A characterization of collective conflict for defeasible argumentation

Teresa ALSINET a,1, Ramón BÉJAR a and Lluís GODÓ b

a Department of Computer Science, University of Lleida, SPAIN
b Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, SPAIN

Abstract.
In this paper we define a recursive semantics for warrant in a general defeasible argumentation framework by formalizing a notion of collective (non-binary) conflict among arguments. This allows us to ensure direct and indirect consistency (in the sense of Caminada and Amgoud) without distinguishing between direct and indirect conflicts. Then, the general defeasible argumentation framework is extended by allowing to attach levels of preference to defeasible knowledge items and by providing a level-wise definition of warranted and blocked conclusions. Finally, we formalize the warrant recursive semantics for the particular framework of Possibilistic Defeasible Logic Programming, characterize the unique output program property and design an efficient algorithm for computing warranted conclusions in polynomial space.

Keywords. Defeasible argumentation, collective conflict, recursive warrant semantics.

1. Introduction and motivation
Possibilistic Defeasible Logic Programming (P-DeLP) [3] is a rule-based argumentation framework which is an extension of Defeasible Logic Programming (DeLP) [10] in which defeasible rules are attached with weights (belonging to the real unit interval [0, 1]) expressing their belief or preference strength and formalized as necessity degrees. As many other argumentation frameworks [8,14], P-DeLP can be used as a vehicle for facilitating rationally justifiable decision making when handling incomplete and potentially inconsistent information. Actually, given a P-DeLP program, justifiable decisions correspond to warranted conclusions (with a maximum necessity degree), that is, those which remain undefeated after an exhaustive dialectical analysis of all possible arguments for and against.

In [6] Caminada and Amgoud propose three rationality postulates which every rule-based argumentation system should satisfy. One of such postulates (called Indirect Consistency) claims that the closure of warranted conclusions with respect to the set of strict rules must be consistent. A number of rule-based argumentation systems are identified in which such postulate does not hold (including DeLP [10] and Prakken & Sartor’s [13], among others). As a way to solve this problem, the use of transposed rules is proposed in [6] to extend the representation of strict rules.

Since the dialectical analysis based semantics of P-DeLP does not satisfy indirect consistency, in [2,1] a level-wise approach to compute warranted conclusions, called

---

1Correspondence to: T. Alsinet. Department of Computer Science, University of Lleida. C/Jaume II, 69. Lleida, Spain. Tel.: +34 973702734; Fax: +34 973702702; E-mail: tracy@diei.udl.cat
level-based P-DeLP, was defined ensuring the indirect consistency postulate without extending the representation of strict rules with transposed rules. In contrast with DeLP and other argument-based approaches [8,14,5,15], the level-based P-DeLP framework does not require the use of dialectical trees as underlying structures for characterizing the semantics for warranted conclusions. The level-based P-DeLP framework distinguishes two types of conflicts between arguments, direct and indirect. Direct conflicts occur when there exists an inconsistency emerging from arguments supporting contradictory literals. Indirect conflicts occur in a given program when there exists an inconsistency emerging from the set of strict rules of the program and a set of non-defeated (due to a direct conflict) arguments. The level-based P-DeLP framework therefore establishes an implicit evaluation order between conflicts, in the sense that if a conclusion is involved in both a direct and indirect conflict, the direct conflict invalidates the indirect one. On the other hand, although the level-based P-DeLP semantics for warranted conclusions is skeptical, in [1] it was shown that some circular definitions of conflict between arguments can arise and they can lead to different extensions of warranted conclusions.

Recently Pollock defined [12] a recursive semantics for defeasible argumentation (without levels of preference) where circular definitions of defeat between arguments were characterized by means of inference-graphs, representing (binary) support and defeat relations between the conclusions of arguments. Following this approach, our aim in this paper is to formally characterize circular definitions of conflict among arguments that cause different extensions of warranted conclusions in the level-based P-DeLP framework. However, because of the above mentioned implicit evaluation order between conflicts and its undesired side-effect, we are in need for a new and general notion of conflict among arguments which, besides of ensuring the Caminda and Amgoud’s rationality postulates, allows us to safely reason about circular definitions of conflict between arguments.

To this end, in this paper we first define a recursive semantics for warranted conclusions in a quite general framework (without levels of strength) by formalizing a new collective (non-binary) notion of conflict between arguments ensuring indirect consistency without distinguishing between direct and indirect conflicts. Second we extend the recursive semantics to an argumentation framework with levels of preference by providing a level-wise definition of warranted and blocked conclusions. A warranted conclusion is a justified conclusion which is only based on warranted information and which does not generate a conflict, while a blocked conclusion is a conclusion which, like warranted conclusions, is only based on warranted information, but it does generate a conflict. Third, we specialize the warrant recursive semantics for the particular framework of P-DeLP, we refer to this formalism as RP-DeLP, characterize the condition under which a program has a unique output based on what we call warrant dependency graph, and design an efficient algorithm for computing warranted conclusions in polynomial space.

2. General defeasible argumentation framework

We will start by considering a rather general framework for defeasible argumentation based on a propositional logic \((\mathcal{L}, \vdash)\) with a special symbol \(\bot\) for contradiction\(^2\). For any set of formulas \(A\), if \(A \vdash \bot\) we will say that \(A\) is contradictory, while if \(A \not\vdash \bot\) we

\(^2\)If not stated otherwise, in this and in the next section \((\mathcal{L}, \vdash)\) may be taken as classical propositional logic.
will say that \( A \) is consistent. A knowledge base (KB) is a triplet \( \mathcal{P} = (\Pi, \Delta, \Sigma) \), where \( \Pi, \Delta, \Sigma \subseteq \mathcal{L} \), and \( \Pi \not\vdash \bot \). \( \Pi \) is a finite set of formulas representing strict knowledge (formulas we take for granted they hold to be true), \( \Delta \) is another finite set of formulas representing the defeasible knowledge (formulas for which we have reasons to believe they are true) and \( \Sigma \) denotes the set of formulas over which arguments can be built. In many argumentation systems, \( \Sigma \) is taken to be a set of literals.

The notion of argument is the usual one. Given a KB \( \mathcal{P} \), an argument for a formula \( \varphi \in \Sigma \) is a pair \( A = \langle A, \varphi \rangle \), with \( A \subseteq \Delta \) such that:

1. \( \Pi \cup A \not\vdash \bot \), and
2. \( A \) is minimal (w.r.t. set inclusion) such that \( \Pi \cup A \vdash \varphi \).

If \( A = \emptyset \), then we will call \( A \) a s-argument (s for strict), otherwise it will be a d-argument (d for defeasible). The notion of subargument is referred to d-arguments and expresses an incremental prove relationship between arguments which is formalized as follows.

**Definition 1 (Subargument)** Let \( \langle B, \psi \rangle \) and \( \langle A, \varphi \rangle \) be two d-arguments such that the minimal sets (w.r.t. set inclusion) \( \Pi_\psi \subseteq \Pi \) and \( \Pi_\varphi \subseteq \Pi \) such that \( \Pi_\psi \cup B \vdash \psi \) and \( \Pi_\varphi \cup A \vdash \varphi \) verify that \( \Pi_\psi \subseteq \Pi_\varphi \). Then, \( \langle B, \psi \rangle \) is a subargument of \( \langle A, \varphi \rangle \), written \( \langle B, \psi \rangle \sqsubseteq \langle A, \varphi \rangle \), when either \( B \subseteq A \) (strict inclusion for defeasible knowledge), or \( B = A \) and \( \Pi_\psi \subseteq \Pi_\varphi \) (strict inclusion for strict knowledge), or \( B = A \) and \( \psi \vdash \varphi \) and \( \varphi \not\vdash \psi \).

A formula \( \varphi \in \Sigma \) will be called justifiable w.r.t. \( \mathcal{P} \) if there exists an argument for \( \varphi \), i.e. there exists \( A \subseteq \Delta \) such that \( \langle A, \varphi \rangle \) is an argument.

The usual notion of attack or defeat relation in an argumentation system is binary. However in certain situations, the conflict relation among arguments is hardly representable as a binary relation. For instance, consider the following KB \( \mathcal{P}_1 = (\Pi, \Delta, \Sigma) \) with
\[
\Pi = \{ a \land b \rightarrow \neg p \}, \quad \Delta = \{ a, b, p \} \quad \text{and} \quad \Sigma = \{ a, b, p, \neg p \}.
\]

Clearly, \( A_1 = \{ p \} \), \( A_2 = \{ b \} \), \( A_3 = \{ a \} \) are arguments that justify \( p \), \( b \) and \( a \) respectively, and which do not pair-wisely generate a conflict. Indeed, \( \Pi \cup \{ a, b \} \not\vdash \bot \), \( \Pi \cup \{ a \} \not\vdash \bot \) and \( \Pi \cup \{ b, p \} \not\vdash \bot \). However the three arguments are collectively conflicting since \( \Pi \cup \{ a, b, p \} \not\vdash \bot \), hence in this \( \mathcal{P}_1 \) there is a non-binary conflict relation among several arguments. In the following we will formalize this notion of collective, or non-binary, conflict among in principle valid arguments and which arises when we compare them with the strict part of the knowledge base.

The following notion of acceptable argument with respect to a set (possibly empty) of justifiable conclusions \( W \) will play a key role. If we think of \( W \) as a consistent set of already warranted conclusions, an acceptable argument captures the idea of an argument which is based on subarguments already warranted.

**Definition 2 (Acceptable argument)** Let \( W \) be a set of justifiable conclusions which is consistent w.r.t. \( \Pi \), i.e. \( \Pi \cup W \not\vdash \bot \). A d-argument \( A = \langle A, \varphi \rangle \) is an acceptable argument for \( \varphi \) w.r.t. \( W \) iff:

\[\text{Notice that if } (\Pi, \Delta, \Sigma) = (\{ r \}, \{ r \rightarrow p \land q, \{ p, q, p \land q \}) \text{ and } A = \{ r \rightarrow p \land q \} \text{ then } A_1 = \{ A, p \}, A_2 = \{ A, q \} \text{ and } A_3 = \{ A, p \land q \} \text{ are arguments for different formulas with a same support and thus, in our framework, } A_3 \sqsubseteq A_1 \text{ and } A_3 \sqsubseteq A_2 \text{ are the subargument relations between arguments } A_1, A_2 \text{ and } A_3 \text{ since } p \land q \not\vdash p, p \land q \not\vdash q, p \not\vdash p \land q \text{ and } q \not\vdash p \land q.\]
1. if $\langle B, \psi \rangle$ is a subargument of $\langle A, \varphi \rangle$ then $\psi \in W$
2. $\Pi \cup W \cup \varphi \not\models \bot$

In the above example, arguments $A_1$, $A_2$ and $A_3$ are acceptable w.r.t. $\Pi$ and the empty set of conclusions $W = \emptyset$. However $A_4 = \langle \{a, b\}, \neg p \rangle$ is an argument for $\neg p$, but $A_3$ is not acceptable w.r.t. $W = \emptyset$ since $A_2$ and $A_3$ are subarguments of $A_4$ but obviously $a, b \not\in W$.

Now we are ready to introduce the notion of collective conflict relative to a consistent set of justifiable conclusions. The idea of defining a warrant semantics on the basis of conflicting sets of arguments was proposed in [16] and [11]. The difference between these approaches and our notion of collective conflict is that in [16] the notion of conflict is not relative to a set of already warranted conclusions and [11] defines a generalization of Dung’s abstract framework with sets of attacking arguments not relative to the strict part of the knowledge base.

**Definition 3 (Conflict among arguments)** Let $P = (\Pi, \Delta, \Sigma)$ be a KB, let $W$ be a consistent set of justifiable conclusions w.r.t. $\Pi$ and let $A_1 = \langle A_1, \varphi_1 \rangle, \ldots, A_k = \langle A_k, \varphi_k \rangle$ be acceptable arguments w.r.t. $W$. We say that the set of arguments $\{A_1, \ldots, A_k\}$ generates a conflict w.r.t. $W$ iff the two following conditions hold:

- (C) The set of argument conclusions $\{\varphi_1, \ldots, \varphi_k\}$ is contradictory w.r.t. $\Pi \cup W$, i.e. $\Pi \cup W \cup \{\varphi_1, \ldots, \varphi_k\} \not\models \bot$.
- (M) The set $\{A_1, \ldots, A_k\}$ is minimal w.r.t. set inclusion satisfying (C), i.e. if $S \subset \{\varphi_1, \ldots, \varphi_k\}$, then $\Pi \cup W \cup S \not\models \bot$.

Consider the previous KB $P_1$. According to Definition 3, it is clear that the set of acceptable arguments $\{A_1, A_2, A_3\}$ for $p, b$ and $a$ respectively generates a (collective) conflict w.r.t. $W = \emptyset$. The intuition is that this collective conflict should block the conclusions $a, b$ and $p$ to be warranted. Now, this general notion of conflict is used to define a recursive semantics for warranted conclusions of a knowledge base. Actually we define below an output of a KB $P = (\Pi, \Delta, \Sigma)$ as a pair $(\text{Warr}, \text{Block})$ of subsets of $\Sigma$ of warranted and blocked conclusions respectively, all of them based on warranted information but, while warranted conclusions do not generate any conflict, blocked conclusions do not.

**Definition 4 (Output for a KB)** An output for a KB $P = (\Pi, \Delta, \Sigma)$ is any pair $(\text{Warr}, \text{Block})$, where $\text{Warr} = s-\text{Warr} \cup d-\text{Warr}$ with $s-\text{Warr} = \{\varphi \mid \Pi \not\models \varphi \} \cap \Sigma$, and $d-\text{Warr}$ and $\text{Block}$ are required to satisfy the following recursive constraints:

1. A d-argument $\langle A, \varphi \rangle$ is called valid (or not rejected) if it satisfies the following two conditions:
   (i) for every $\langle B, \psi \rangle \sqsupseteq \langle A, \varphi \rangle, \psi \in d-\text{Warr}$,
   (ii) $\langle A, \varphi \rangle$ is acceptable w.r.t. the set $W = \{\psi \mid \langle B, \psi \rangle \sqsupseteq \langle A, \varphi \rangle\}$.

2. For every valid argument $\langle A, \varphi \rangle$ we have that
   - $\varphi \in d-\text{Warr}$ whenever there does not exist a set of valid arguments $G$ such that
     (i) $\langle A, \varphi \rangle \not\sqsubseteq \langle C, \chi \rangle$ for all $\langle C, \chi \rangle \in G$
     (ii) $G \cup \{\langle A, \varphi \rangle\}$ generates a conflict w.r.t. $W = \{\psi \mid \text{there exists} \langle B, \psi \rangle \sqsupseteq \langle D, \gamma \rangle \text{ for some} \langle D, \gamma \rangle \in G \cup \{\langle A, \varphi \rangle\}\}$. 

The intuition underlying this definition is as follows: an argument \( \langle A, \varphi \rangle \) is either warranted or blocked whenever for each subargument \( \langle B, \psi \rangle \) of \( \langle A, \varphi \rangle \), \( \psi \) is warranted; then, it is eventually warranted if \( \varphi \) is not involved in any conflict, otherwise it is blocked.

**Example 5** Consider the KB \( \mathcal{P}_2 = (\Pi, \Delta, \Sigma) \), with
\[
\Pi = \{ a \rightarrow y, b \land c \rightarrow \neg y \}, \quad \Delta = \{ a, b, c, \neg c \} \quad \text{and} \quad \Sigma = \{ a, b, c, \neg c, y, \neg y \}.
\]
According to Definition 4, \( s-\text{Warr} = \emptyset \) and the arguments \( \{\{a\}, \}, \{\{b\}, \}, \{\{c\}, \} \) and \( \{\neg c, \neg c\} \) are valid. Now, for every such valid argument there exists a set of valid arguments which generates a conflict w.r.t. \( W = \emptyset \); indeed both sets of valid arguments \( \{\{a\}, \}, \{\{b\}, \}, \{\{c\}, \} \) and \( \{\{c\}, \{\neg c\}, \{\neg c\}\} \) generate a conflict (since \( \Pi \cup \{a, b, c\} \vdash \bot \) and \( \Pi \cup \{c, \neg c\} \vdash \bot \)). Therefore, \( a, b, c \) and \( \neg c \) are blocked conclusions. On the other hand, the arguments \( \{a, b, \neg c\}, \{a, y\} \) and \( \{b, c, \neg y\} \) are not valid since they are based on conclusions which are not warranted. Hence \( y \) and \( \neg y \) are considered as rejected conclusions. Thus, the (unique) output for \( \mathcal{P} \) is the pair \( (\text{Warr}, \text{Block}) = (\emptyset, \Delta) \). Intuitively this output for \( \mathcal{P} \) expresses that all conclusions in Block are valid, however all together are contradictory w.r.t. \( \Pi \).

The KB \( \mathcal{P}_2 \) was also considered in [2], where direct conflicts were evaluated before indirect conflicts, and thus, every blocked literal invalidated all rules in which that literal occurred. Hence, in [2], \( c \) and \( \neg c \) were considered as blocked conclusions but \( a \) and \( y \) were warranted conclusions.

Next we prove that if \( (\text{Warr}, \text{Block}) \) is an output for a KB the set \( \text{Warr} \) of warranted conclusions satisfies indirect consistency and the closure postulate (in the sense of Caminada and Amgoud) with respect to the strict knowledge.

**Proposition 6 (Indirect consistency)** Let \( \mathcal{P} = (\Pi, \Delta, \Sigma) \) be a KB and let \( (\text{Warr}, \text{Block}) \) be an output for \( \mathcal{P} \). Then, \( \Pi \cup \text{Warr} \nvdash \bot \).

**Proof:** Suppose that \( \Pi \cup \text{Warr} \vdash \bot \). Obviously, it should be that \( \Pi \cup W \vdash \bot \) for some \( W \subseteq d-\text{Warr} \). However, by Definition 4, for every \( \varphi \in d-\text{Warr} \) there does not exist a set \( W' \subseteq d-\text{Warr} \) such that \( \Pi \cup W' \cup \{\varphi\} \vdash \bot \), and therefore, \( \Pi \cup W \nvdash \bot \) for all \( W \subseteq d-\text{Warr} \).

**Proposition 7 (Closure)** Let \( \mathcal{P} = (\Pi, \Delta, \Sigma) \) be a KB and let \( (\text{Warr}, \text{Block}) \) be an output for \( \mathcal{P} \). If \( \Pi \cup \text{Warr} \vdash \varphi \) with \( \varphi \in \Sigma \), then \( \varphi \in \text{Warr} \) whenever there exists an acceptable argument for \( \varphi \) w.r.t. \( \text{Warr} \).

**Proof:** Suppose that for some \( W \subseteq \text{Warr} \), \( \Pi \cup W \vdash \varphi \) and \( \varphi \notin \text{Warr} \), for some formula \( \varphi \in \Sigma \) such that there exists an argument \( \langle A, \varphi \rangle \) satisfying that if \( \langle B, \psi \rangle \) is a subargument of \( \langle A, \varphi \rangle \) then \( \psi \in W \). On the one hand, if \( A = \emptyset \), it must be that \( \Pi \vdash \varphi \). Then, as \( \varphi \in \Sigma \), \( \varphi \in \text{Warr} \). On the other hand, as \( \Pi \cup \text{Warr} \nvdash \bot \) and \( \Pi \cup W \vdash \varphi \), \( \Pi \cup W \cup \{\varphi\} \nvdash \bot \), and thus, \( \langle A, \varphi \rangle \) is a valid argument for \( \varphi \). Then, if \( \varphi \notin \text{Warr} \), there exists a set of valid arguments \( G \) such that (i) \( \langle A, \varphi \rangle \nsubseteq \langle C, \chi \rangle \) for all \( \langle C, \chi \rangle \in G \), and (ii) \( G \cup \{\langle A, \varphi \rangle\} \) generates a conflict w.r.t. \( \Pi \), \( \Pi \cup W' \vdash \langle B, \psi \rangle \nsubseteq \langle D, \gamma \rangle \) for some \( \langle D, \gamma \rangle \in G \cup \{\langle A, \varphi \rangle\} \). According to Definition 3, if \( G \cup \{\langle A, \varphi \rangle\} \) generates a conflict w.r.t. \( W' \),
then $\Pi \cup W' \cup \{\varphi\} \cup \{\psi \mid \langle B, \psi \rangle \in G\} \vdash \bot$ (condition (C)), and $\Pi \cup W' \cup S \not\vdash \bot$, for all set $S \subset \{\varphi\} \cup \{\psi \mid \langle B, \psi \rangle \in G\}$ (condition (M)). Now, as $W \subset W'$ and $\Pi \cup W \not\vdash \varphi$, if $\Pi \cup W' \cup \{\varphi\} \cup \{\psi \mid \langle B, \psi \rangle \in G\} \vdash \bot$, then $\Pi \cup W' \cup \{\psi \mid \langle B, \psi \rangle \in G\} \vdash \bot$, and thus, $\varphi \in Warr$. 

We remark that, as it will be discussed in Section 4, a KB may have multiple outputs. For instance, consider the KB $P_3 = (\Pi, \Delta, \Sigma)$ with $\Pi = \emptyset$, $\Delta = \{p, q, \neg p \vee \neg q\}$ and $\Sigma = \{p, q, \neg p, \neg q\}$. Then, one can check that there are two outputs, $Warr_1 = \{p\}$, $Block_1 = \{q, \neg q\}$, and $Warr_2 = \{q\}$, $Block_2 = \{p, \neg p\}$.

3. Extending the framework with a preference ordering on arguments

In the previous section, we have considered knowledge bases containing formulas describing knowledge at two epistemic levels, strict and defeasible. A natural extension is to introduce several levels of defeasibility or preference among different pieces of defeasible knowledge.

A stratified knowledge base (sKB) is a tuple $P = (\Pi, \Delta, \preceq, \Sigma)$, such that $(\Pi, \Delta, \Sigma)$ is a KB (in the sense of the previous section) and $\preceq$ is a suitable total pre-order on the set of defeasible formulas $\Delta$. Suitable means for us that this pre-order is representable by a necessity measure defined on the set of formulas of $L$, namely $\varphi \preceq \psi$ iff $N(\varphi) \leq N(\psi)$ for each $\varphi, \psi \in \Delta \cup \Pi$, where $N$ is a mapping $N : L \rightarrow [0, 1]$ such that

1. $N(\top) = 1$, $N(\bot) = 0$,
2. $N(\varphi \land \psi) = \min\{N(\varphi), N(\psi)\}$, and further
3. $N(\varphi) = 1$ iff $\Pi \vdash \varphi$

Then we define the strength of an argument $\langle A, \varphi \rangle$, written $s(\langle A, \varphi \rangle)$, as follows:

$$s(\langle A, \varphi \rangle) = 1 \text{ if } A = \emptyset, \text{ and } s(\langle A, \varphi \rangle) = \min\{N(\psi) \mid \psi \in A\}, \text{ otherwise.}$$

Since we are considering several levels of strength among arguments, the intended construction of the sets of conclusions $Warr$ and $Block$ is done level-wise, starting from the highest level and iteratively going down from one level to next level below. If $1 > \alpha_1 > \ldots > \alpha_p \geq 0$ are the strengths of $d$-arguments that can be built within a sKB $P = (\Pi, \Delta, \preceq, \Sigma)$, we define $d-Warr = \{d-Warr(\alpha_1), \ldots, d-Warr(\alpha_p)\}$ and $Block = \{Block(\alpha_1), \ldots, Block(\alpha_p)\}$, where $d-Warr(\alpha_i)$ and $Block(\alpha_i)$ are respectively the sets of warranted and blocked justifiable conclusions with strength $\alpha_i$. Then, we safely write $d-Warr(> \alpha_i)$ to denote $\cup_{\beta > \alpha_i} d-Warr(\beta)$, and analogously for $Block(> \alpha_i)$, defining $d-Warr(> \alpha_1) = s-Warr$ and $Block(> \alpha_1) = \emptyset$.

Definition 8 (Output for a sKB) An output for a sKB $P = (\Pi, \Delta, \preceq, \Sigma)$ is any pair $(Warr, Block)$, where $Warr = s-Warr \cup d-Warr$ with $s-Warr = \{\varphi \mid \Pi \vdash \varphi \cap \Sigma$, and $d-Warr$ and $Block$ are required to satisfy the following recursive constraints:

1. A d-argument $\langle A, \varphi \rangle$ of strength $\alpha_i$ is called valid (or not rejected) if it satisfies the following three conditions:

$^4$Actually, several $N$’s may lead to a same pre-order $\preceq$, but we can take any of them to define the degree of strength since only the relative ordering is what matters.

$^5$Notice that if $\langle A, \varphi \rangle$ is an acceptable argument w.r.t. $d-Warr(> \alpha_i)$, then $\langle A, \varphi \rangle$ is valid whenever condition (ii) holds.
Then, at level $k$, the argument $A, \varphi$ is valid when there does not exist a set $G$ of valid arguments of strength greater than $\varphi$(i) for every subargument $B, \psi \sqsubseteq A, \varphi$ of strength $\alpha$, $\psi \in d$-Warr($\alpha$);
(ii) $(A, \varphi)$ is acceptable w.r.t.
$W = d$-Warr($\varphi$) $\cup \{\psi \mid (B, \psi) \sqsubseteq (A, \varphi) \text{ and } s((B, \psi)) = \alpha\}$;
(iii) $\varphi \not\in d$-Warr($\varphi$) $\cup$ Block($\varphi$) and $\{\varphi, \psi\} \not\vdash \bot$ for all $\psi \in \text{Block}(\varphi)$.

2. For every valid argument $(A, \varphi)$ of strength $\alpha$, we have that
- $(A, \varphi) \in \text{d-Warr}(\alpha)$ whenever there does not exist a set $G$ of valid arguments of strength $\alpha$, such that
  (i) $G \cup \{C, \chi\}$ for all $(C, \chi) \in G$
  (ii) $G \cup \{(A, \varphi)\}$ generates a conflict w.r.t. $W = d$-Warr($\alpha$) $\cup \{\psi \mid$ there exists $(B, \psi) \sqsubseteq (D, \gamma)$ for some $(D, \gamma) \in G \cup \{(A, \varphi)\}\}$
- otherwise, $\varphi \in \text{Block}(\alpha)$.

There are two main remarks when considering several levels of strength among arguments. On the one hand a d-argument $(A, \varphi)$ of strength $\alpha$, is valid whenever there does not exist a different valid argument for $\varphi$ of strength greater than $\alpha$, and $\varphi$ is consistent with each valid argument of strength greater than $\alpha$. On the other hand, a valid argument $(A, \varphi)$ of strength $\alpha$, becomes blocked as soon as it leads to some conflict among arguments of strength $\alpha$.

Example 9 Consider the KB $P_1$ in the previous section
$\mathcal{P} = \{a \land b \rightarrow \neg p\}$, $\Delta = \{a, b, p\}$ and $\Sigma = \{a, b, p, \neg p\}$.
extended with levels of defeasibility as follows: $\{a, b\} \prec p$. Assume $\alpha_1$ is the level of $p$ and $\alpha_2$ the level of $a$ and $b$, obviously with $1 > \alpha_1 > \alpha_2$. According to Definition 8, $s$-Warr = $\emptyset$ and the argument for $(\{p\}, p)$ is the only valid argument with strength $\alpha_1$. Then, at level $\alpha_1$, we get $d$-Warr($\alpha_1$) $\cup \{\psi \mid (A, \psi) \sqsubseteq (A, \varphi)\}$ and $\text{Block}(\alpha_1) = \emptyset$. At level $\alpha_2$, we have that $\text{d-Warr}(\alpha_2) = \emptyset$ and $\text{Block}(\alpha_2) = \{a, b\}$. Notice that the argument $(\{a, b\}, \neg p)$ is not a valid argument since it is based on $a$ and $b$ and $a, b \not\in d$-Warr($\alpha_2$).

Example 10 Consider the KB $P_2$ of Example 5:
$\mathcal{P} = \{a \rightarrow y, b \land c \rightarrow \neg y\}$, $\Delta = \{a, b, c, \neg c\}$, and $\Sigma = \{a, b, c, \neg c, y, \neg y\}$.
extended with three levels of defeasibility as follows: $\neg c \prec c \prec \{a, b\}$. Assume $\alpha_1$ is the level of $a$ and $b$, $\alpha_2$ is the level of $c$, and $\alpha_3$ is the level of $\neg c$, with $1 > \alpha_1 > \alpha_2 > \alpha_3$. Then, $s$-Warr = $\emptyset$ and, at level $\alpha_1$, we have not only the conclusions $a$, $b$ and $y$ with valid arguments not generating conflict, but also $(\{a, b\}, \neg c)$ is a valid argument for $\neg c$ which does not generate conflict. Therefore, $d$-Warr($\alpha_1$) $\cup \{a, b, y, \neg c\}$ and $\text{Block}(\alpha_1) = \emptyset$. At level $\alpha_2$, we have arguments for $c$ and $\neg y$. Since $\text{d-Warr}(\alpha_1) \cup \{c\} \vdash \bot$, the argument $(\{c\}, c)$ is not acceptable w.r.t. $d$-Warr($\alpha_1$), and thus, $c$ is a rejected conclusion. Then, as the argument $(\{b, c\}, \neg y)$ for $\neg y$ is based on $c$, $\neg y$ is also a rejected conclusion, and therefore $d$-Warr($\alpha_2$) $\cup \{\neg y\} = \emptyset$. Finally, at level $\alpha_3$ we have the argument $(\{\neg c\}, \neg c)$, but since $\neg c$ is already in $d$-Warr($\alpha_1$), we also have $d$-Warr($\alpha_3$) $\cup \{\neg c\} = \emptyset$. 


4. A particular case: recursive P-DeLP

In this section we particularize the framework and recursive warrant semantics for stratified knowledge bases defined in the previous section to the case of the P-DeLP programs. As mentioned in Section 1, P-DeLP is a rule-based argumentation system extending the well-known DeLP system in which weights are attached to defeasible rules expressing their belief or preference strength and formalized as necessity degrees. For a detailed description of the P-DeLP argumentation system based on dialectical trees the reader is referred to [3].

Although the original syntax and inference of P-DeLP are a bit different (e.g. the weights are explicit in the formulas and arguments), here we will present them in a way so to adapt them to the framework introduced in the previous sections. We will refer to this particular framework as RP-DeLP. Hence we define the logic \( (\mathcal{L}_R, \vdash_R) \) underlying RP-DeLP as follows. The language of RP-DeLP is inherited from the language of logic programming, including the notions of atom, literal, rule and fact. Formulas are built over a finite set of propositional variables \( p, q, ... \) which is extended with a new (negated) atom \( \sim p \) for each original atom \( p \). Atoms of the form \( \sim p \) will be referred as literals, and if \( P \) is a literal, we will use \( \sim P \) to denote \( \sim p \) if \( P \) is an atom \( p \), and will denote \( p \) if \( P \) is a negated atom \( \sim p \). Formulas of \( \mathcal{L}_R \) consist of rules of the form \( Q \leftarrow P_1 \land \ldots \land P_k \), where \( Q, P_1, \ldots, P_k \) are literals. A fact will be a rule with no premises. We will also use the name clause to denote a rule or a fact. The inference operator \( \vdash_R \) is defined by instances of the modus ponens rule of the form: \( \{ Q \leftarrow P_1 \land \ldots \land P_k, P_1, \ldots, P_k \} \vdash_R Q \). A set of clauses \( \Gamma \) is contradictory, denoted \( \Gamma \vdash \bot \), if, for some atom \( q \), \( \Gamma \vdash_R q \) and \( \Gamma \vdash_R \sim q \).

A RP-DeLP program \( \mathcal{P} \) is just a stratified knowledge base \( (\Pi, \Delta, \preceq, \Sigma) \) over the logic \( (\mathcal{L}_R, \vdash_R) \), where \( \Sigma \) consists of the set of all literals of \( \mathcal{L}_R \). We will assume that \( \preceq \) is representable by a necessity measure \( N \), so we will often refer to numerical weights for defeasible clauses and arguments rather than to the pre-ordering \( \preceq \). Also, for the sake of a simpler notation we will get rid of \( \Sigma \) from a program specification.

As we have mentioned in the previous section, in some cases the output \( (\text{Warr}, \text{Block}) \) for a stratified knowledge base in general, and for a RP-DeLP program in particular, is not unique, due to circular definitions of warranty that emerge when considering conflicts among arguments. Such circular definitions of warranty are characterized next by means of what we call warrant dependency graph of a RP-DeLP program. In [12] a similar graph structure, called inference-graph, was defined to represent inference (support) and defeat relations among arguments allowing to detect circular defeat relations when considering recursive semantics for defeasible reasoning. The main difference between both approaches is that in our case we handle collective conflicts among arguments in order to preserve direct consistency among warranted conclusions and indirect consistency with respect to the strict knowledge.

In the following, given a RP-DeLP program \( \mathcal{P} = (\Pi, \Delta) \) with preference levels \( 1 > \alpha_1 > \ldots > \alpha_m > 0 \), if \( W \) denotes a set of justifiable literals, we will denote by \( W(\alpha) \) the subset of literals \( Q \) from \( W \) for which there exist an argument \( \langle A, Q \rangle \) with maximum strength \( \alpha \), \( W(\geq \alpha) = \bigcup_{\beta \geq \alpha} W(\beta) \), and \( W(> \alpha) = \bigcup_{\beta > \alpha} W(\beta) \) with \( W(> \alpha_1) = W(1) \).

**Definition 11 (Warrant dependency graph)** Let \( \mathcal{P} = (\Pi, \Delta) \) be a RP-DeLP program and let \( W \) be a set of justifiable conclusions consistent with \( \Pi \). Let \( A_1 = \ldots \)
\{A_1, Q_1\}, \ldots, A_k = \{A_k, Q_k\} be acceptable arguments of a same strength \(\alpha\) w.r.t. \(W\) such that for all \(i, Q_i \notin W(\geq \alpha)\). Moreover, let \(B_1 = \{B_1, P_1\}, \ldots, B_n = \{B_n, P_n\}\) be arguments of the same strength \(\alpha\) such that for all \(j, P_j \notin W(\geq \alpha), P_j \notin \{Q_1, \ldots, Q_k\}\), and there exists an argument \(S \in \{A_1, \ldots, A_k\}\) with \(S \subseteq B_j\). Then, the warrant dependency graph \((V, E)\) for \(\{A_1, \ldots, A_k\}\) w.r.t. \(W\) and \(\{B_1, \ldots, B_n\}\) is defined as follows:

1. For every literal \(L \in \{Q_1, \ldots, Q_k\} \cup \{P_1, \ldots, P_n\}\), the set of vertices \(V\) includes one vertex \(v_L\).
2. For every pair of literals \(\{L_1, L_2\}\) such that \(L_1 \not\sim L_2\) with \(L_1 \in \{P_1, \ldots, P_n\}\) and \(L_2 \in \{Q_1, \ldots, Q_k\}\), the set of directed edges \(E\) includes one edge \((v_{L_1}, v_{L_2})^6\).
3. For every pair of literals \(\{L_1, L_2\}\) such that \(L_1 \in \{Q_1, \ldots, Q_k\}\), \(L_2 \in \{P_1, \ldots, P_n\}\) and the argument of \(L_1\) is a subargument of the argument of \(L_2\), the set of directed edges \(E\) includes one edge \((v_{L_1}, v_{L_2})^7\).
4. For every strict rule \(L \leftarrow L_1 \land \ldots \land L_m\) of \(\Pi\) such that
   - either \(\sim L \in W(\geq \alpha)\) or \(\sim L \in \{Q_1, \ldots, Q_k\}\), and
   - for every \(L_i\) \((i = 1, \ldots, m)\), either \(L_i \in W(\geq \alpha)\) or \(L_i \in \{Q_1, \ldots, Q_k\} \cup \{P_1, \ldots, P_n\}\),
   the set of directed edges \(E\) includes one edge \((v_{L_1}, v_{L_2})^8\) for every pair of literals \(\{L_i, L_j\} \subseteq \{L_1, \ldots, L_m\}\) with \(L_i \in \{P_1, \ldots, P_n\}\) and \(L_j \in \{Q_1, \ldots, Q_k\}\), whenever the argument of \(L_j\) is not a subargument of the argument of \(L_i\).
5. Elements of \(V\) and \(E\) are only obtained by applying the above construction rules.

Intuitively, the warrant dependency graph for a set of arguments represents conflict and support dependences among arguments in \(\{A_1, \ldots, A_k\}\) and arguments in \(\{B_1, \ldots, B_n\}\) w.r.t. a set of justified conclusions \(W\).

**Example 12** Consider a RP-DeLP program defined from the KB \(\mathcal{P}_3\) of Section 2; i.e. a RP-DeLP program with an empty set of strict clauses and the following set of defeasible clauses with just one defeasibility level:
\[
\Delta = \{p, q, \sim p \leftarrow q, \sim q \leftarrow p\}.
\]
Now, consider the empty set of conclusions \(W = W(1) = \emptyset\) and arguments for conclusions \(p\) and \(q\); i.e. \(A_1 = \{\{p\}, p\}\) and \(A_2 = \{\{q\}, q\}\). Finally, consider the arguments for conclusions \(\sim p\) and \(\sim q\); i.e. \(B_1 = \{\{q, \sim p \leftarrow q\}, \sim p\}\) and \(B_2 = \{\{p, \sim q \leftarrow p\}, \sim q\}\).

Figure 1 (a) shows the warrant dependency graph for \(A_1\) and \(A_2\) w.r.t. \(W = \emptyset, B_1,\) and \(B_2\). Conflict and support dependences between literals are represented as dashed and solid arrows, respectively. The cycle of the graph expresses that (1) the warranty of \(p\) depends on a (possible) conflict with \(\sim p\); (2) the support of \(\sim p\) depends on \(p\) (i.e. the validity of \(\sim p\) depends on the warranty of \(p\)); (3) the warranty of \(q\) depends on a (possible) conflict with \(\sim q\); and (4) the support of \(\sim q\) depends on \(p\) (i.e. the validity of \(\sim q\) depends on the warranty of \(p\)).

Consider now the RP-DeLP program \(\mathcal{P}_4 = (\Pi, \Delta)\), with
\[
\Pi = \{y, \sim y \leftarrow p \land r, \sim y \leftarrow q \land s\}\] and \(\Delta = \{p, q, r \leftarrow q, s \leftarrow p\}\)
with just one defeasibility level. Moreover consider the set of justified conclusions \(W =\)

---

6The directed edge \((v_{L_1}, v_{L_2})^6\) represents a conflict dependence of \(L_2\) w.r.t. \(L_1\).
7The directed edge \((v_{L_1}, v_{L_2})^7\) represents a support dependence of \(L_2\) w.r.t. \(L_1\).
8The directed edge \((v_{L_1}, v_{L_2})^8\) represents a conflict dependence of \(L_j\) w.r.t. \(L_i\).
\[ W(1) = \{y\} \text{ and arguments for conclusions } p \text{ and } q; \text{ i.e.} \]

\[ A_1 = \{p\}, p \text{ and } A_2 = \{q\}, q. \]

Finally, consider arguments for conclusions \( r \) and \( s \); i.e.

\[ B_1 = \{q, r \leftarrow q\}, r \} \text{ and } B_2 = \{p, s \leftarrow p\}, s. \]

Figure 1 (b) shows the warrant dependency graph for \( A_1 \) and \( A_2 \) w.r.t. \( W, B_1, \) and \( B_2. \) The cycle of the graph expresses that (1) the warranty of \( p \) depends on a (possible) conflict with \( r; \) (2) the support of \( r \) depends on \( q \) (i.e. the validity of \( r \) depends on the warranty of \( q; \) (3) the warranty of \( q \) depends on a (possible) conflict with \( s; \) and (4) the support of \( s \) depends on \( p \) (i.e. the validity of \( s \) depends on the warranty of \( p \)).

![Figure 1](image)

The characterization of the unique output property for a program \( P = (\Pi, \Delta) \) is done level-wise, starting from the highest level and iteratively going down from one level to next level below. For every level it consists in checking whether for some literal \( L, \) the warranty of \( L \) recursively depends on itself based on the topology of a warrant dependency graph defined as follows.

**Definition 13 (Graph for a literal)** Let \( P = (\Pi, \Delta) \) be a RP-DeLP program, let \( (Warr, \text{Block}) \) be an output for \( P \) and let \( L \) be a literal such that \( L \in \text{Warr}(\alpha), \) for some level \( \alpha. \) The graph for \( L \) w.r.t. \( \text{Warr} \) is the warrant dependency graph \( (V, E) \) for arguments \( \{A_1, \ldots, A_k\} \) w.r.t. \( W \) and \( \{B_1, \ldots, B_n\} \) where

- \( W = \text{Warr}(\geq \alpha) \setminus \{L\}, \)
- \( A_1 = \{A_1, Q_1\}, \ldots, A_k = \{A_k, Q_k\} \) are all arguments with strength \( \alpha^9 \) that are acceptable w.r.t. \( W \) (according to Definition 2) and such that \( Q_j \notin \text{Warr}(\geq \alpha) \) and \( Q_j \neq \text{Block}(\geq \alpha), \) and
- \( B_1 = \{B_1, P_1\}, \ldots, B_n = \{B_n, P_n\} \) are all arguments with strength \( \alpha \) that satisfy the following conditions \( ^{10} \): (i) \( P_j \notin \text{Warr}(\geq \alpha) \) and \( P_j \notin \{Q_1, \ldots, Q_k\}, \)
- (ii) \( P_j \neq \text{Block}(\geq \alpha), \)
- (iii) for all \( \langle C, R \rangle \subset B_j \) with strength \( \beta > \alpha, R \in \text{Warr}(\beta) \) and for all \( \langle C, R \rangle \subset B_j \) with strength \( \alpha, R \in \{Q_1, \ldots, Q_k\} \cup \{P_1, \ldots, P_n\}, \)
- (iv) \( \Pi \cup \text{Warr}(\geq \alpha) \cup \{\langle C, R \rangle \subset B_j \} \cup \{P_j\} \neq \perp, \)
- (v) there exists an argument \( S \in \{A_1, \ldots, A_k\} \) such that \( S \subset B_j, \) and

\(^{9}\)Remark that for all argument \( A_j \in \{A_1, \ldots, A_k\} \) with \( Q_j \neq L, A_j \) does not depend on \( L \) and either \( Q_j \in \text{Warr}(\alpha) \) or \( Q_j \in \text{Block}(\alpha). \)

\(^{10}\)Remark that for all argument \( B_j \in \{B_1, \ldots, B_n\}, \) either \( B_j \) depends on \( L \) and \( (P_j, \alpha) \in \text{Warr}(\alpha) \cup \text{Block}(\alpha) \) or \( B_j \) depends on some \( Q_j \in \text{Block}(\alpha). \)
(vi) for every argument \( S \in \{A_1, \ldots, A_k\} \) such that \( S \subseteq B_j \), there does not exist a set of arguments \( G \subseteq \{A_1, \ldots, A_k\} \setminus \{S\} \) such that \( G \cup \{S\} \) generates a conflict w.r.t. \( W \).

**Proposition 14** (RP-DeLP program with unique output) Let \( \mathcal{P} = (\Pi, \Delta) \) be a RP-DeLP program and let \((\text{Warr}, \text{Block})\) be an output for \( \mathcal{P} \). \((\text{Warr}, \text{Block})\) is the unique output for \( \mathcal{P} \) iff for all literal \( L \in \text{Warr} \) there is no cycle in the graph for \( L \) w.r.t. \( \text{Warr} \).

Intuitively, given a literal \( L \) such that \( L \in \text{Warr}(\alpha) \), for some program preference level \( \alpha \), Definition 13 builds the warrant dependency graph for \( L \) and all acceptable arguments \( \{A_1, \ldots, A_k\} \) of strength \( \alpha \) that do not depend on \( L \) w.r.t. arguments \( \{B_1, \ldots, B_n\} \) of strength \( \alpha \) whose supports depend on \( L \) or on some argument in \( \{A_1, \ldots, A_k\} \). Then, according to Definition 11, the existence of a cycle expresses that the warranty of the argument for \( L \) depends on the validity of some \( B \in \{B_1, \ldots, B_n\} \), which depends on the warranty of some \( L' \in \{A_1, \ldots, A_k\} \) with \( L \neq L' \), which in turn depends on the validity of some \( B' \in \{B_1, \ldots, B_n\} \) with \( B' \neq B \), which in turn depends on the warranty of \( L \). Thus, for arguments of \( L \) and \( L' \) there does not exist a (unique) conflict evaluation order. Obviously, for RP-DeLP programs with unique output the set of warranted conclusions for every level \( \alpha \) can be (computed) defined by an unique conflict evaluation order between arguments. Next we show that programs of Example 12 have multiple outputs.

**Example 15** According to Definition 8, \( \text{Output}_1 = (\text{Warr}_1, \text{Block}_1) \) with \( \text{Warr}_1 = \{p\} \) and \( \text{Block}_1 = \{q, \sim q\} \) is an output for program \( \mathcal{P}_3 \) of Example 12. Then, according to Definition 13, Figure 1 shows the graph for \( p \) w.r.t. \( \text{Warr}_1 \); i.e. the warrant dependency graph for arguments \( \{A_1, A_2\} \) w.r.t. \( W \) and \( \{B_1, B_2\} \) with

\[
\begin{align*}
A_1 &= \{(p), p\}, A_2 &= \{(q), q\}, W = \emptyset, \\
B_1 &= \{(q, \sim p \leftarrow q), \sim p\} \text{ and } B_2 &= \{(p, \sim q \leftarrow p), \sim q\}.
\end{align*}
\]

Therefore, according to Proposition 14, \( \text{Output}_1 \) is not the unique output for \( \mathcal{P}_3 \) since there is a cycle in the graph for \( p \) w.r.t. \( \text{Warr}_1 \). Notice that \( \text{Output}_2 = (\text{Warr}_2, \text{Block}_2) \) with \( \text{Warr}_2 = \{q\} \) and \( \text{Block}_2 = \{p, \sim p\} \), is also an output for program \( \mathcal{P}_3 \) and the graph for \( q \) w.r.t. \( \text{Warr}_2 \) also contains a cycle.

Consider now the RP-DeLP program \( \mathcal{P}_4 \) of Example 12. According to Definition 8, \( \text{Output}_1 = (\text{Warr}_1, \text{Block}_1) \) with \( \text{Warr}_1 = \{y, p\} \) and \( \text{Block}_1 = \{q, s\} \), is an output for \( \mathcal{P}_4 \). Then, according to Definition 13, Figure 1 shows the graph for \( p \) w.r.t. \( W = \text{Warr}_1(1) = \{y\} \) proving that the output for \( \mathcal{P}_4 \) is not unique. Indeed, notice that \( \text{Output}_2 = (\text{Warr}_2, \text{Block}_2) \) with \( \text{Warr}_2 = \{y, q\} \) and \( \text{Block}_2 = \{p, r\} \), is also an output for program \( \mathcal{P}_4 \) and the graph for \( q \) w.r.t. \( \text{Warr}_2 \) also contains a cycle.

One of the main advantages of the warrant recursive semantics for RP-DeLP is from the implementation point of view. Actually, warrant semantics based on dialectical trees and, in general, rule-based argumentation frameworks like DeLP [7,9], might consider an exponential number of arguments with respect to the number of rules of a given program. In contrast, in our framework, at least for the particular case of RP-DeLP programs with unique output, it is not necessary to explicitly compute all the possible arguments for a given literal to check whether it is warranted, as we can implement an algorithm\(^{11}\) (not shown here due to space limitations) with a worst-case complexity in \( P^{NP} \).

\(^{11}\) Details can be found in the extended version at http://ia.udl.cat/ramon/comma2010full.pdf
5. Conclusions and future work

In this paper we have introduced a new recursive semantics for determining the warranty status of arguments in defeasible argumentation. The distinctive features of this semantics, e.g. with respect to Pollock’s critical link semantics, are: (i) it is based on a non-binary notion of conflict in order to preserve consistency with the strict knowledge and (ii) besides the set of warranted and rejected conclusions, we introduce the set of blocked conclusions, which are those conclusions which are based on warranted information but they generate a conflict with other already warranted conclusions of the same strength.

As future work we plan to formalize the maximal ideal output for RP-DeLP programs which will allow us to characterize the relationship between this unique output based on the recursive warrant semantics and the output of DeLP [10] and other general argumentation frameworks [5,4] based on the use of dialectical trees as underlying structures for characterizing the semantics of warranted conclusions.

Acknowledgments  Authors are thankful to the anonymous reviewers for their helpful comments. Research partially funded by the Spanish MICINN projects MULOG2 (TIN2007-68005-C04-01/02) and ARINF (TIN2009-14704-C03-01/03), CONSOLIDER (CSD2007-0022), and ESF Eurocores-LogICCC/MICINN (FFI2008-03126-E/FILO), and the grant JC2009-00272 from the Ministerio de Educación.

References