Prioritized norms in formal argumentation

BEISHUI LIAO, *Center for the study of Language and Cognition, Zhejiang University Hangzhou, 310028 China.*
E-mail: baiseliao@zju.edu.cn

NIR OREN, *Department of Computing Science University of Abereden AB24 3UE United Kingdom.*
E-mail: n.oren@abdn.ac.uk

LEENDERT VAN DER TORRE, *CSC Lab Maison du Nombre Belval, Eschr-sur-Alzette Luxembourg.*
E-mail: leon.vandertorre@uni.lu

SERENA VILLATA, *Laboratoire I3S Campus Tech - Building Templier 1 06902 Sophia Antipolis Cedex.*
E-mail: villata@i3s.unice.fr

Abstract

To resolve conflicts amongst norms, various non-monotonic formalisms can be used to perform prioritized normative reasoning. Meanwhile, formal argumentation provides a way to represent non-monotonic logics. In this paper we propose a representation of prioritized normative reasoning by argumentation. Using hierarchical abstract normative systems (HANS), we define three kinds of prioritized normative reasoning approaches called Greedy, Reduction and Optimization. Then, after formulating an argumentation theory for a HANS, we show that for a totally ordered HANS, Greedy and Reduction can be represented in argumentation by applying the weakest link and the last link principles, respectively, and Optimization can be represented by introducing additional defeats capturing the idea that for each argument that contains a norm not belonging to the maximal obeyable set then this argument should be rejected.

Keywords: deontic logic, formal argumentation, norms, normative systems, normative multiagent systems.

1 Introduction

Since the work of Alchourrón and Makinson [1] on hierarchies of regulations and their logic, in which a partial ordering on a code of laws or regulations is used to overcome logical imperfections in the code itself, reasoning with prioritized norms has been a central challenge in deontic logic [12, 4, 24].

The goal of this paper is to study the open issue of reasoning with priorities over norms through the lens of argumentation theory [10]. More precisely, we focus on reasoning with the abstract normative system proposed by Tosatto *et al.* [28], which in turn is based on Makinson and van der Torre's [20] approach to input/output logic. In this system, an abstract norm is represented by an ordered pair (a, x), where the body of the norm a is thought of as an input, representing some kind of condition or situation, and the head of the norm x is thought of as an output, representing what the norm tells us to be obligatory in that situation a. As a consequence, an abstract normative system is a directed

Hierarchical Abstract
Normative System

1 : (w, h)
2 : $(w, \neg o)$
3 : (h, o)

Extensions

Greedy : $\{h, \neg o\}$
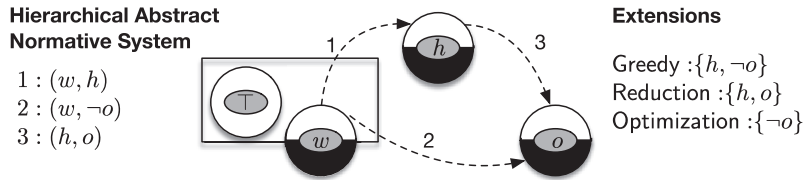Reduction : $\{h, o\}$
Optimization : $\{\neg o\}$

FIGURE 1. The Order puzzle example, represented using the graphical notation of Tosatto *et al.* [29] with edges annotated by norm strength.

graph (L, N) together with a context $C \subseteq L$ capturing base facts, where L is a set of nodes and $N \subseteq L \times L$ is the set of abstract norms. When the edge of an abstract normative system is associated with a number to indicate its priority over the other norms in the system, we obtain a hierarchical abstract normative system (HANS), which will be formally defined and studied in the remainder of this paper.

Let us clarify how a HANS is defined by considering the well-known Order Puzzle [17] example from the deontic logic literature, which revolves around three norms.

'Suppose that there is an agent, called Corporal O'Reilly, and that he is subject to the commands of three superior officers: a Captain, a Major, and a Colonel. The Captain, who does not like to be cold, issues a standing order that, during the winter, the heat should be turned on. The Major, who is concerned about energy conservation, issues an order that during the winter, the window should not be opened. And the Colonel, who does not like to be too warm and does not care about energy conservation, issues an order that, whenever the heat is on the window should be opened.'

Let w, h and o, respectively, denote the propositions that it is winter, the heat is turned on and the window is open. There are three norms (w, h), $(w, \neg o)$ and (h, o).

These three norms are visualized in Figure 1, extending the graphical notation described in the study by Tosatto *et al.* [29] by associating edges with numbers denoting priorities of norms. These priorities are obtained from the rank of the issuer, since Colonels outrank Majors and Majors outrank Captains. Within the figure, each circle denotes a proposition; the light part of the circle is the proposition itself, while the dark part denotes a negated proposition. Dashed lines represent the conditional obligations. Within Figure 1, the line from the light part of w to the dark part of o denotes $(w, \neg o)$. The box on the left represents the context in the example containing $\top$ and w.

In formalizing examples, one problem in applied logic is that the representation may be challenged. For example, it may be argued that the Colonel implies that if the window is closed, then the heating should be turned off. However, in normative systems, such pragmatic considerations are usually not part of the detachment procedure [16], with only explicitly given norms and commands being considered. Therefore, any such additional interpretations or other pragmatic concerns are out of the scope of this paper.

The central notion of inference in normative systems is called detachment. For example, in the Order Puzzle, the fundamental question is whether we can detach o, $\neg o$ or both. In the example, the formulas which can be derived from a normative system are obligations. In general, permissions and institutional facts can also be detached from normative systems. A detachment procedure therefore defines the way deontic facts are derived from a normative system. Different detachment procedures have been defined and studied in deontic logic, as well as in other rule-based systems. Moreover,

even in hierarchical normative systems, not all conflicts may be resolved. In such a case, the detachment procedure may derive several so-called extensions, each representing a set of obligations, permissions and institutional facts.

Abstract normative systems [28] were introduced as a common core representation for normative systems, which are still expressive enough to define the main detachment procedures. In particular, analogous to the main input/output logics, they have factual detachment built in, and have reasoning by cases, deontic detachment and identity as optional inference patterns [3, 16]. Such systems are called 'abstract' because negation is the only logical connective that is defined in the language. Furthermore, Tosatto *et al*. [28] considered elements and anti-elements rather than literals and propositions. It is straightforward to define more connectives within such systems, and it is also possible to define structured normative systems where the abstract elements are instantiated with logical formulas, e.g. with formulas of a propositional or modal logic. The latter more interesting representation of logical structure is analogous to the use of abstract arguments in formal argumentation. An advantage of abstract normative systems over structured ones is that the central inference of detachment can be visualized by walking paths in the graph. In other words, inference is represented by graph reachability. For example, in the Order puzzle, node o is reachable from the context, and thus it can be detached. Moreover, a conflict is represented by a node where both its light and the dark side are reachable from the context as in the case of node o in Figure 1.

There are several optional inference patterns for abstract normative systems because, as is well known, most principles of deontic logic have been criticized for some examples and applications. However, the absence of the same inference patterns is often criticized as well due to the lack of explanations and predictions of the detachment procedures resulting from these patterns. Therefore, current approaches to represent and reason with normative systems, such as input/output logic as well as abstract normative systems, do not restrict themselves to a single logic, but define a family of logics which can be embedded within them. Deciding which logic to use in a specific context depends on the requirements of the application. Similarly, with regards to permissions, there is an even larger diversity of deontic logics [15] which adopt different representations. For example, for each input/output logic, various notions of permission have been defined in terms of their relation to obligation. We refer to the work of Parent and van der Torre [24] for further explanation, discussion and motivation of this topic.

Now let us consider how various detachment procedures might apply norms in hierarchical normative systems in different orders, and result in different outcomes or extensions. Since detachment operates differently over permissions and obligations, and since the former introduces significant extra complexity, we only consider regulative norms and obligation detachment procedures in this paper. We examine three approaches describing well-known procedures defined in the literature [32, 5, 14]. These procedures all share one important property, namely that the context itself is not necessarily part of the procedure's output. It is precisely this feature which distinguishes input/output logics from traditional rule-based languages like logic programming or default logic [6, 26]. Such traditional rule-based languages where the input is part of the output are called throughput operators in input/output logic research. Therefore, the procedures we consider can naturally be captured using input/output logics. We refer to the three approaches as Greedy, Reduction and Optimization.

Greedy: the context contains propositions that are known to hold. This procedure always applies the norm with the highest priority that does not introduce inconsistency to an extension and the context. Here we say that a norm is applicable when its body is in the context or has been produced by other norms and added to the extension. In the Order puzzle example, we begin with the context $\{w\}$, and $(w, \neg o)$ is first applied. Then (w, h) is applied. Finally, (h, o) cannot be applied as this would result in a conflict, and so, by using Greedy, we obtain the extension $\{h, \neg o\}$.
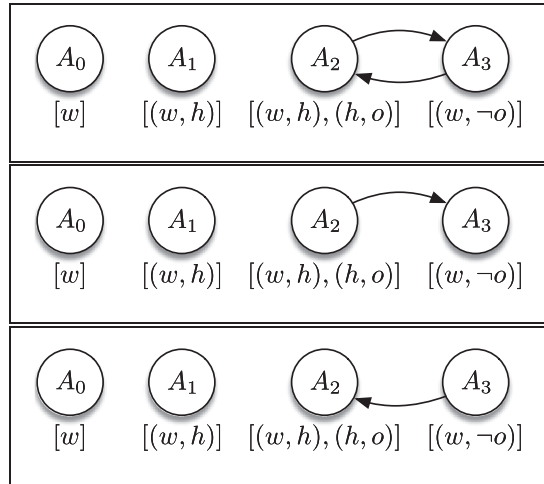
FIGURE 2. Top: the AF obtained from the order puzzle hierarchical normative system with the four arguments and attacks between them visualized as directed arrows. Middle: the AF obtained when last link is applied. Bottom: the AF obtained when weakest link is applied.

Reduction: in this approach, a candidate extension is guessed. All norms that are applicable according to this candidate extension are selected and transformed into unconditional or body-free norms. For example, a norm (a, b) selected in this way is transformed to a norm $(\top, b)$. The modified HANS with the transformed norms is then evaluated using Greedy. The candidate extension is selected as an extension by Reduction if it is identified as an extension according to this application of Greedy. In our example we select a candidate extension $\{h, o\}$, obtaining a set of body-free norms $\{(\top, h), (\top, \neg o), (\top, o)\}$. The priorities assigned to these norms are carried through from the original HANS, and are therefore 1, 2 and 3, respectively. After applying Greedy, we get an extension of Reduction: $\{h, o\}$. However, if we had selected the candidate extension $\{h, \neg o\}$, then this new extension would not appear in Greedy as $(\top, \neg o)$ has a lower priority than $(\top, o)$, and the latter is therefore not an extension of Reduction.

Optimization: in terms of Hansen's prioritized conditional imperatives, a set of maximally obeyable (i.e. minimally violated) norms is selected by choosing norms in order of priority, which are consistent with the context. Once these norms are selected, Greedy is applied to identify the extension. In our example, the maximal set of obeyable norms is $\{(h, o), (w, \neg o)\}$. Optimization therefore detaches the unique extension $\{\neg o\}$.

We can also consider the example in terms of formal argumentation. Given a HANS, we may construct an argumentation framework (AF). The top part of Figure 2 illustrates the AF obtained for the Order puzzle, which does not consider the priority relation between arguments.

An AF is a directed graph in which nodes denote arguments, and edges denote attacks between arguments. In the setting of a HANS, an argument is represented as a path within the directed graph starting from a node in the context. In this example, there are four arguments A_0, A_1, A_2 and A_3 represented as $[w]$, $[(w, h)]$, $[(w, h), (h, o)]$ and $[(w, \neg o)]$, respectively. Since the conclusions of A_2 and A_3 are inconsistent, A_2 attacks A_3 and vice versa. To obtain correct conclusions, we must take priorities between arguments into account, transforming attacks into defeats. This in turn requires lifting the priorities given for the constituents of an argument to a priority for the argument itself,

and two different principles have been commonly used for such a lifting. The last link principle ranks an argument based on the strength of its last inference, while the weakest link principle ranks an argument based on the strength of its weakest inference. In the order example, if the last link principle is applied, then $[(w, h), (h, o)]$ defeats $[(w, \neg o)]$. The corresponding AF is illustrated in the middle portion of Figure 2. If the weakest link principle is used instead, then $[(w, \neg o)]$ defeats $[(w, h), (h, o)]$. The corresponding AF is illustrated at the bottom of Figure 2. As a result, the former principle allows us to conclude $\{h, o\}$, while the latter concludes $\{h, \neg o\}$. In turn, the first result coincides with that obtained by Reduction, while the second is the same as that obtained by Greedy.

Inspired by the example above, we wish to investigate the links between the three detachment procedures for prioritized normative reasoning and argumentation theory. More specifically, our main research question is as follows:

> *How can the three detachment procedures (Reduction, Greedy and Optimization) proposed in the context of abstract normative reasoning be represented in formal argumentation?*

To answer this research question, we propose a formal framework to connect hierarchical normative reasoning with argumentation theory. More precisely, our framework represents the above-mentioned detachment procedures by lifting priorities from norms to arguments, with the underlying goal of making as few commitments as possible to specific argumentation systems. For this reason, we build on a structured AF which admits undercuts and rebuts between arguments, and allows for priorities between norms making up arguments. We show that variants of approaches to lifting priorities from rules to arguments allow us to capture both Greedy and Reduction, while the introduction of additional defeats allows us to obtain Optimization.

Some preliminary results found in this paper were originally presented in DEON 2016 [19]. We keep, but extend, the representation results for Greedy and Reduction (previously called the 'Brewka–Eiter construction'). To capture Optimization (previously called 'Hansen's construction') the DEON 2016 paper adds new arguments through the introduction of additional permissive norms. In the current paper, this is replaced with an approach which introduces additional defeats. The reasons for this replacement are twofold. First, the preliminary approaches for Optimization from the earlier work [19] cannot fully capture the idea that each argument containing a norm not belonging to the maximal obeyable set should be rejected. Second, permissive norms were used as a means to support the representation, which makes the approach more complicated.

The layout of the remainder of the paper is as follows: Section 2 formalizes the above-mentioned three detachment procedures of hierarchical normative reasoning (i.e. Greedy, Reduction and Optimization). In Section 3, we introduce an argumentation theory for a HANS. Section 4, 5 and 6 show how Greedy, Reduction and Optimization can be represented in argumentation. Finally, in Section 7 we discuss open problems and compare the proposed approach with related work, and in Section 7 we point out possible directions for future work.

2   HANS

In this section we formally introduce the notion of HANS and the three detachment procedures to compute their normative conclusions. A HANS captures the context of a system and the norms in force in such a system. There is an element in the universe called $\top$, contained in every context. In this paper, we consider only a finite universe. A HANS also encodes a ranking function over the norms to allow for the resolution of conflicts.

Based on the notion of abstract normative system defined by Tosatto *et al.* [28], a HANS can be defined as follows:

DEFINITION 2.1 (HANS)
A HANS is a tuple $\mathcal{H} = \langle L, N, C, r \rangle$, where

- $L = \text{E} \cup \{\neg e \mid e \in \text{E}\} \cup \{\top\}$ is the universe, a set of literals based on some finite set E of atomic elements;
- $N \subseteq L \times L$ is a finite set of regulative norms;
- $C \subseteq L$ is a subset of the universe, called a context, such that $\top \in C$ and for all e in E, $\{e, \neg e\} \nsubseteq C$;
- $r : N \to I\!N$ is a function from norms to natural numbers.

Regulative (ordinary) norms are of the kind 'if you turn on the heat, then you should open the window'. These norms are conditional norms, requiring some condition to hold (e.g. turning on the heat) before their conclusions can be drawn.

We write (a, x) for a regulative norm, where $a, x \in L$ are the antecedent and conclusion of the norm, respectively. Given (a, x), we use $r(a, x)$ to denote $r((a, x))$. Let $u, v \in N$ be two norms; we say that v is at least as preferred as u (denoted $u \leq v$) if and only if $r(u)$ is not larger than $r(v)$ (denoted $r(u) \leq r(v)$), where $r(u)$ is also called the rank of u. We write $u < v$ or $v > u$ if and only if $u \leq v$ and $v \nleq u$. Given a norm $u = (a, x)$, we write $ant(u)$ for a to represent the antecedent of the norm, and $cons(u)$ for x to represent the consequent of the norm. Given a set of norms $S \subseteq N$, we use $cons(S)$ to denote $\{cons(u) \mid u \in S\}$. We say that a HANS is totally ordered if and only if the ordering $\leq$ over N is anti-symmetric, transitive and total. Due to the finiteness of universe, the set of norms is finite. Note that given this assumption, the notion of total ordering here is identical to that of the full prioritization in Brewka and Eiter's [7] and Hansen's [13] work, and of the linearized ordering of Young *et al.* [32]. For $a \in L$, we write $\bar{a} = \neg a$ if and only if $a \in \text{E}$, and $\bar{a} = e$ for $e \in \text{E}$ if and only if $a = \neg e$. For a set $S \subseteq L$, we say that S is consistent if and only if there exist no $e_1, e_2 \in S$ such that $e_1 = \overline{e_2}$. To exemplify the notions of a HANS, consider the following example:

EXAMPLE 2.2 (Order puzzle)
In terms of Definition 2.1, the set of norms and priorities that are visualized in Figure 1 can be formally represented as a HANS $\mathcal{H} = \langle L, N, C, r \rangle$, where $L = \{w, h, o, \neg w, \neg h, \neg o, \top\}$, $N = \{(w, h), (h, o), (w, \neg o)\}$, $C = \{w, \top\}$, $r(w, h) = 1$, $r(h, o) = 3$, $r(w, \neg o) = 2$.

In the HANS setting, the three detachment procedures for prioritized normative reasoning can be defined as follows:

Firstly, Greedy detachment for a HANS always applies the norm with the highest priority amongst those which can be applied, if this does not bring inconsistency to the extension and the context. To formally define the notion of Greedy, we first introduce the following notions of paths and consistent paths:

DEFINITION 2.3 (Path, consistent path)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a HANS.

- A path in $\mathcal{H}$ from x_1 to x_n is a sequence of norms $(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)$ such that $\{(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)\} \subseteq N$, $n \geq 2$, and all norms of the sequence are distinct.

- A path in $\mathcal{H}$ from x_1 to x_n with respect to $R \subseteq N$ is a sequence of norms $(x_1, x_2), (x_2, x_3), \ldots,$ (x_{n-1}, x_n) such that $\{(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)\} \subseteq R, n \geq 2$, and all norms of the sequence are distinct.
- A consistent path in $\mathcal{H}$ from x_1 to x_n (with respect to R) is a path $(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)$ such that $\{x_1, x_2, \ldots, x_n\}$ is consistent.

Then, the unique extension of a totally ordered HANS by Greedy can be defined as follows: by first selecting a set of applicable norms with the highest priority in each step, and then collecting the final vertex of each path with respect to the set of selected norms.

DEFINITION 2.4 (Greedy)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a totally ordered HANS. For all $R \subseteq N$, let $R(C) = \{x \mid$ there is a path in $\mathcal{H}$ from an element in C to x with respect to $R\}$, and $Appl(N, C, R) := \{(a, x) \in N \setminus R \mid a \in C \cup cons(R),$ $\{x, \bar{x}\} \not\subseteq C \cup cons(R)\}$. The extension of $\mathcal{H}$ by Greedy, written as $Greedy(\mathcal{H})$, is the set $R(C)$ such that $R = \cup_{i=0}^{\infty} R_i$ is built inductively as follows:

$$R_0 = \emptyset$$
$$R_{i+1} = R_i \cup \max(N, C, R_i, r),$$

where $\max(N, C, R_i, r) = \{u \in Appl(N, C, R_i) \mid \forall v \in Appl(N, C, R_i) : r(u) \geq r(v)\}$.

In terms of Definition 2.4, the unique extension of the Order puzzle can be constructed as follows:

EXAMPLE 2.5 (Extension by Greedy)
Given $\mathcal{H}$ in Example 2.2, by Greedy, it holds that $R_0 = \emptyset$, $R_1 = \{(w, \neg o)\}$, $R_2 = \{(w, \neg o), (w, h)\}$, $R = \{(w, \neg o), (w, h)\}$. So, $Greedy(\mathcal{H}) = \{h, \neg o\}$.

Reduction is defined by first using a candidate extension to get a modified HANS, and then applying Greedy to it. If the candidate extension is an extension according to this application of Greedy, then it is an extension of the original HANS by Reduction. Formally, we have the following definition:

DEFINITION 2.6 (Reduction)
Given a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$, and a set X, let $\mathcal{H}^X = \langle L, N', C, r' \rangle$, where $N' = \{(\top, x_2) \mid (x_1, x_2) \in N, x_1 \in C \cup X\}$ is the set of ordinary norms, and $r'(\top, x_2) = \max(r(x_1, x_2) \mid (x_1, x_2) \in N, x_1 \in C \cup X)$ for all $(x_1, x_2) \in N$ are priorities over norms. An extension of $\mathcal{H}$ by Reduction is a set U such that U is $Greedy(\mathcal{H}^U)$. The set of extensions of $\mathcal{H}$ by Reduction is denoted as $Reduction(\mathcal{H})$.

According to Definition 2.6, it can be the case that different norms have the same consequent. To avoid the duplication of multiple body-free norms, only a single norm with the highest priority is used.

To exemplify the notion of Reduction, consider again the Order puzzle in Example 2.2.

EXAMPLE 2.7 (Extensions by Reduction)
By using Reduction, given $X = \{h, o\}$, we have $\mathcal{H}^X = \langle L, N', C, r' \rangle$, where $N' = \{(\top, h), (\top, o), (\top, \neg o)\}$, $r'(\top, h) = 1$, $r'(\top, o) = 3$ and $r'(\top, \neg o) = 2$. Since $X \in Greedy(\mathcal{H}^X)$, and no other set can be an extension, we have that $Reduction(\mathcal{H}) = \{\{h, o\}\}$.
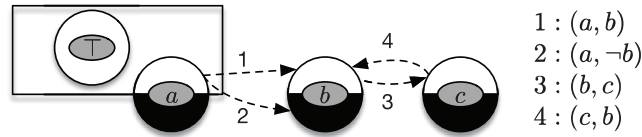
FIGURE 3. The HANS of Example 2.7 containing the two Reduction extensions $\{b, c\}$ and $\{\neg b\}$.

Note that a totally ordered HANS might have more than one extension, as illustrated by the following example:

EXAMPLE 2.8 (Multiple extensions by Reduction)
Given the HANS in Figure 3, assume that we have a context $C = \{a\}$. We then consider $X_1 = \{b, c\}$ and $X_2 = \{\neg b\}$. In the first case, we have $\mathcal{H}^{X_1} = \langle L, N', C, r' \rangle$, where $N' = \{(\top, b), (\top, \neg b), (\top, c)\}$ and $r'(\top, b) = 4$, $r'(\top, c) = 3$, $r'(\top, \neg b) = 2$. Here, $Greedy(\mathcal{H}^{X_1}) = \{b, c\}$, i.e. X_1.

In the second case, we obtain $\mathcal{H}^{X_2} = \langle L, N', C, r' \rangle$, where $N' = \{(\top, b), (\top, \neg b)\}$ and $r'(\top, b) = 1$, $r'(\top, \neg b) = 2$. Now, $Greedy(\mathcal{H}^{X_2}) = \{\neg b\}$, i.e. X_2. In this case, we therefore obtain two extensions using Reduction.

Finally, Optimization can be defined by first choosing a set of maximally obeyable norms, and then applying Greedy to it.

DEFINITION 2.9 (Optimization)
Given a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$, let $T = (u_1, u_2, \ldots, u_n)$ be the linear order on N such that $u_1 > u_2 > \ldots > u_n$. We define a sequence of sets of norms $R_0 \ldots R_n$ as follows:

$$R_0 = \emptyset$$

$$R_{i+1} = \begin{cases} lR_i \cup \{u_i\}, & \text{if } C \cup R(C) \text{ is consistent where } R = R_i \cup \{u_i\} \\ R_i, & \text{else.} \end{cases}$$

The unique extension of $\mathcal{H}$ by Optimization is $R_n(C)$ and denoted as $Optimization(\mathcal{H})$.

According to Definition 2.9, the unique extension of the Order puzzle can be constructed as follows:

EXAMPLE 2.10 (Extension by Optimization)
Regarding $\mathcal{H}$ in Example 2.2, by Optimization, let $u_1 = (h, o)$, $u_2 = (w, \neg o)$ and $u_3 = (w, h)$, and $T = (u_1, u_2, u_3)$. Then, it holds that $R_0 = \emptyset$, $R_1 = \{u_1\}$, $R_2 = \{u_1, u_2\}$ and $R = R_3 = R_2 = \{u_1, u_2\}$. So, we obtain that $Optimization(\mathcal{H}) = \{\neg o\}$.

3 Argumentation theory for a HANS

In this section, we introduce an argumentation theory on prioritized norms. Given a HANS, we first define arguments and defeats between them, then compute extensions of arguments in terms of Dung's [10] theory, and from these, obtain conclusions.

3.1 Arguments

In a HANS, an argument is an acyclic path in the graph starting in an element of the context. We assume minimal arguments—no norm can be applied twice in an argument and no redundant norm is included in an argument. We use $concl(\alpha)$ to denote the conclusion of an argument α, and $concl(E) = \{concl(\alpha) \mid \alpha \in E\}$ for the conclusions of a set of arguments E.

DEFINITION 3.1 (Arguments and sub-arguments)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a HANS.
A context argument in $\mathcal{H}$ is an element $a \in C$, and its conclusion is $concl(a) = a$.
An ordinary argument is a consistent path α in $\mathcal{H}$ from $x \in C$ to some $y \in L$. Moreover, we have that $concl(\alpha) = y$.
The sub-arguments of argument $[u_1, \ldots, u_n]$ are, for $1 \leq i \leq n$, $[u_1, \ldots, u_i]$. Note that context arguments do not have sub-arguments.

The set of all arguments constructed from $\mathcal{H}$ is denoted as $Arg(\mathcal{H})$. For readability, $[(a_1, a_2), \ldots, (a_{n-1}, a_n)]$ may be written as $(a_1, a_2, \ldots, a_{n-1}, a_n)$. The set of sub-arguments of an argument α is denoted as $sub(\alpha)$. When α is a sub-argument of β, we say that β is a super-argument of α. The set of super-arguments of α is denoted as $sup(\alpha)$. Furthermore, the set of proper sub-arguments of α is defined as $psub(\alpha) = sub(\alpha) \setminus \{\alpha\}$. For an ordinary argument $\alpha = [u_1, \ldots, u_n]$, we call u_n the *top norm* of α. Finally, in each ordinary argument, the norm with lowest priority is called the *weakest norm* of the argument. Formally, we have the following definition for this latter concept:

DEFINITION 3.2 (The weakest norm of an argument)
Let $\mathcal{H}$ be a totally ordered HANS. For an ordinary argument $\alpha = [u_1, \ldots, u_n]$ in $\mathcal{H}$, if $r(u_i) = \min(r(u_j) \mid 1 \leq j \leq n)$, then u_i is called the weakest norm of α.

3.2 Defeat relation between arguments

We follow the tradition in much of preference-based argumentation [2, 21], where *attack* captures a relation amongst arguments which ignores preferences, and *defeat* is a preference-aware relation on which the semantics is based. To define the defeat relation amongst prioritized arguments, we assume that *only* the priorities of the norms are used to compare arguments. In other words, we assume a lifting of the ordering on norms to a binary relation on sequences of norms (i.e. arguments), written as $\alpha \succeq \beta$, where α and β are two arguments, indicating that α is at least as preferred as β.

There is no common agreement about the best way to lift $\geq$ to $\succeq$. In argumentation, there are at least two ways to introduce weights. As an AF consists of a set of arguments and an attack relation between them, we can either assign weights to arguments, or we can assign weights to attacks. Traditionally, weights are assigned to arguments. Two common approaches to give the strength of an argument are the *weakest link* and the *last link* principles, combined with the elitist and democratic ordering [21]. For example, in the weakest link principle the weight of the argument is the weight of the weakest rule used in the argument. However, Young *et al.* [32] showed that elitist weakest link cannot be used to calculate $\succeq$ for Greedy, and proposes a *disjoint elitist order* which ignores shared rules. It is worth noticing that the strength of an argument may depend on the argument it is attacking, as identified by Young *et al.* [32]. Based on this idea, we define the orderings between arguments by assigning a strength to the attacks between the arguments to reflect the priority of the norms used in the arguments, following the same insights of the weakest link and last link principles

(denoted as $\succeq_w$ and $\succeq_l$, respectively). Defining the weakest link ordering in the same way as Young *et al.* [32], we have the following definition:

DEFINITION 3.3 (Weakest link and last link)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a HANS, and $\alpha = [u_1, \ldots, u_n]$ and $\beta = [v_1, \ldots, v_m]$ be two arguments in $Arg(\mathcal{H})$. Let $\Phi_1 = \{u_1, \ldots, u_n\}$ and $\Phi_2 = \{v_1, \ldots, v_m\}$. By the weakest link principle, $\alpha \succeq_w \beta$ iff $\exists v \in \Phi_2 \setminus \Phi_1$ such that $\forall u \in \Phi_1 \setminus \Phi_2, v \leq u$. By the last link principle, $\alpha \succeq_l \beta$ iff $u_n \geq v_m$.

   When the context is clear, we write $\succeq$ for $\succeq_w$, or $\succeq_l$. We write $\alpha \succ \beta$ for $\alpha \succeq \beta$ without $\beta \succeq \alpha$. The following proposition shows the transitivity of the two relations $\succeq_w$ and $\succeq_l$.

PROPOSITION 3.4 (Transitivity)
It holds that the relations $\succeq_w$ and $\succeq_l$ are transitive.

PROOF. Let $\alpha = [u_1, \ldots, u_n]$, $\beta = [v_1, \ldots, v_m]$ and $\gamma = [w_1, \ldots, w_k]$ be three arguments in $Arg(\mathcal{H})$. For the case of the weakest link, let $\Phi_1 = \{u_1, \ldots, u_n\}$, $\Phi_2 = \{v_1, \ldots, v_m\}$ and $\Phi_3 = \{w_1, \ldots, w_k\}$, $n, m, k \geq 1$. Let $x_{12} \in \Phi_1 \cap \Phi_2$, $x_{13} \in \Phi_1 \cap \Phi_3$, and $x_{23} \in \Phi_2 \cup \Phi_3$, $x_1 \in \Phi_1 \setminus (\Phi_2 \cup \Phi_3)$, $x_2 \in \Phi_2 \setminus (\Phi_1 \cup \Phi_3)$, and $x_3 \in \Phi_3 \setminus (\Phi_1 \cup \Phi_2)$. Assume that $\alpha \succeq_w \beta$ and $\beta \succeq_w \gamma$. There are only the following four possible cases:
   Case 1: There exists $x_{23} \in \Phi_2$, for all $x_{13}, x_1 \in \Phi_1$, $x_{23} \leq x_{13}$ and $x_{23} \leq x_1$; and since $x_{13} \in \Phi_3$, assume that for all $x_{12}, x_2 \in \Phi_2$, $x_{13} \leq x_{12}$, $x_{13} \leq x_2$. It follows that $x_{23} \leq x_{12}$ and $x_{23} \leq x_1$. Since $x_{23} \in \Phi_3$, it means that there exists $x_{23} \in \Phi_3$ such that $x_{23} \leq x_{12}$ and $x_{23} \leq x_1$, where $x_{12}, x_1 \in \Phi_1$. Hence, $\alpha \succeq_w \gamma$.
   Case 2: There exist $x_{23} \in \Phi_2$ and $x_3 \in \Phi_3$, for all $x_{13}, x_1 \in \Phi_1$, and $x_{12}, x_2 \in \Phi_2$: $x_{23} \leq x_{13}$, $x_{23} \leq x_1$, $x_3 \leq x_{12}$, $x_3 \leq x_2$. In this case, there are in turn only the following two possible sub-cases: either $x_{23} \leq x_3$ or $x_{23} > x_3$. If $x_{23} \leq x_3$, since $x_3 \leq x_{12}$, it holds that $x_{23} \leq x_{12}$. Since $x_{23} \leq x_1$ and $x_{23} \leq x_{12}$, it holds that $\alpha \succeq_w \gamma$. Second, if $x_{23} > x_3$, since $x_{23} \leq x_1, x_3 \leq x_1$. Since $x_3 \leq x_{12}$ and $x_3 \leq x_1$, $\alpha \succeq_w \gamma$.
   Case 3: There exists $x_2 \in \Phi_2$, for all $x_{13}, x_1 \in \Phi_1$, $x_2 \leq x_{13}$ and $x_2 \leq x_1$; and since $x_{13} \in \Phi_3$, assume that for all $x_{12}, x_2 \in \Phi_2$, $x_{13} \leq x_{12}$, $x_{13} \leq x_2$. In this case, there are in turn only the following two possible sub-cases: $x_{23} \leq x_{13}$ or $x_3 \leq x_{13}$, or $x_{23} > x_{13}$ and $x_3 > x_{13}$. If $x_{23} \leq x_{13}$ and $x_3 \leq x_{13}$, it holds that $\alpha \succeq_w \gamma$. If $x_{23} > x_{13}$ and $x_3 > x_{13}$, it holds that $x_{13} < x_{23}$ and $x_{13} \leq x_2$, and therefore $\beta \succeq_w \alpha$. Contradiction.
   Case 4: There exist $x_2 \in \Phi_2$ and $x_3 \in \Phi_3$, for all $x_{13}, x_1 \in \Phi_1$, and $x_{12}, x_2 \in \Phi_2$: $x_2 \leq x_{13}$, $x_2 \leq x_1, x_3 \leq x_{12}, x_3 \leq x_2$. In this case, since $x_3 \leq x_{12}$ and $x_3 \leq x_1$, it holds that $\alpha \succeq_w \gamma$.
   For the last link, if $\alpha \succeq_l \beta$ and $\beta \succeq_l \gamma$, then $u_n \geq v_m$ and $v_m \geq w_k$. It follows that $u_n \geq w_k$, and therefore $\alpha \succeq_l \gamma$. So, it holds that $\succeq_l$ is transitive. □

   Given a way to lift the ordering on norms to an ordering on arguments, the notion of defeat can be defined as follows:

DEFINITION 3.5 (Defeat amongst arguments)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a HANS. For all $\alpha, \beta \in Arg(\mathcal{H})$,

- α **attacks** β iff β has a sub-argument β' such that
    (i) $concl(\alpha) = \overline{concl(\beta')}$;

- α **defeats** β iff β has a sub-argument β' such that
  - (i) $concl(\alpha) = \overline{concl(\beta')}$ and
  - (ii) α is a context argument, or α is an ordinary argument and $\alpha \succeq \beta$.

The set of defeats between the arguments in $Arg(\mathcal{H})$ based on a preference ordering $\succeq$ is denoted as $Def(\mathcal{H}, \succeq)$.

In what follows, an argument $\alpha = [u_1, \ldots, u_n]$ with ranking on norms is denoted as $u_1 \ldots u_n : r(\alpha)$, where $r(\alpha) = (r(u_1), \ldots, r(u_n))$.

The notions of arguments and defeat relations between arguments by the weakest link and the last link principles, respectively, can be illustrated by the following example:

EXAMPLE 3.6 (Order puzzle continued)
Consider the HANS in Example 2.2. We have the following arguments (visually presented in the top of Figure 2):

$$A_0 : w \text{ (context argument)}$$
$$A_1 : (w, h) : (1) \text{ (ordinary argument)}$$
$$A_2 : (w, h)(h, o) : (1, 3) \text{ (ordinary argument)}$$
$$A_3 : (w, \neg o) : (2) \text{ (ordinary argument)}$$

We have that A_2 attacks A_3 and vice versa, and there are no other attacks amongst the arguments. Moreover, A_2 defeats A_3 by the last link principle (Figure 2, middle), and A_3 defeats A_2 by the weakest link principle (Figure 2, bottom).

3.3 Argument extensions and conclusion extensions

Given a set of arguments $\mathcal{A} = Arg(\mathcal{H})$ and a set of defeats $\mathcal{R} = Def(\mathcal{H}, \succeq)$, we get an AF $\mathcal{F} = (\mathcal{A}, \mathcal{R})$.

Following the notions of abstract argumentation by Dung [10], we say that a set $B \subseteq \mathcal{A}$ is *admissible*, if and only if it is *conflict-free* and it can defend each argument within the set. A set $B \subseteq \mathcal{A}$ is *conflict-free* if and only if there exist no arguments α and β in B such that $(\alpha, \beta) \in \mathcal{R}$. Argument $\alpha \in \mathcal{A}$ is *defended* by a set $B \subseteq \mathcal{A}$ (in such a situation α can also be said to be *acceptable* with respect to B) if and only if for all $\beta \in \mathcal{A}$, if $(\beta, \alpha) \in \mathcal{R}$, then there exists $\gamma \in B$ such that $(\gamma, \beta) \in \mathcal{R}$. Based on the notion of admissible sets, some other extensions can be defined. Formally, we have the following:

DEFINITION 3.7 (Conflict-freeness, defence and extensions)
Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF, and $B \subseteq \mathcal{A}$ a set of arguments.

- B is *conflict-free* if and only if $\nexists \alpha, \beta \in B$, such that $(\alpha, \beta) \in \mathcal{R}$.
- An argument $\alpha \in \mathcal{A}$ is defended by B (equivalently α is *acceptable* with respect to B), if and only if $\forall (\beta, \alpha) \in \mathcal{R}$, $\exists \gamma \in B$, such that $(\gamma, \beta) \in \mathcal{R}$.
- B is *admissible* if and only if B is conflict-free, and each argument in B is defended by B.
- B is a *complete* extension if and only if B is admissible and each argument in $\mathcal{A}$ that is defended by B is in B.
- B is a *preferred* extension if and only if B is a maximal (with respect to set-inclusion) complete extension.
- B is a *grounded* extension if and only if B is the minimal (with respect to set-inclusion) complete extension.

- B is a *stable* extension if and only if B is conflict-free, and $\forall \alpha \in \mathcal{A} \setminus B, \exists \beta \in B$ such that $(\beta, \alpha) \in \mathcal{R}$.

A semantics describes the set of extensions one wishes to obtain. We use $sem \in \{cmp, prf, grd, stb\}$ to denote the complete, preferred, grounded and stable semantics, respectively. A set of argument extensions of $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ is denoted as $sem(\mathcal{F})$. We write *Outfamily* for the set of conclusions from the extensions of the argumentation theory [30].

DEFINITION 3.8 (Conclusion extensions)
Given a HANS $\mathcal{H} = \langle L, N, C, r \rangle$, let $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq))$ be the AF constructed from $\mathcal{H}$. The conclusion extensions, written as *Outfamily*$(\mathcal{F}, sem)$, are the conclusions of the ordinary arguments within argument extensions.

$$Outfamily(\mathcal{F}, sem) = \{\{concl(\alpha) \mid \alpha \in S, \alpha \text{ is an ordinary argument}\} \mid S \in sem(\mathcal{F})\}$$

The following example shows that for a HANS, when adopting different principles for lifting the priorities over norms to those over arguments, the resulting AFs as well as their conclusion extensions may be different.

EXAMPLE 3.9 (Order puzzle in argumentation)
According to Example 3.6, let $\mathcal{A} = \{A_0, \ldots, A_3\}$. We have $\mathcal{F}_1 = (\mathcal{A}, \{(A_2, A_3)\})$, where $A_2 \succeq_l A_3$; and $\mathcal{F}_2 = (\mathcal{A}, \{(A_3, A_2)\})$, where $A_3 \succeq_w A_2$. For all $sem \in \{cmp, prf, grd, stb\}$, *Outfamily*$(\mathcal{F}_1, sem) = \{\{h, o\}\}$ and *Outfamily*$(\mathcal{F}_2, sem) = \{\{h, \neg o\}\}$.

We now turn our attention to the properties of the argumentation theory for a HANS.
First, according to Definition 3.5, we have the following proposition, capturing the relation between the attack/defeat on an argument and on its super-arguments.

PROPOSITION 3.10 (Super-argument attack and defeat)
Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF constructed from a HANS. For all $\alpha, \beta \in \mathcal{A}$, if α attacks β, then α attacks all super-arguments of β; if α defeats β, α defeats all super-arguments of β.

PROOF. When α attacks β, according to Definition 3.5, β has a sub-argument β' such that $concl(\alpha) = concl(\beta')$. Let γ be an super-argument of β. It follows that β' is a sub-argument of γ. Hence, α attacks γ.

When α defeats β, according to Definition 3.5, β has a sub-argument β' such that $concl(\alpha) = concl(\beta')$ and α is a context argument, or α is an ordinary argument and $\alpha \succ \beta'$. Since β' is a sub-argument of γ, α defeats γ.   □

Second, corresponding to properties of sub-argument closure and direct consistency in *ASPIC*$^+$ [21], we have the following two properties:

PROPOSITION 3.11 (Closure under sub-arguments)
Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF constructed from a HANS. For all $sem \in \{cmp, prf, grd, stb\}, \forall E \in sem(\mathcal{F})$, if an argument $\alpha \in E$, then $sub(\alpha) \subseteq E$.

PROOF. For every $\beta \in sub(\alpha)$, since α is acceptable with respect to E, it holds that β is acceptable with respect to E. This is because for each $\gamma \in \mathcal{R}$, if γ defeats β, then according to Proposition 3.10, γ defeats α; since $\alpha \in E$, there exists an $\eta \in E$ such that η defeats γ. Given that β is acceptable with respect to E and E is a complete extension, it holds that $\beta \in E$.   □

Since all norms in a HANS are defeasible, we only need to discuss direct consistency and contextual consistency.

PROPOSITION 3.12 (Direct consistency)
Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF constructed from a HANS. For all *sem* $\in \{cmp, prf, grd, stb\}$, $\forall E \in sem(\mathcal{F})$, $\{concl(\alpha) \mid \alpha \in E, \alpha$ is an ordinary argument$\}$ is consistent.

PROOF. Assume that there exist $\alpha, \beta \in E$ such that $concl(\alpha) = \overline{concl(\beta)}$. Since both α and β are ordinary arguments, α attacks β and β attacks α. If $\alpha \succeq \beta$ then α defeats β. Otherwise, β defeats α. In both cases, E is not conflict-free, contradicting the fact that E is a complete extension.   □

PROPOSITION 3.13 (Contextual consistency)
Let $\mathcal{F} = (\mathcal{A}, \mathcal{R})$ be an AF constructed from a HANS $\mathcal{H} = \langle L, N, C, r \rangle$. For all *sem* $\in \{cmp, prf, grd, stb\}$, $\forall E \in sem(\mathcal{F})$, $C \cup \{concl(\alpha) \mid \alpha \in E, \alpha$ is an ordinary argument$\}$ is consistent.

PROOF . Since C is consistent, we only need to verify that for all $a \in C$, for all $\alpha \in E$, $\{a, concl(\alpha)\}$ is consistent. We use proof by contradiction. Assume the contrary, i.e. $concl(\alpha) = \overline{a}$. It follows that a defeats α, and therefore $\alpha \notin E$. Contradiction. So the assumption is false, i.e. $C \cup \{concl(\alpha) \mid \alpha \in E, \alpha$ is an ordinary argument$\}$ is consistent. This completes the proof.   □

In the next sections we present representation results for the Greedy, Reduction and Optimization approaches introduced in Section 2, identifying equivalences between these approaches and the argument semantics based descriptions of a HANS.

4   Representation results for Greedy

Based on the idea introduced in Section 1, for a totally ordered HANS, we have the following lemma and proposition:

LEMMA 4.1 (Unique extension of Greedy)
Given a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$ and the corresponding AF $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_w))$, it holds that $\mathcal{F}$ is acyclic, and therefore has a unique extension under stable semantics.

PROOF. Since $\mathcal{H}$ is totally ordered, under $\succeq_w$, the relation $\succeq_w$ amongst arguments is acyclic. Assume the contrary. Then, there exist three distinct $\alpha, \beta, \gamma \in Arg(\mathcal{H})$ such that $\alpha \succeq_w \beta$, $\beta \succeq_w \gamma$ and $\gamma \succeq_w \alpha$. According to Definition 3.3, when $\mathcal{H}$ is totally ordered, it holds that $\alpha \succ_w \beta$, $\beta \succ_w \gamma$ and $\gamma \succ_w \alpha$. According to Proposition 3.4, $\alpha \succ_w \gamma$, contradicting $\gamma \succ_w \alpha$. Hence, $\mathcal{F}$ is acyclic, and therefore has a unique extension under stable semantics.   □

PROPOSITION 4.2 (Greedy is weakest link)
Given a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$ and the corresponding AF $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_w))$, it holds that $\{Greedy(\mathcal{H})\} = Outfamily(\mathcal{F}, stb)$.

PROOF. Our proof is constructive: given the Greedy extension, we show how to construct a stable extension of the AF whose conclusions coincide with the Greedy extension. Since the AF has only one extension (Lemma 4.2), this completes the proof.

By Definition 2.4, $Greedy(\mathcal{H})$ is the set of elements x such that there is a path in $\mathcal{H}$ from an element in C to x with respect to a set of norms $R \subseteq N$, inductively defined by $R = \cup_{i=0}^{\infty} R_i$, where $R_0 = \emptyset$ and $R_{i+1} = R_i \cup \max(N, C, R_i, r)$, $\max(N, C, R_i, r) = \{u \in Appl(N, C, R_i) \mid \forall v \in Appl(N, C, R_i) : r(u) \geq r(v)\}$ and $Appl(N, C, R) := \{(a, x) \in N \setminus R \mid a \in C \cup cons(R), \{x, \overline{x}\} \nsubseteq C \cup cons(R)\}$.

Let $E = \{a \in Arg(\mathcal{H}) \mid a \in C$ is a context argument$\} \cup \{[u_1, \ldots, u_n] \in Arg(\mathcal{H}) \mid n \geq 1, ant(u_1) \in C, \{ant(u_2) \ldots, ant(u_n), cons(u_n)\} \subseteq Greedy(\mathcal{H})\}$. From the consistency of the extensions (Proposition 3.11) and the construction of Greedy we know that there is a consistent path to any element of $Greedy(\mathcal{H})$, and thus for any argument of $Greedy(\mathcal{H})$ there is an ordinary argument with that element as its conclusion, and thus $Greedy(\mathcal{H}) = \{concl(\alpha) \mid \alpha \in E, \alpha$ is an ordinary argument$\}$. Now we only need to prove that E is a stable extension under weakest link.

We use proof by contradiction. Assume that E is not a stable extension. Given that $Greedy(\mathcal{H})$ is consistent, and thus the extension E is conflict-free , and given that context arguments defeat all conflicting ordinary arguments and are thus included in E, there must be $\alpha = [u_1, \ldots, u_n] \in Arg(\mathcal{H}) \backslash E$, such that α is not defeated by any argument in E. Since $\alpha \notin E$, there exist $\alpha' = [u_1, \ldots, u_{j-1}] \in E$ and $\alpha'' = [u_1, \ldots, u_j] \notin E$. Let $S \subseteq E$ be the set of arguments conflicting with α'', thus each of them has a conclusion $\overline{cons(u_j)}$. So, α'' defeats each argument $\beta = [v_1, \ldots, v_m]$ in S using weakest link, i.e. $\alpha'' \succeq_w \beta$. If $\Phi_1 = \{u_1, \ldots, u_n\}$ and $\Phi_2 = \{v_1, \ldots, v_m\}$, then according to Definition 3.2, $\exists v \in \Phi_2 \setminus \Phi_1$ such that $\forall u \in \Phi_1 \setminus \Phi_2, v \leq u$. Let $M \subseteq N$ be the set of all these norms v, i.e. all norms that occur in the arguments of S which are smaller than all the norms in α'', excluding norms that occur in α'' itself.

We now consider the point in the construction of the greedy extension where an element of M was added to it, we consider $Appl(N, C, R)$ and $\max(N, C, R_i, r)$, and we derive the contradiction. Let i be the lowest index such that R_i does not contain a norm of M. At this moment, $Appl(N, C, R_i)$ also contains a norm in α'', and $\max(N, C, R_i, r)$ contains an element of M. However, by definition, all norms of M are ranked lower than the norms in α''. Contradiction. So the assumption is false, i.e. E is a stable extension. This completes the proof. □

Note that Proposition 4.2 corresponds to Theorem 5.3 of Young *et al*. [32]. This correspondence arises as follows. First, in the argumentation theory for a HANS, we use disjoint elitist order to compare sets of norms, while in the argumentation theory for prioritized default logic (PDL), Young *et al*. [32] use a new order called a structure-preference order, which takes into account the structure of how arguments are constructed. Since in the setting of HANS, arguments are acyclic paths, it is not necessary to use the structure-preference order to compare arguments. Second, due to the different ways of constructing AFs, the proof of Proposition 4.2 differs from that of Theorem 5.3 of Young *et al*. [32]. The former considers the order of the applicability of norms in the proof, while the latter uses the mechanism defined in the structure-preference order.

5   Representation result for Reduction

According to Brewka and Eiter [7], Reduction is based on the following two points:

1) *The application of a rule with non-monotonic assumptions means jumping to a conclusion. This conclusion is yet another assumption which has to be used globally in the programme for the issue of deciding whether a rule is applicable or not.*
2) *The rules must be applied in an order compatible with the priority information.*

This *global* view of deciding whether a rule is applicable coincides with the last link principle of lifting a preference relation between rules to a priority relation between resulting arguments. According to Definition 2.6 and the argumentation theory for a HANS, we have the following representation result:

PROPOSITION 5.1 (Reduction is last link)
Given a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$ and the corresponding AF $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_l))$, it holds that $Reduction(\mathcal{H}) = Outfamily(\mathcal{F}, stb)$.

PROOF. We prove $Reduction(\mathcal{H}) = Outfamily(\mathcal{F}, stb)$ by first proving that the left-hand side is a subset of the right-hand side, and then by proving that the right-hand side is a subset of the left-hand side.

We first prove $Reduction(\mathcal{H}) \subseteq Outfamily(\mathcal{F}, stb)$.

Assume a set $U \in Reduction(\mathcal{H})$. According to Definition 2.6 we have $Greedy(\mathcal{H}^U)$ where $\mathcal{H}^X = \langle L, N', C, r' \rangle$, $N' = \{(\top, x_2) \mid (x_1, x_2) \in N, x_1 \in C \cup X\}$ is the set of ordinary norms, and $r'(\top, x_2) = \max(r(x_1, x_2) \mid (x_1, x_2) \in N, x_1 \in C \cup X)$ for all $(x_1, x_2) \in N$ are priorities over norms.

We now construct a stable extension of $\mathcal{F}$ such that its conclusions are exactly U. Let E be the set of arguments that can be constructed from elements of U, just like in the proof of Greedy. So $E = \{a \in Arg(\mathcal{H}) \mid a \in C$ is a context argument$\} \cup \{[u_1, \ldots, u_n] \in Arg(\mathcal{H}) \mid n \geq 1, ant(u_1) \in C, \{ant(u_2) \ldots, ant(u_n), cons(u_n)\} \subseteq U\}$. From the consistency of the extensions (Proposition 3.11) and the construction of Reduction we know that there is a consistent path to any element of U, and thus for any argument of U there is an ordinary argument with that element as its conclusion, and thus $U = \{concl(\alpha) \mid \alpha \in E, \alpha$ is an ordinary argument$\}$. Now we only need to prove that E is a stable extension.

We use proof by contradiction, just like in the proof of Greedy. Assume that E is not stable. Given that context arguments defeat all conflicting ordinary arguments, there must be a $\beta = [v_1, \ldots, v_n] \in Arg(\mathcal{H}) \setminus E$, such that β is not defeated by any argument in E. Since $\beta \notin E$, there exists $\beta' = [v_1, \ldots, v_j]$ in E and $\beta'' = [v_1, \ldots, v_{j+1}]$ not in E. Let $S \subseteq E$ be the set of arguments attacking β'', thus each of them has a conclusion $\overline{cons(v_{j+1})}$. Since β'' is not defeated by any argument in E, β'' defeats each argument in S. Then, the last link of β'' is higher than the last links of all arguments of S. Consequently, the rank of v_{j+1} is higher than the rank of all norms with consequent $\overline{cons(v_{j+1})}$. But this means that in N', the rank of $(\top, cons(v_{j+1}))$ is higher than the rank of $(\top, \overline{cons(v_{j+1})})$. Then by the construction of Reduction, i.e. $Greedy(\mathcal{H}^U)$, we have that $cons(v_{j+1})$ must be in U, a contradiction. Thus the assumption is false, i.e. E is stable, and this completes the proof that $Reduction(\mathcal{H}) \subseteq Outfamily(\mathcal{F}, stb)$.

To complete our proof, we now show that $Reduction(\mathcal{H}) \supseteq Outfamily(\mathcal{F}, stb)$. Let E be a stable extension of $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_l))$. Thus E is conflict-free, closed under sub-arguments, and for all arguments not in E, there is an argument in E defeating it. In particular, for any $\beta' = [v_1, \ldots, v_j]$ in E and $\beta'' = [v_1, \ldots, v_{j+1}]$ not in E, there is an argument in E defeating β''.

Moreover, let $U = \{concl(\alpha) \mid \alpha \in E, \alpha$ is an ordinary argument$\}$. From the contextual consistency of stable extensions (Proposition 3.12), we know that $U \cup C$ is consistent. Since defeat is based on last link, it means that for every norm applicable in U whose conclusion c is not in U, there is a higher-ranked norm applicable in U whose conclusion is $\bar{c}$. Consider $\mathcal{H}^U = \langle L, N', C, r' \rangle$, $N' = \{(\top, x_2) \mid (x_1, x_2) \in N, x_1 \in C \cup U\}$ is the set of ordinary norms, and $r'(\top, x_2) = \max(r(x_1, x_2) \mid (x_1, x_2) \in N, x_1 \in C \cup U)$ for all $(x_1, x_2) \in N$ are priorities over norms. Assume there is an $x \in U$ with $(\top, x), (\top, \bar{x}) \in N'$. Due to the above the rank of $(\top, x)$ is higher than the rank of $(\top, \bar{x})$, and thus $x \in Greedy(\mathcal{H}^U)$.

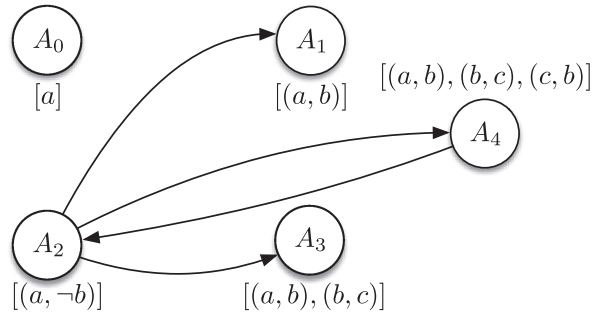FIGURE 4. The AF obtained for the HANS of Figure 3.

Thus $U = Greedy(\mathcal{H}^U)$, i.e. $U \in Reduction(\mathcal{H})$, and that completes our proof.   □

Now, let us consider the following three examples which show that by Reduction, a HANS and the corresponding AF may have a unique extension, multiple extensions or an empty extension.

EXAMPLE 5.2 (Order puzzle, Reduction)
Consider Example 3.9 when the last link principle is applied, A_2 defeats A_3. Then, we have *Outfamily*$(\mathcal{F}_1, stb) = \{\{h, o\}\}$, which is equal to *Reduction*$(\mathcal{H})$.

EXAMPLE 5.3 (Reduction, multiple extensions)
Consider the HANS of Figure 3 when the last link principle is applied. We obtain the AF shown in Figure 4, written as $\mathcal{F}_3$, yielding *Outfamily*$(\mathcal{F}_3, stb) = \{\{b, c\}, \{\neg b\}\}$. Note that here, we have two distinct stable extensions.

Since stable extensions do not necessarily exist for all AFs, the Reduction of a HANS might not exist.

EXAMPLE 5.4 (Reduction, empty extension)
Consider the HANS $\mathcal{H} = \langle L, N, C, r \rangle$, where $N = \{(c, d), (p, \neg d), (z, \neg c), (\neg d, \neg z), (r, z), (b, c)\}$, $C = \{b, r, p\}$ and $r(c, d) = 5$, $r(p, \neg d) = 4$, $r(z, \neg c) = 6$, $r(\neg d, \neg z) = 2$, $r(r, z) = 1$ and $r(b, c) = 0$. When the last link principle is applied, this HANS yields the AF shown in Figure 5, which has no stable extension.

6   Representation result for Optimization

In principle, Optimization is realized by adding norms in order of priority which are consistent with the context, until no more norms can be added, obtaining a maximal set of obeyable norms, following which conclusions can be computed. For each norm that does not belong to the maximal set of obeyable norms, if it is in a consistent path of a HANS $\mathcal{H}$, it is the weakest norm of this path. In terms of the terminology of argumentation, a consistent path of $\mathcal{H}$ is an argument of the corresponding AF. Hence, for an ordinary argument, if its top norm is the weakest norm that does not belong to the maximal obeyable set, then this argument should not be accepted. In other words, this argument should be defeated by some accepted argument. Furthermore, when an argument is rejected, all its
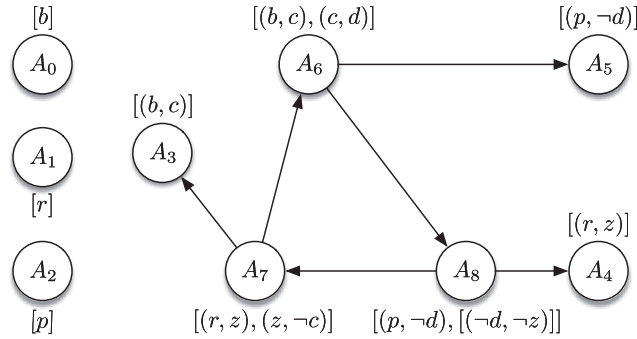
FIGURE 5. An AF with no stable extension.

super-arguments should be rejected. For instance, as illustrated in Figure 1, the maximal obeyable set of norms is $\{(h, o), (w, \neg o)\}$. The norm (w, h) does no belong to this set. So, the arguments containing (w, h), i.e. $[(w, h)]$ and $[(w, h), (h, o)]$, should be rejected.

In this paper, given an argument α, a sub-argument of α is called the *weakest sub-argument* of α, if it is an ordinary argument and its top norm is the weakest norm of α.

DEFINITION 6.1 (Weakest sub-argument)
Given a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$, and the corresponding AF $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_w))$, for all $\alpha = [u_1, \ldots, u_n] \in Arg(\mathcal{H})$, if u_i is the weakest norm of α, then the ordinary sub-argument of α whose top norm is u_i is called the weakest sub-argument of α.

Then, the set of super-arguments of α'—the weakest sub-argument of α— is called the weakest arguments with respect to α, formally defined as follows:

DEFINITION 6.2 (Weakest argument)
Given a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$, and the corresponding AF $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_w))$, for all $\alpha = [u_1, \ldots, u_n] \in Arg(\mathcal{H})$, let α' be the weakest sub-argument of α. Then, the set of weakest arguments with respect to α is defined as $warg(\alpha) = sup(\alpha')$.

The notion of weakest arguments can be illustrated by the following example:

EXAMPLE 6.3 (Order puzzle, weakest arguments)
Consider the AF in Figure 2. When applying the weakest link principle, A_3 defeats A_2. According to Definition 6.2, $warg(A_1) = \{A_1\}$, $warg(A_2) = \{A_1, A_2\}$ and $warg(A_3) = \{A_3\}$.

Based on the concept of weakest arguments, when an argument α defeats another argument β according to the weakest link principle, there are two different cases, namely that α is in the set of weakest arguments with respect to β, or that it is not.

First, consider the case where α is in $warg(\beta)$. In this case, it means that the weakest norm of β is also the weakest norm of α.

PROPOSITION 6.4 (Shared weakest norm)
Given $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_w))$, for all $(\alpha, \beta) \in Def(\mathcal{H}, \succeq_w)$, if u is the weakest norm of β and $\alpha \in warg(\beta)$, then u is the weakest norm of α.

PROOF. We use proof by contradiction. Assume that u is not the weakest norm of α. Then, there exists u' such that $u > u'$. Since u' does not belong to β, it holds that $\beta \succ_w \alpha$. This contradicts $(\alpha, \beta) \in Def(\mathcal{H}, \succeq_w)$. Thus the assumption is false, i.e. u is the weakest norm of α. $\square$

Given that α and β share a weakest norm, according to the definition of Optimization, this weakest norm does not belong to the maximal obeyable set. Hence, both α and β cannot be accepted. In this paper we use an auxiliary argument aux to defeat all arguments in $warg(\beta)$ that should not be accepted.

The second case occurs when α is not in $warg(\beta)$. Here, the weakest norm of α is superior to the weakest norm of β. So, if α is accepted, then all arguments in $warg(\beta)$ should be rejected. For this purpose, we may use α to defeat each argument in $warg(\beta)$. In addition, since the defeat from α to each super-argument of β is already in $Arg(\mathcal{H})$, only the arguments in $warg(\beta) \cap psub(\beta)$ should be added to $Arg(\mathcal{H})$.

According to the above observations, an expanded AF of $\mathcal{F}$ with auxiliary defeats on weakest arguments is defined as follows:

DEFINITION 6.5 (Expanded AF with additional defeats on weakest arguments)
Let $\mathcal{F} = (Arg(\mathcal{H}), Def(\mathcal{H}, \succeq_w))$ be an AF that is constructed from a totally ordered HANS $\mathcal{H} = \langle L, N, C, r \rangle$, and aux be an auxiliary argument such that aux $\notin Arg(\mathcal{H})$. The expanded AF of $\mathcal{F}$ with auxiliary defeats on weakest arguments is $\mathcal{F}' = (Arg(\mathcal{H}) \cup \{aux\}, Def(\mathcal{H}, \succeq_w) \cup \Phi_1 \cup \Phi_2)$, where $\Phi_1 = \cup_{(\alpha, \beta) \in Def(\mathcal{H}, \succeq_w) \wedge \alpha \notin warg(\beta)} \{(\alpha, \gamma) \mid \gamma \in warg(\beta) \cap psub(\beta)\}$ and $\Phi_2 = \cup_{(\alpha, \beta) \in Def(\mathcal{H}, \succeq_w) \wedge \alpha \in warg(\beta)} \{(aux, \gamma) \mid \gamma \in warg(\beta)\}$.

In Definition 6.5, Φ_1 is the set of defeats from α that is not a weakest argument with respect to β, and Φ_2 is the set of defeats from the auxiliary argument aux when α is a weakest argument with respect to β.

The following lemma and proposition show that Optimization can be represented in formal argumentation by using weakest link together with auxiliary defeats.

LEMMA 6.6 (Unique extension of Optimization)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a totally ordered HANS, and $\mathcal{F}'$ be an AF of $\mathcal{H}$ with additional defeats on weakest arguments presented in Definition 6.2. It holds that $\mathcal{F}'$ is acyclic, and therefore has a unique extension under stable semantics.

PROOF. According to Lemma 4.1, $\mathcal{F}$ is acyclic. The addition of Φ_2 does not produce cycles. Meanwhile, for all $(\alpha, \gamma) \in \Phi_1$, since α defeats β and γ is a weakest argument with respect to β, $\alpha \succeq_w \gamma$. So, the addition of Φ_2 does not produce cycles either. As a result, $\mathcal{F}'$ is acyclic, and therefore has a unique extension under stable semantics. $\square$

PROPOSITION 6.7 (Optimization is weakest link plus auxiliary defeats)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a totally ordered HANS, and $\mathcal{F}'$ be an AF of $\mathcal{H}$ with additional defeats on weakest arguments presented in Definition 6.2. It holds that $Optimization(\mathcal{H}) = \{concl(\alpha) \mid \alpha \in E \setminus \{aux\}, \alpha$ is an ordinary argument$\}$, where E is the unique stable extension of $\mathcal{F}'$.
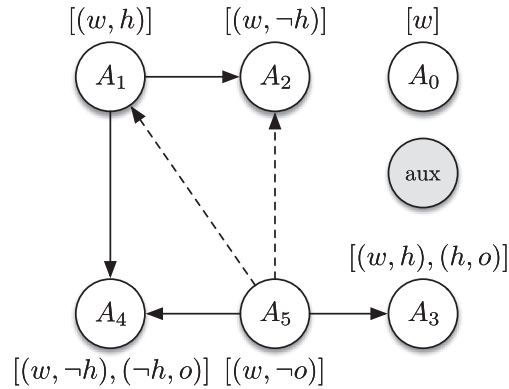
FIGURE 6. AF expanded by an auxiliary argument and two auxiliary defeats (denoted using dashed lines).

PROOF. This proof is similar to that of Greedy: given the Optimization extension, we show how to construct a stable extension of the expanded AF whose conclusions coincide with the Optimization extension. Since the AF has only one extension (Lemma 6.6), this completes the proof.

By Definition 2.9, $Optimization(\mathcal{H})$ is the set of elements x such that there is a path in $\mathcal{H}$ from an element in C to x with respect to a set of norms $R \subseteq N$. Given $T = (u_1, u_2, \ldots, u_n)$ the linear order on N such that $u_1 > u_2 > \ldots > u_n$, R is inductively defined by $R = R_n$, where $R_0 = \emptyset$; $R_{i+1} = R_i \cup \{u_i\}$ if $C \cup R(C)$ is consistent where $R = R_i \cup \{u_i\}$, and $R_{i+1} = R_i$ otherwise.

Let $E = \{aux\} \cup \{\alpha \in Arg(\mathcal{H}) \mid \alpha \in C$ is a context argument$\} \cup \{[u_1, \ldots, u_n] \in Arg(\mathcal{H}) \mid n \geq 1, ant(u_1) \in C, \{ant(u_2) \ldots, ant(u_n), cons(u_n)\} \subseteq Optimization(\mathcal{H})\}$. From the consistency of the extensions (Proposition 3.11) and the construction of Optimization we know that there is a consistent path to any element of $Optimization(\mathcal{H})$, and thus for any argument of $Optimization(\mathcal{H})$ there is an ordinary argument with that element as its conclusion, and thus $Optimization(\mathcal{H}) = \{concl(\alpha) \mid \alpha \in E \setminus \{aux\}, \alpha$ is an ordinary argument$\}$.

Now we only need to prove that E is a stable extension of $\mathcal{F}'$.

We use proof by contradiction. So assume that E is not a stable extension. On the one hand, given that $Optimization(\mathcal{H})$ is consistent, and thus the extension E is conflict-free, and given that context arguments defeat all conflicting ordinary arguments and are thus included in E, there must be $\alpha = [u_1, \ldots, u_n] \in Arg(\mathcal{H}) \setminus E$, such that α is not defeated by any argument in E. On the other hand, since $\alpha \notin E$, there exists a weakest norm u_j of α such that $u_j \notin R$. Let i be the index where u_j can not be added to R_i since $C \cup R(C)$ is consistent where $R = R_{j+1} = R_i \cup \{u_j\}$. Let $\alpha' = [u_1, \ldots, u_k]$ where $k \geq j$, and $\beta = [v_1, \ldots, v_m]$ such that $\{v_1, \ldots, v_m\} \subseteq R_{i+1}$ and $v_m = \overline{u_k}$. Since $u_j < v$ for all $v \in R_i$, it holds that β defeats α'. Then, we have the following two possible cases. If $u_j \in \{v_1, \ldots, v_m\}$, then α is defeated by $aux \in E$. Contradiction. Otherwise, $u_j \notin \{v_1, \ldots, v_m\}$. In this case, $\{v_1, \ldots, v_m\} \subseteq R_i$ and $C \cup R(C)$ is consistent where $R = R_i$. So, β is in E. Contradiction. So the assumption is false, i.e. E is a stable extension. This completes the proof. □

Now, let us consider a revised version of the Order puzzle as follows:

EXAMPLE 6.8 (Order puzzle, Optimization)
Let $\mathcal{H}' = \langle L, N, C, r \rangle$ be a HANS, where $L = \{w, h, o, \neg w, \neg h, \neg o, \top\}$, $N = \{(w, h), (w, \neg h), (h, o), (\neg h, o), (w, \neg o)\}$, $C = \{w, \top\}$, $r(w, h) = 1$, $r(w, \neg h) = 0$, $r(h, o) = 3$, $r(\neg h, o) =$
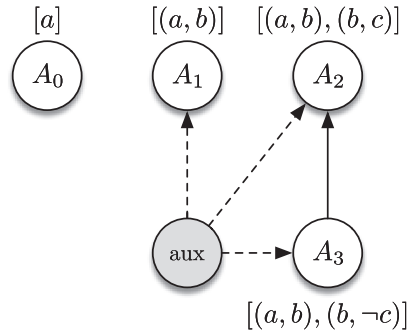
FIGURE 7. AF expanded by an auxiliary argument and three auxiliary defeats (denoted using dashed lines).

4, $r(w, \neg o) = 2$. The maximal set of obeyable norms is $R = \{(\neg h, o), (h, o), (w, \neg o)\}$ and so *Optimization*$(\mathcal{H}') = \{\neg o\}$. Figure 6 illustrates the AF obtained from this HANS, by adding an auxiliary argument and two auxiliary defeats. In this example, since A_5 is neither in $warg(A_3) = \{A_1, A_3\}$ nor in $warg(A_4) = \{A_2, A_4\}$, the two auxiliary defeats are from A_5 to A_1 and A_2, respectively. Then, under stable semantics, the expanded AF has a unique extension $E = \{aux, A_0, A_5\}$. As a result, the set of conclusions $\{concl(\alpha) \mid \alpha \in E \setminus \{aux\}, \alpha \text{ is an ordinary argument}\} = \{concl(A_5)\} = \{\neg o\}$.

Finally, let us consider an expanded AF where auxiliary defeats are from the auxiliary argument aux.

EXAMPLE 6.9 (Empty Optimization)
Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a HANS, where $L = \{a, b, c, \neg a, \neg b, \neg c, \top\}$, $N = \{(a, b), (b, c), (b, \neg c)\}$, $C = \{a, \top\}$, $r(a, b) = 1$, $r(b, c) = 2$, $r(b, \neg c) = 3$. The maximal set of obeyable norms is $R = \{(b, c), (b, \neg c)\}$ and so *Optimization*$(\mathcal{H}) = \emptyset$. Figure 7 illustrates the AF obtained from this HANS, by adding an auxiliary argument and three auxiliary defeats. In this example, since A_3 is in $warg(A_2) = \{A_1, A_2, A_3\}$, the three auxiliary defeats are from aux. Then, under stable semantics, the expanded AF has a unique extension $E = \{aux, A_0\}$. As a result, the set of conclusions $\{concl(\alpha) \mid \alpha \in E \setminus \{aux\}, \alpha \text{ is an ordinary argument}\} = \emptyset$.

7 Discussions and related work

In the previous sections, after defining three detachment procedures for totally ordered HANS motivated by the Order puzzle example, we formulated an argumentation theory to represent them. The role of examples in the study of logic has a long and rich history. Traditionally, a logic was proposed to model some example problem, following which examples were introduced to highlight paradoxes or inconsistencies in the logic, whereupon a new logic was proposed to address these problems, and the cycle repeated. While this approach has significantly enriched the field, it is not without problems. For example, there is still a debate regarding deontic detachment within the community, as in some cases, deontic detachment intuitively holds, and in other cases it does not [23]. Given this, we do not seek to claim that the detachment procedures we present in this paper are in any sense the 'right' logics. Instead, our goal is to answer the following questions:

  (i) What are the general properties of systems considered relevant to some problem?
  (ii) Given an application, what choices should be made in order to obtain a solution?
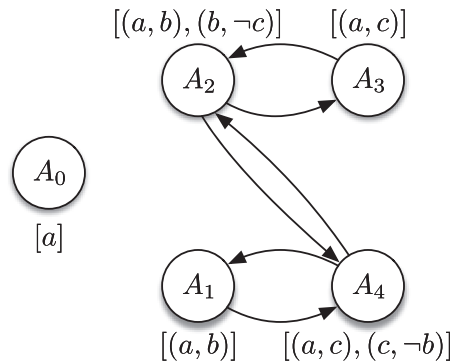
FIGURE 8. AF constructed from a pre-ordered HANS.

In this paper, the systems we considered are the three different detachment procedures, encoded in the general framework of HANS. The property we considered is then the conclusions that one can draw from each of the detachment procedures in the context of prioritized norms, which we describe in the context of an argumentation system.

Our results then characterize the outputs of Greedy, Reduction and Optimization in terms of argumentation for a totally ordered HANS, allowing one to decide which approach is relevant to their needs by understanding the effects of each approach through the argumentation literature. The semantics associated with each approach also sheds light on the complexity of computing conclusions in the normative context.

Furthermore, it is important to note that our representation results only hold when the HANS is totally ordered. This is illustrated by the following example, which shows that Greedy does not match the results of weakest link when one does not have a total order over preferences.

Let $\mathcal{H} = \langle L, N, C, r \rangle$ be a HANS, where $N = \{(a, b), (a, c), (b, \neg c), (c, \neg b)\}$ $C = \{a\}$ and $r(a, b) = 1$, $r(a, c) = 1$, $r(b, \neg c) = 2$ and $r(c, \neg b) = 2$. On the one hand, by Greedy, there are two extensions $\{b, \neg c\}$ and $\{c, \neg b\}$. On the other hand, by the weakest link principle, the AF constructed from $\mathcal{H}$ is illustrated in Figure 8. Under stable semantics, there are three extensions $\{A_0, A_1, A_2\}$, $\{A_0, A_1, A_3\}$ and $\{A_0, A_3, A_4\}$. So, there are three conclusion extensions $\{b, \neg c\}$, $\{b, c\}$, $\{c, \neg b\}$. As a result, the set of conclusions obtained by Greedy is not equal to the one obtained by argumentation using the weakest link principle. We leave identifying valid representation results for HANS containing preference pre-orders as an avenue for future research.

Regarding related work, Young *et al.* [32] endowed Brewka's PDL with argumentation semantics using the *ASPIC*$^+$ framework for structured argumentation [22]. More precisely, their goal is to define a preference ordering over arguments $\succsim$, based on the strict total order over defeasible rules defined to instantiate *ASPIC*$^+$ to PDL, so as to ensure that an extension within PDL corresponds to the justified conclusions of its *ASPIC*$^+$ instantiation. Several options are investigated, and they demonstrate that the standard *ASPIC*$^+$ *elitist* ordering cannot be used to calculate $\succsim$ as there is no correspondence between the argumentation-defined inferences and PDL, and the same holds for a disjoint elitist preference ordering. The authors come up with a new argument preference ordering definition which captures both preferences over arguments and also *when* defeasible rules become applicable in the arguments' construction, leading to the definition of a strict total order on defeasible rules and corresponding non-strict arguments. Their representation theorem shows that a

correspondence always exists between the inferences made in PDL and the conclusions of justified arguments in the *ASPIC*$^+$ instantiation under stable semantics.

Brewka and Eiter [7] consider programmes supplied with priority information, which is given by a supplementary strict partial ordering of the rules. This additional information is used to solve potential conflicts. Moreover, their idea is that conclusions should be only those literals that are contained in at least one answer set. They propose to use preferences on rules for selecting a subset of the answer sets called the *preferred answer sets*. In their approach, a rule is applied unless it is defeated via its assumptions by rules of higher priorities. Our definition (Definition 2.6) and the original formalism of Brewka and Eiter [7] are different, in the sense that in our definition we do not make use of default negation to represent the exceptions, i.e. the defeasibility of a (strict) rule. Rather, we use defeasible rules and the notion of the applicability of such rules. This means that the correct translation of the Order Puzzle of Example 2.2 ends up with the following logic programme[1]:

r_0 : w.
r_1 : h :- not ¬h, w.
r_2 : ¬o :- not o, w.
r_3 : o :- not ¬o, h.
$r_0 < r_3 < r_2 < r_1$.

If preferences are disregarded, then this logic programme has two answer sets: $\{w, h, ¬o\}$ and $\{w, h, o\}$. Thus, considering preferences, the latter is the unique preferred answer set. After dropping the context w from the answer set, we get an extension $\{h, o\}$, which is identical to the result obtained in Example 2.7.

Dung [11] presents an approach to deal with contradictory conclusions in defeasible reasoning with priorities. More precisely, he starts from the observation that often, the proposed approaches to defeasible reasoning with priorities (e.g. [5, 21, 27]) sanction contradictory conclusions, as exemplified by *ASPIC*$^+$ using the weakest link principle together with the elitist ordering which returns contradictory conclusions with respect to its other three attack relations, and the conclusions reached with the well-known approach of Brewka and Eiter [7]. Dung shows then that the semantics for any complex interpretation of default preferences can be characterized by a subset of the set of stable extensions with respect to the normal attack relation assignments, i.e. a normal form for ordinary attack relation assignments. In the setting of this paper, the notion of 'normal attack relation' could be defined as follows. Let $\alpha = (a_1, \ldots, a_n)$ and $\beta = (b_1, \ldots, b_m)$ be arguments constructed from a HANS. Since we have no Pollock style undercutting argument (as in *ASPIC*$^+$) and each norm is assumed to be defeasible, α is said to normally attack argument β if and only if β has a sub-argument β' such that $concl(\alpha) = \overline{concl(\beta')}$, and $r((a_{n-1}, a_n)) \geq r((b_{m-1}, b_m))$. According to Definitions 3.3 and 3.5, the normal defeat relation is equivalent to the defeat relation using the last link principle in this paper.

Kakas *et al.* [18] present a logic of arguments called *argumentation logic*, where the foundations of classical logical reasoning are represented from an argumentation perspective. More precisely, their goal is to integrate into the single argumentative representation framework both classical reasoning, as in propositional logic, and defeasible reasoning.

You *et al.* [31] define a prioritized argumentative characterization of non-monotonic reasoning, by casting default reasoning as a form of prioritized argumentation. They illustrate how the parameterized formulation of priority may be used to allow various extensions and modifications to default reasoning.

---

[1]Note that in Nrewka and Eiter's [7] original formulation $r_0 < r_3$ means that r_0 has higher priority than r_3.

We, and all these approaches, share the idea that an argumentative characterization of Non Monotonic Reasoning formalisms, like PDL in Young *et al.*'s [32] case and HANS in our approach, contributes to make the inference process more transparent to humans. However, the targeted Non Monotonic Reasoning formalism is different, leading to different challenges in the representation results. To the best of our knowledge, no other approach addressed the challenge of an argumentative characterization of prioritized normative reasoning.

The reason we study prioritized normative reasoning *in the setting of formal argumentation* is twofold. First, formal argumentation has been recognized as a popular research area in Artificial Intelligence, thanks to its ability to make the inference process more intuitive and provide natural explanations for the reasoning process [8]; its flexibility in dealing with the dynamics of the system; and its appeal in sometimes being more computationally efficient than competing approaches. Second, while some progress has been made on the use of priorities within argumentation (e.g. [2, 21]), how to represent different approaches for prioritized normative reasoning in argumentation is still a challenging issue.

8 Conclusions

In this paper we embedded three approaches to prioritized normative reasoning—namely the Greedy [32], Reduction [5] and Optimization [13] approaches—within the framework of a HANS. Within such a system, conditional norms are represented by a binary relation over literals, and priorities are represented by natural numbers. HANS provide an elegant visualization of a normative system, with conflicts shown as two paths to a proposition and its negation. Since both conflicts and exceptions can be encoded, such systems are inherently non-monotonic. In Dung [10], the author pointed out that 'many of the major approaches to non-monotonic reasoning in Artificial Intelligence and logic programming are different forms of argumentation', and inspired by this, we described how arguments can be instantiated as paths through a HANS; demonstrated that this instantiation satisfies certain desirable properties; and described how attacks and defeats between these arguments can be identified. Defeats in particular are dependent on the priorities associated with the arguments, and several different techniques have been proposed to lift priorities from argument components—made up of norms in the context of our work—to the arguments themselves [21]. We demonstrated that for a total ordering of priorities, lifting priorities to arguments based on the weakest link principle, evaluated using the stable semantics, is equivalent to Greedy; that lifting priorities to arguments based on last link and using the stable semantics is equivalent to Reduction; and that the Optimization approach can be encoded by an argumentation system which uses weakest link together with the stable semantics, and which introduces additional defeats capturing implicit conflicts between arguments.

This last result—which requires a relatively complex argumentative representation—opens up an interesting avenue for future work, namely in determining which non-monotonic logics can be easily captured through standard formal argumentation techniques, and which require additional rules or axioms in order to be represented. We note that on the argumentation side, work on bipolar argumentation (e.g. [9]) has considered introducing additional defeats between arguments based on some notion of support, and we intend to investigate how the additional defeats we introduced can be categorized in such frameworks.

Apart from our representation results, the use of argumentation allows us to make some useful observations, such as that Reduction will sometimes not reach any conclusions. Furthermore, argumentation can be used to provide explanation [8]. When implemented, a system building on our approach can help users understand what norms they should comply with, and why. For large

normative systems, the use of stable semantics to compute Reduction and Optimization results in a high computational overhead, while Greedy is computationally efficient. Ultimately, selecting the correct reasoning procedure thus requires giving consideration to both reasoning complexity, and the domain in which the system will be used.

In closing, our main observations can be summarized as follows. First, from a normative systems perspective, we know there are many logics of prioritized rules/norms, and we consider only three here. The choice we make (Greedy, Reduction and Optimization) may seem arbitrary. However, many other examples, in particular detachment procedures not satisfying defeasible deontic detachment, are much easer to characterize, while the three throughput variants of Greedy, Reduction and Optimization can be derived from the existing results. Furthermore, these three alternatives display quite diverse behaviour, and are illustrative of the various kind of approaches around.

Second, the results we present are interesting and promising, but the work on representing prioritized rules/norms using argumentation has only begun, and there are many open issues. In particular, the restriction to totally ordered systems must be relaxed in future work.

Third, given the large number of possibilities and the vast existing literature on normative rules/norms, a different methodology is needed for dealing with prioritized rules/norms in formal argumentation.

Finally, one may wonder why our results have not been shown before, given the long-standing discussion on weakest versus last link at least since the work of Pollock [25], and the central role of prioritized rules in many structured argumentation theories like ASPIC$^+$. The reason, we believe, is that it is easier to study these issues on a small fragment, like HANS, than on a very general theory like ASPIC$^+$.

Acknowledgements

References

[1] C. E. Alchourrón and D. Makinson. Hierarchies of regulations and their logic. In *New Studies in Deontic Logic*, R. Hilpinened., pp. 125–148. Springer, 1981.

[2] L. Amgoud and S. Vesic. Handling inconsistency with preference-based argumentation. In*Scalable Uncertainty Management—4th International Conference, SUM 2010, Toulouse, France, September 27–29, 2010. Proceedings, Lecture Notes in Computer Science*, A. Deshpande and A. Hunter, eds, **6379**, pp. 56–69. Springer, 2010.

[3] G. Boella and L. W. N. van der Torre. Permissions and obligations in hierarchical normative systems. In *Proceedings of the 9th International Conference on Artificial Intelligence and Law, ICAIL 2003 Edinburgh, Scotland, UK, June 24–28, 2003*, pp. 109–118. ACM, 2003.

[4] G. Boella, L. W. N. van der Torre and H. Verhagen. Ten challenges for normative multiagent systems. In *Programming Multi-Agent Systems, 31.08. - 05.09.2008*, Dagstuhl, 2008.

[5] G. Brewka. Preferred subtheories: An extended logical framework for default reasoning. In

Proceedings of the 11th International Joint Conference on Artificial Intelligence. Detroit, MI, USA, August 1989, N. S. Sridharaned., pp. 1043–1048. AAAI, 1989.

[6] G. Brewka. Reasoning about priorities in default logic. In *Proceedings of the 12th National Conference on Artificial Intelligence, Seattle, WA, USA, July 31 - August 4, 1994*, Volume **2**, B. Hayes-Roth and R. E. Korf, eds, pp. 940–945. AAAI, 1994.

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

[8] M. Caminada, R. Kutlak, N. Oren and W. W. Vasconcelos. Scrutable plan enactment via argumentation and natural language generation. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, pp. 1625–1626. IFAAMAS, 2014.

[9] C. Cayrol and M.-C. Lagasquie-Schiex. An axiomatic approach to support in argumentation, In *Theory and Applications of Formal Argumentation: Third International Workshop*. E. Black, S. Modgil and N. Oren, eds, pp. 74–91. Springer, 2015.

[10] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, **77**, 321–358, 1995.

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

[12] J. Hansen. Deontic logics for prioritized imperatives. *Artificial Intelligence & Law*, **14**, 1–34, 2006.

[13] J. Hansen. Prioritized conditional imperatives: problems and a new proposal. *Autonomous Agents and Multi-Agent Systems*, **17**, 11–35, 2006.

[14] J. Hansen. Reasoning about permission and obligation. In *David Makinson on Classical Methods for Non-Classical Problems*, S. O. Hanssoned., pp. 287–333. Springer, 2014.

[15] S. O. Hansson. The varieties of permission. In *Handbook of Deontic Logic and Normative Systems*, D. Gabbay, J. Horty, X. Parent, R. van der Meyden and L. van der Torre, eds, pp. 195–240. College Publications, 2013.

[16] R. Hilpinen and P. Mcnamara. Deontic logic: A historical survey and introduction. In *Handbook of Deontic Logic and Normative Systems*, D. Gabbay, J. Horty, X. Parent, R. van der Meyden and L. van der Torre, eds, pp. 3–136. College Publications, 2013.

[17] J. Horty. Defaults with priorities, *Journal of Philosophical Logic*, **36**, 367–413, 2007.

[18] A. C. Kakas, F. Toni and P. Mancarella. Argumentation for propositional logic and nonmonotonic reasoning. In *Proceedings of the 29th Italian Conference on Computational Logic, Torino, Italy, June 16–18, 2014. Vol. 1195 of CEUR Workshop Proceedings*, L. Giordano, V. Gliozzi and G. L. Pozzato, eds, pp. 272–286, 2014.

[19] B. Liao, N. Oren, L. van der Torre and S. Villata. Prioritized norms and defaults in formal argumentation. In *Proceedings of the 13th International Conference on Deontic Logic and Normative Systems (DEON2016)*, pp. 139–154. College Publications, 2016.

[20] D. Makinson and L. W. N. van der Torre. Input/output logics. *J. Philosophical Logic*, **29**, 383–408, 2000.

[21] S. Modgil and H. Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, **195**, 361–397, 2013.

[22] S. Modgil and H. Prakken. The ASPIC$^+$ framework for structured argumentation: A tutorial. *Argument & Computation*, **5**, 31–62, 2014.

[23] D. Nute, ed. *Defeasible Deontic Logic*, vol. 263, Springer Science & Business Media, 1997.

[24] X. Parent and L. van der Torre. Input/output logic. In *Handbook of Deontic Logic and Normative Systems*, D. Gabbay, J. Horty, X. Parent, R. van der Meyden and L. van der Torre, eds, pp. 499–544. College Publications, 2013.

[25] J. L. Pollock. *Cognitive Carpentry*. Bradford/MIT Press, 1995.

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

[27] T. Schaub and K Wang. A comparative study of logic programs with preference. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI 2001, Seattle, Washington, USA, August 4–10, 2001*, B. Nebeled., pp. 597–602. AAAI, 2001.

[28] S. C. Tosatto, G. Boella, L. W. N. van der Torre and S. Villata. Abstract normative systems: Semantics and proof theory. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Thirteenth International Conference, KR 2012, Rome, Italy, June 10–14, 2012*, G. Brewka, T. Eiter and S. A. McIlraith, eds, pp. 358–368, 2012.

[29] S. C. Tosatto, G. Boella, L. W. N. van der Torre and S. Villata. Visualizing normative systems: An abstract approach. In *Deontic Logic in Computer Science—11th International Conference, DEON 2012, Bergen, Norway, July 16–18, 2012. Proceedings*, pp. 16–30. Springer, 2012.

[30] L. W. N. van der Torre and S. Villata. An ASPIC-based legal argumentation framework for deontic reasoning. In *Computational Models of Argument—Proceedings of COMMA 2014, Atholl Palace Hotel, Scottish Highlands, UK, September 9–12, 2014. Vol. 266 of Frontiers in Artificial Intelligence and Applications*, S. Parsons, N. Oren, C. Reed and F. Cerutti, eds, pp. 421–432, 2014.

[31] J. You, X. Wang and L. Yuan. Nonmonotonic reasoning as prioritized argumentation. *IEEE Transactions on Knowledge and Data Engineering*, **13**, 968–979, 2001.

[32] A. P. Young, S. Modgil and O. Rodrigues. Prioritised default logic as rational argumentation. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, pp. 626–634. IFAAMAS, 2016.