_{K_{sh}}, patt_2)$ where $patt_2 = \emptyset$. $\mathcal{F}_2$ has two stable extensions that deliver the two scenarios in the 2nd choice in Ex. 4.*

$\mathcal{F}_2$ in Example 5 shows that not every p-attack relation yields the expected and intuitive semantics for a DKB. The extreme case of $patt_2 = \emptyset$ is equivalent to not considering any preference in the DKB at all. Nevertheless, all PAFs—as defined in Definition 5- satisfies both the closure and subargument closure postulates introduced in Caminada and Amgoud (2007); Modgil and Prakken (2013); Martinez, Garcia, and Simari (2006); Amgoud (2014). We adapt these postulates to our notation below.

Note that for any set $S \subseteq ARP_K$, we define $cnl(S) = cnl(S \cap AR_K)$.

- $patt$ satisfies the *consistency postulate* (resp. *closure postulate*) iff for each complete extension E of $\mathcal{F}_K$, $cnl(E)$ is consistent (resp. closed);
- $patt$ satisfies the *subargument closure postulate* iff for each complete extension E of $\mathcal{F}_K$, E contains all subarguments of its arguments.

Before continuing, we introduce some additional notions.

An argument A is said to be *generated by* a set of arguments $S \subseteq AR_K$ if and only if all basic defeasible subarguments of A are subarguments of arguments in S. A is said to be *generated by* a set of arguments from ARP_K if it is generated by the arguments from AR_K in this set. The next theorem shows the closure and subargument closure postulates are satisfied with any $patt$.

Theorem 2. *Let $\mathcal{R}$ be a rule based-system, $K{\in}\mathcal{C}_{\mathcal{R}}$, and $\mathcal{F}{=}(ARP_K, datt_K, patt)$ a PAF of K. If E is a complete extension of $\mathcal{F}$ then E contains all arguments generated by E.*

Theorems 1-2 show that PAFs present a reasonable generalization of the semantics of basic DKBs and can satisfy two of the basic postulates proposed for characterizing attack relations in structured argumentation. Definition 5 indicates that there are distinct preference attack relations for the same knowledge base. Example 5 shows that it is important to correctly identify this relation. The key question is then *given K, what should be selected as the intuitive p-attack relation?* We next addresses this question.

Regular Properties

Attack relations have been studied in Dung (2016) and Dung and Thang (2018) for dealing with unconditional preferences. They introduced the regular properties for attack relations and proved that under reasonable conditions on the

KBs, these properties guarantee the intuitively expected behavior of the attack relations including the satisfaction of the basic postulates introduced in Caminada and Amgoud (2007); Amgoud (2014); Modgil and Prakken (2013); Martinez, Garcia, and Simari (2006).

To save space, from now until the end of this section, we assume an arbitrary but fixed consistent rule-based system $\mathcal{R} = (RS, RD, RP, RT)$, $K \in \mathcal{C}_\mathcal{R}$, and $\mathcal{F} = (ARP_K, datt_K, patt)$ a PAF of K.

Even though the consistency postulate is very intuitive, it does not give any insight into the structure of the attack relations. We next adapt the inconsistency-resolving property from Dung (2016) and Dung and Thang (2018) to conditional preferences to shed light on the structure of attack relations underlining the consistency postulate.

Definition 6. *$patt$ satisfies the* inconsistency-resolving property ***(IR)*** *for K iff for each finite set of arguments $S \subseteq AR_K$, if $cnl(S)$ is inconsistent then there is some $(A, B) \in dbut_K \cup dcut_K$ s.t. both A, B are generated by S and there is no $P \in AR_{p,K}$ s.t. $(P, (A, B)) \in patt$.*

Intuitively, the IR property states that the inconsistency between arguments is rooted in the conflict between them w.r.t. rebutting or undercutting relations.

Theorem 3. *If $patt$ satisfies **(IR)** and E is a complete extension of $\mathcal{F}$ then $cnl(E)$ is consistent.*[6]

The next property focuses on a minimal interpretation of preferences. Specifically, in situations when $d_0 \prec d_1$ holds and both d_0 and d_1 are applicable but accepting both d_0, d_1 is not possible, then d_1 should be preferred.

Definition 7. *$patt$ satisfies the* effective rebut property ***(ER)*** *iff for each $P \in AR_{p,K}$ and $(A_0, A_1) \in dbut_K$ such that each A_i, $i = 0, 1$ contains exactly one defeasible rule d_i, it holds that $(P, (A_0, A_1)) \in patt$ iff $(d_0 \prec d_1) \in cnl(P)$.*

Preference attacks propagate from stronger to weaker rebuts. For illustration, let consider the rebuts (N_1, A_2) and (C_1, A_2) in Ex. 1. As C_1 is based on a fact $Inno(S)$ while N_1 is based on a defeasible conclusion about $Inno(S)$, it is clear that C_1 is stronger an argument then N_1. Hence the rebuttal of A_2 by C_1 is stronger than the rebuttal of A_2 by N_1. We say that N_1 is a weakening of C_1.

Let $A, B \in AR_K$ and AS be a set of domain arguments. B is said to be **a weakening of** A by AS iff

- $A = [\alpha]$ for $\alpha \in BE$, and ($B = [\alpha]$ or $B \in AS$ with $cnl(B) = \alpha$), or
- $A = [A_1, \ldots, A_n, r]$ and $B = [B_1, \ldots, B_n, r]$, $n \geq 0$ where each B_i is a weakening of A_i by AS, or
- $A = [A_1, \ldots, A_n]$ and $B = [B_1, \ldots, B_n]$, $n \geq 1$ where each B_i is a weakening of A_i by AS.

By $A \downarrow AS$ we denote the set of all weakenings of A by AS. For simplicity, we often say that A is a *strengthening* of B (by AS) if B is a weakening of A (by AS).

For illustration, consider the arguments in Example 1. For $AS = \{A_s\}$, we have $C_1 \downarrow AS = \{C_1, N_1\}$.

Suppose A' is a weakening of A and A' directly rebuts B. Hence A also directly rebuts B. Since A is a strengthening of A', a rebut from A against B should be stronger than a rebut from A' against B. Therefore, if a preference-argument P attacks (A, B), it should also attack (A', B). Similarly if A directly rebuts B then A also represents a stronger rebuttal against any weakening B' of B than against B. Therefore, if a preference argument P attacks (A, B'), it should also attack (A, B).

Definition 8. *$patt$ satisfies the* property of attack monotonicity ***(AM)*** *iff for all $A, B, P \in AR_K$, and for each weakening A' of A, for each weakening B' of B, the following assertions hold:*

- *If $(P, (A, B)) \in patt$ then $(P, (A', B)) \in patt$.*
- *If $(P, (A, B')) \in patt$ then $(P, (A, B)) \in patt$.*

Let us look at Example 1 again for illustration. Suppose $patt$ satisfies both (**ER**) and (**AM**) properties. (**ER**) implies that X_1 attacks (C_1, A_2), i.e., $(X_1, (C_1, A_2)) \in patt$. Since N_1 is a weakening of C_1, (**AM**) dictates that X_1 also attacks (N_1, A_2) (i.e. $(X_1, (N_1, A_2)) \in patt$).

The next property extends the link-oriented property, which focuses on identifying the culprit links within arguments and directs the attacks against these culprits.

Definition 9. *$patt$ satisfies the* link-oriented property ***(LO)*** *iff for all arguments $A, B, B', P \in AR_K$ such that (i) $B' \in B \downarrow BS$ for some $BS \subseteq AR_K$, (ii) $(P, (A, B)) \in patt$, and (iii) A does not rebut any argument in BS, it holds that $(P, (A, B')) \in patt$.*

A key property for attack relations is that the attacks depend only on the structure of the arguments involved. This property holds obviously for attack relations $datt_K$ that are based on the rebut and undercut relations. We will introduce shortly the *context-independence* (**CI**) property that relates the p-attack relations among different DKBs of a rule-based system and states that the attacks depend only on the structure of the arguments involved. But we first need to define the notion of attack relation assignments.

Definition 10. *A* preference-attack relation assignment *w.r.t. a rule-based system $\mathcal{R}$ is a mapping Π assigning to each knowledge base $K \in \mathcal{C}_\mathcal{R}$ a preference-attack relation $\Pi(K) \subseteq AR_{p,K} \times dbut_K$.*

Definition 11. *(***Context-Independence (CI)***) A preference-attack relation assignment Π for a rule-based system $\mathcal{R}$ is said to satisfy the property of* context-independence (CI) *iff for any two knowledge bases $K, K' \in \mathcal{C}_\mathcal{R}$ and for any three arguments A, B, P from $AR_K \cap AR_{K'}$, it holds that $(P, (A, B)) \in \Pi(K)$ iff $(P, (A, B)) \in \Pi(K')$.*

We can now present the concept of the regular preference-attack relation assignments.

For simplicity, we refer to the properties of context-independence, effective rebuts, inconsistency-resolving, attack monotonicity and link-orientation as *regular properties*.

Let π be one of the regular properties except the context-independent one. For ease of reference, we say that a preference attack relation assignment Π satisfies π if for each knowledge base $K \in \mathcal{C}_\mathcal{R}$, $\Pi(K)$ satisfies π.

Definition 12. *(***Regular Preference-Attack Relation Assignments***) A preference-attack relation assignment Π for a*

[6]Note that $cnl(E) = cnl(E \cap AR_K)$.

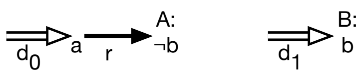

Figure 3: patt and consistency postulate

rule-based system $\mathcal{R}$ is said to be regular *iff it satisfies all regular properties.*

The set of all regular preference-attack relation assignments for $\mathcal{R}$ is denoted by $RPAA_{\mathcal{R}}$.

Semilattice Structure of Regular Assignments of Preference-Attack Relations

Let $\mathcal{R} = (RS, RD, RP, RT)$ be a rule-based system. For $\Pi, \Pi' \in RPAA_{\mathcal{R}}$, define $\Pi \subseteq \Pi'$ iff $\forall K \in \mathcal{C}_{\mathcal{R}}: \; \Pi(K) \subseteq \Pi'(K)$. It is obvious that $(RPAA_{\mathcal{R}}, \subseteq)$ is a partial order.

Definition 13. *Let $\mathcal{A}$ be a non-empty set of assignments of preference-attack relations. Define $\sqcap\mathcal{A}$ by:*

$$\forall K \in \mathcal{C}_{\mathcal{R}}: \quad (\sqcap\mathcal{A})(K) = \bigcap\{\, \Pi(K) \mid \Pi \in \mathcal{A}\,\}$$

The following simple lemma and theorem present a deep insight into the structure of regular attack assignments.

Lemma 1. *Let $\mathcal{A}$ be a non-empty set of assignments of preference-attack relations.*

1. *Suppose P is a regular property and every preference-attack relation assignment $\Pi \in \mathcal{A}$ satisfies P. Then $\sqcap\mathcal{A}$ also satisfies P.*
2. *If the preference-attack relations assignments in $\mathcal{A}$ are regular then $\sqcap\mathcal{A}$ is also regular.*

Because of the second item in Lemma 1, we have that $\sqcap\mathcal{A}$ is the infimum of $\mathcal{A}$ w.r.t. $(RPAA_{\mathcal{R}}, \subseteq)$. It follows immediately

Theorem 4. $(RPAA_{\mathcal{R}}, \subseteq)$ *is a complete lower semilattice.*

The following example adapted from example 8 in Dung and Thang (2018) illustrates that the set $RPAA_{\mathcal{R}}$ of regular assignments of preference attack relations could be empty.

Example 6. *Consider the DKB K with empty base of evidence where all strict and defeasible rules of K appearing in Fig. 3 and K contains an unique unconditional preference rule $\pi :\rightarrow d_0 \prec d_1$. We can check that AR_K consists of A and B (Fig. 3), $C = [d_0]$, and a preference argument $X = [\pi]$. $dbut_K = \{(A, B)\}$. By definition, $datt_K = \{((A, B), B)\}$. It is not difficult to see that $patt = \{(X, (A, B))\}$ is the unique preference attack relation satisfying the effective rebut property. Therefore the corresponding PAF has an unique stable extension $E = \{A, B, C, X\}$. However, $cnl(E)$ is inconsistent.*

Canonical preference-attack relation assignment is defined next.

Definition 14. *Suppose the set $RPAA_{\mathcal{R}}$ of all regular attack relation assignments for $\mathcal{R}$ is not empty. The* canonical assignment of preference-attack relations *of $\mathcal{R}$ denoted by $\mathbf{CP}_{\mathcal{R}}$ is defined by: $\mathbf{CP}_{\mathcal{R}} = \sqcap RPAA_{\mathcal{R}}$.*

In other words, we could say that the canonical attack relation assignment is uniquely characterized by the regular properties coupled with the minimal removal principle.

It turns out that even though in general regular attack relation assignments do not exist, their existence is guaranteed under fairly general and sensible conditions. We will discuss these conditions in the next section.

Normal Preference Attack Relation Assignment

Definition 15. *A preference attack relation patt w.r.t. K is said to be* normal *iff for any arguments $A, B, X \in AR_K$, $(X, (A, B)) \in patt$ if and only if there exists $d \in ldr(A)$ such that $(d \prec last(B)) \in cnl(X)$.*

A preference-attack relation assignment Π w.r.t. a rule-based system $\mathcal{R}$ is normal *iff for any knowledge base $K \in \mathcal{C}_{\mathcal{R}}$, $\Pi(K)$ is normal.*

We denote the normal assignment of preference attack relations by Π_{nr}.

It is easy to check that $patt_1$ in Ex. 5 is indeed normal.

Lemma 2. *The normal preference-attack relation assignment Π_{nr} satisfies in general all regular properties except the inconsistency-resolving one.*

Furthermore, the following holds.

Lemma 3. *If the canonical preference attack relation assignment $CP_{\mathcal{R}}$ exists then $CP_{\mathcal{R}} \subseteq \Pi_{nr}$.*

In general, the canonical and normal assignments of preference attack relations are distinct (see Ex. 7 below adapted from Dung and Thang (2018)).

Example 7. *Let $\mathcal{R} = (RS, RD, RP, RT)$ such that $RS \cup RP$ consists of the strict and defeasible rules appearing in Fig. 4 (where a bar on an arrow indicates that the conclusion of the rule is negated) and RP contains an unique preference rule $\pi :\rightarrow d_2 \prec d_3$.*

Let $\overline{\emptyset}$ be an assignment of preference attack relations assigning to each $K \in \mathcal{C}_{\mathcal{R}}$, the empty preference attack relation.

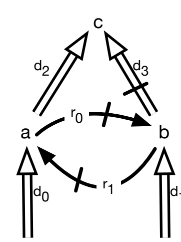

Figure 4: *Rebut-Redundance*

We show that $\overline{\emptyset}$ is also the canonical assignment of preference attack relations of $\mathcal{R}$.

It is easy to check that $\overline{\emptyset}$ satisfies the properties of attack monotonicity, link orientation, and context independence.

Since the set $\{a, b\}$ is inconsistent, there is no knowledge base $K \in \mathcal{C}_{\mathcal{R}}$ such that $\{a, b\} \subseteq BE_K$. Therefore arguments $[[a], d_2]$ and $[[b], d_3]$ never coexist w.r.t. the same knowledge base. Hence the effective rebut property is always satisfied.

We show that $\overline{\emptyset}$ satisfies the inconsistency-resolving property. Let $S \subseteq AR_K$, $K \in \mathcal{C}_{\mathcal{R}}$, be finite such that $cnl(S)$ is inconsistent. Therefore there are two contradictory arguments X, Y generated by S. Since BE_K is consistent, at least one of them (say X) is defeasible. We show that at least one of X, Y is basic defeasible. Suppose none of X, Y are basic defeasible. Therefore X is defeasible but not basic defeasible. Since r_0, r_1 are the only strict rules in $\mathcal{R}$, X must be of one of the following two forms: $X = B_0 = [[d_0], r_0]$ or $X = B_1 = [[d_1], r_1]$. Let $X = B_0 = [[d_0], r_0]$. Hence $cnl(Y) = b$. Because Y is not basic defeasible, and there is no strict rule in $\mathcal{R}$ whose conclusion is b, $Y = [b]$. Therefore

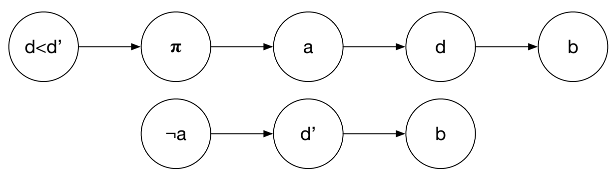

Figure 5: Preference Incoherence

both arguments $[b, r_1]$ and $[d_0]$ are generated w.r.t. S. Similarly, we can prove that if $X = B_1 = [[d_1], r_1]$ then arguments $[a, r_0]$ and $[d_1]$ are generated w.r.t. S. Obviously $[b, r_1]$ rebuts $[d_0]$ (at $[d_0]$) and $[a, r_0]$ rebsut $[d_1]$ (at $[d_1]$). Since there is no preference attack, the inconsistency-resolving property holds obviously. Thus $\overline{\emptyset}$ is regular and hence the canonical assignment of preference attack relations.

Let $K_0 = (\mathcal{R}, \emptyset)$ and $A_0 = [d_0]$, $A_1 = [d_1]$, $A_2 = [[d_0], d_2]$, $A_3 = [[d_1], d_3]$.

It is obvious that $\Pi_{nr}(K_0) = \{P, (A_2, A_3))\}$ where $P = [\pi]$. Thus it is obvious that $\Pi_{nr} \neq \overline{\emptyset}$.

We next discuss two rather intuitive properties of a DKB K, which guarantee that $\Pi_{nr}(K)$ satisfies (**IR**).

The first property addresses the sensibility of preference rules in a rule-based system.

Example 8. *Consider a rule-based system consisting of only two defeasible rules $d : b \Rightarrow a \quad d' : b \Rightarrow \neg a$ together with a preference rule $\pi : \ a \rightarrow d \prec d'$*

The preference rule π intuitively states that "if a holds and both rules d, d' are applicable then pick d'." But for a to hold, d should already be applied and so d' can not be applied. The preference rule is obviously counter-intuitive.

To rule out counter-intuitive rule-based systems like the one above, we introduce below a new notion of general dependency graph between literals and rules.

Definition 16. *The* general dependency graph *of $\mathcal{R}$, denoted by $DG(\mathcal{R}) = (V_{\mathcal{R}}, E_{\mathcal{R}})$, is defined as follows:*

- *Each vertice in $V_{\mathcal{R}}$ is either (i) a rule from $RS \cup RD \cup RP$; or (ii) a preference atom or (iii) a domain literal appearing in some rule in $RS \cup RD \cup RP$.*
- *$(x, y) \in E_{\mathcal{R}}$ iff (i) y is a rule d and x = hd(d); or (ii) x is a rule d and y is an element of $bd(d)$.*

The general dependency graph of the rule-based system in Ex.8 is given in Fig. 5.

The next definition provides a characterization of rule-based system which does not admit the type of counter-intuitive dependence as shown in Ex. 8.

Definition 17. *$\mathcal{R}$ is said to be* preference-stratified *(or just* p-stratified *for short) if there is a ranking function $rank$ from the set of defeasible rules into the set of natural numbers such that the following conditions holds:*

1. *Defeasible rules occurring in the head of the same preference rule have the same rank.*
2. *If there is a path from a vertice ν to a vertice μ in $DG(\mathcal{R})$, where ν is a preference atom containing a defeasible rule d and μ is a defeasible rule d', then $rank(d) > rank(d')$.*

The rule-based system in Ex. 8 is obviously not preference-stratified. The rule based system in Ex. 1 is p-stratified where defeasible rules d_1, d_2 are given rank 1 and the rest is given rank 0. Furthermore, rule-based systems with unconditional preferences are p-stratified.

The second property on $\mathcal{R}$ is on the strict rules of $\mathcal{R}$ adapted from Dung (2016) and Dung and Thang (2018).

Definition 18. *$\mathcal{R}$ is said to satisfy the* self-contradiction property *iff for each minimal inconsistent set of domain literals $S \subseteq \mathcal{L}$, for each $l \in S$, it holds: $S \vdash_{RS} \neg l$.*

The next theorem relates (**IR**) and the self-contradiction property.

Theorem 5. *If $\mathcal{R}$ is p-stratified and satisfies the self-contradictory property then Π_{nr} also satisfies (**IR**) and hence is regular.*

An interesting question arisen immediately is under which conditions the canonical and normal preference attack relation assignments coincide. Such a result has been given for the case of unconditional preferences in Dung and Thang (2018).

As the regular properties for preference attack relation assignments generalize the properties with the same names for unconditional preferences, it is expected that similar result also holds w.r.t. the framework of preference attack relation assignments. We show below that it is indeed the case.

Let $\mathcal{R} = (RS, RD, RP)$ be a consistent unconditional rule-based system. Since $\mathcal{R}$ is consistent, it is obvious that the transitive closure of RP is a strict partial order. For simplicity, we assume that RP coincides with its transitive closure and hence is a strict partial order.

We say a domain literal λ *directly depends* on a domain literal β iff there is a rule $r \in RS \cup RD$ such that $\lambda = hd(r)$ and $\beta \in bd(r)$.

λ *depends* on β iff $\lambda = \beta$ or λ depends on α that directly depends on β.

The set of all sentences in $\mathcal{L}$ on which λ depends is denoted by $\Delta(\lambda)$. For a set $S \subseteq \mathcal{L}$, $\Delta(S)$ is the union of $\Delta(\lambda)$ for $\lambda \in S$.

Theorem 6. *Let $\mathcal{R}$ be an unconditional rule-based system satisfying the self-contradiction property. Furthermore for each defeasible rule $d \in RD$, if there exists $d' \prec d$ in RP then the set $\Delta(bd(d)) \cup \Delta(\neg hd(d))$ is consistent. Then the canonical attack relation assignment $\mathbf{CP}_{\mathcal{R}}$ and the normal attack relation assignment Π_{nr} coincide.*

Discussion and Conclusion

We study defeasible knowledge bases with conditional preferences between defaults (DKB). We introduce the notion of a preference-based argumentation frameworks (PAF), which views direct rebut and undercut attacks among arguments as arguments and contains a preference-attack relation between preference arguments and direct rebuts and undercuts.

We show that PAF generalizes the semantics of basic knowledge bases to DKBs.

We propose the notion of a preference attack relation assignment, which maps each knowledge base to a preference attach relation, and discuss five regular properties, of such an assignment: the context-independence, effective rebuts, inconsistency-resolving, attack monotonicity, and link-orientation properties.

We prove that the set of all regular preference attack relation assignments forms a complete lower-semilattice under the $\subseteq$ relation.

We also define the notion of a normal preference attack relation assignment that in general satisfies all but the inconsistency-resolving property and specify conditions under which normal preference attack relation assignments coincide with regular assignments.

Prakken and Sartor (1997) has studied conditional preferences of defeasible rules. As pointed out in Caminada and Amgoud (2007), the proposed system does not satisfy the consistency postulate. Further, conditional preferences are not used to define attacks against attacks. Antoniou (2004) has studied the use of conditional preferences in the context of defeasible logics.

Delgrande, Schaub, and Tompits (2003) viewed preferences as specifying application orders of rules. In Dung (2016), we show for unconditional preferences that this view is more conservative than our approach in the sense that extensions following DST concept are also extensions in our approach but not vice versa. It turns out that this result also holds for the general case of conditional preferences.

Our approach to dealing with conditional preferences generalizes the approach in Dung and Thang (2018) where the regular properties of preference attack relation assignments generalize the properties with the same names for attack relation assignments for unconditional preferences.

References

Amgoud, L., and Cayrol, C. 2002. Infering from inconsistency in preference-based argumentation framework. *Int J. Automated Reasoning* 29(2):197–215.

Amgoud, L. 2014. Postulates for logic-based argumentation systems. *Int J. Approximate Reasoning* 55(9):2028–2048.

Antoniou, G. 2004. Defeasible logic with dynamic priorities. *International Journal of Intelligent Systems* 19.

Baroni, P.; Cerutti, F.; Giacomin, M.; and Guida, G. 2011. AFRA: argumentation framework with recursive attacks. *Int. J. Approx. Reasoning* 52(1):19–37.

Brewka, G., and Eiter, T. 1999. Preferred answer sets for extended logic programs. *Artificial Intelligence* 109:297–356.

Brewka, G. 1989. Preferred subtheories: An extended logical framework for default reasoning. In *Proc of IJCAI'89*, 1043–1048. Morgan Kaufmann.

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

Davey, B. A., and Priestley, H. A. 2002. *Introduction to Lattices and Order*. Cambridge University Press.

Delgrande, J.; Schaub, T.; and Tompits, H. 2003. A framework for compiling preferences in logic programs. *Theory and Practice of Logic Programming* 129–187.

Dung, P. M., and Thang, P. M. 2018. Fundamental properties of attack relations in structured argumentation with priorities. *Artificial Intelligence* 255:1–42.

Dung, P. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.* 77(2):321–358.

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

Gabbay, D. M. 2009. Semantics for higher level attacks in extended argumentation frames part 1: Overview. *Studia Logica* 93(2-3):357–381.

Garcia, A., and Simari, G. 2004. Defeasible logic programming: An argumentative approach. *TPLP* 4(1-2):95–138.

Hanh, D.; Dung, P.; Hung, N.; and Thang, P. M. 2011. Inductive defence for skeptixcal semantics of extended argumentation. *J of Logic and Computation* 21(2):307–349.

Martinez, D. C.; Garcia, A. J.; and Simari, G. R. 2006. On acceptability in abstract argumentation frameworks with an extended defeat relation. In P.E. Dunne, T. J. M. B.-C., ed., *Proc. of Int. Conf. on Computational models of arguments*. IOS Press.

Modgil, S., and Prakken, H. 2013. A general account of argumentation with preferences. *Artificial Intelligence* 197:361–397.

Modgil, S., and Prakken, H. 2014. The aspic+ framework for structured argumenttion: a tutorial. *J. Arguments and Computation* 5:31–62.

Modgil, S. 2009. Reasoning about preferences in argumentation frameworks. *Artif. Intell.* 173(9-10):901–934.

Prakken, H., and Sartor, G. 1997. Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics (1997)* 7(1):25–75.

Prakken, H. 2010. An abstract framework for argumentation with structured arguments. *J. Arguments and Computation* 1.

Schaub, T., and Wang, K. 2001. A comparative study of logic programs with preferences. In *Proc of IJCAI, 2001*. Morgan kaufmann.

Young, A. P.; Kokciyan, N.; Sassoon, I. K.; Modgil, S.; and Parsons, S. 2018. Instantiating metalevel argumentation frameworks. In *7th International Conference on Computational Models of Argument, Warsaw*.