Sequent Rules for Reasoning and Conflict Resolution in Conditional Norms

Agata Ciabattoni Björn Lellmann

TU Wien, Vienna, Austria {agata, lellmann}@logic.at

Abstract

We introduce a sequent based method for reasoning with deontic assumptions using specificity and superiority for conflict resolution. Starting from a base logic, we apply strengthening of the antecedent to the assumptions wherever possible unless this would yield an inconsistency. The method applies to logics with an arbitrary finite number of dyadic deontic operators of type MP or MD (in the sense of Chellas) with inclusions among the operators. We illustrate the method using various examples. An implementation is also available.

1 Introduction

Legal, ethical, religious and behavioral norms often have a conditional form. A common way for formalising such conditional norms is via dyadic deontic operators, historically introduced to represent Contrary-To-Duty (CTD) obligations, i.e. obligations which are applicable only if another norm is violated. Although the dyadic representation can solve notorious CTD paradoxes, it also introduces new difficulties; in particular, how to reason on the conditions (i.e., the second argument of dyadic operators) without reintroducing possible deontic conflicts. Roughly speaking, a deontic conflict occurs when two or more obligations/prohibitions cannot be mutually realized.

Various general conflict resolution principles are considered in the literature. Here we focus on two major ones, widely used in law and AI: specificity and superiority. The former, known in law as lex specialis derogat legi generali, states that specific obligations/prohibitions override more general ones, while the latter refers to prioritized obligations/prohibitions coming from normative authorities of different strength (lex superior) or, e.g., in a different chronological order (lex posterior).

In this article we extend the most basic dyadic deontic logics with a general and purely syntactic mechanism for reasoning on the conditions of deontic

 $^{^\}star$ This work was partly supported by WWTF Project MA16-28 and by BRISE-Vienna (UIA04-081), a European Union Urban Innovative Actions project.

assumptions, resolving conflicts using specificity on conditions and superiority between assumptions. The mechanism generalizes and extends to superiority the calculus introduced in [8] for a particular logic (see also Sec. 6.1).

Our starting point are logics based on finite combinations of operators \heartsuit , which are dyadic versions of non-normal (upwards or downwards) monotone modal logic M, extended with the (dyadic versions of) the axioms $P \neg (\heartsuit(\bot/B) \land \heartsuit(\top/B))$ or $D \neg (\heartsuit(A/B) \land \heartsuit(\neg A/B))$. For an upwards monotone, i.e., obligation type, operator this yields, e.g., the dyadic version of minimal deontic logic MP from [6]. Although well behaved, these logics are not useful for reasoning on the conditions of deontic formulae. E.g, for a downwards monotone, i.e., prohibition type, operator \mathcal{F} we can derive $\mathcal{F}(\mathsf{park}/\top) \to \mathcal{F}(\mathsf{park} \land \mathsf{ride}/\top), \text{ but not } \mathcal{F}(\mathsf{park}/\top) \to \mathcal{F}(\mathsf{park}/\neg\mathsf{permit}).$ The naive solution of adding unrestricted strengthening of the antecedent, i.e., an unrestricted downwards monotonicity rule for the second argument, quickly leads to conflicting norms, and in presence of axiom D to a contradiction. To avoid this, we consider sequent rules incorporating a limited form of strengthening of the antecedent / downwards monotonicity "up to conflicting assumptions". Starting from prima-facie deontic assumptions and propositional background facts, our sequent rules intuitively permit to derive every formula resulting from strengthening the antecedent, unless this would lead to an inconsistency over the base logic. Deontic conflicts are resolved using specificity and superiority. The resulting system satisfies the disjunctive response of [10], see Ex. 4.3, and can be used to model permissions as exceptions as well as some forms of CTD reasoning, see Ex. 4.2, Sect. 6.2 and Rem. 6.3.

As in sequent calculi for non-monotonic logics [3,25], our rules use statements expressing that certain sequents are not derivable. In contrast with other calculi for non-monotonicity in normative reasoning like [13,29], our calculi enjoy cut-elimination, which yields decidability and complexity results. A further corollary is that we can define the set of consequences of deontic assumptions iteratively, thus avoiding fixed-point constructions like those in [17].

The generality of our system is demonstrated with case studies including the logic simulating the reasoning of the Mīmāṃsā school from [7,8], a modelling of permissions as exceptions, and the operators of sanction and violation.

The system is implemented in the Prolog system deonticProver2.0 (http://subsell.logic.at/bprover/deonticProver/version2.0/). For any finite set of dyadic operators of type M, MP, or MD, with (possible) inclusions, the system constructs sequent rules to deal with specificity and superiority, and uses them to answer the question: Given an input of deontic assumptions and background facts, which conditional norms are in force, i.e., which formulae are derivable? In addition to a web interface for the prover, the website contains a number of examples and illustrates the behaviour of the system with respect to some standard deontic puzzles and paradoxa.

2 Restricted strengthening of the antecedent

Before delving into the technicalities, we briefly illustrate the intuitions behind our approach. As mentioned above, given deontic assumptions such as

- (i) You ought not to eat with your fingers,
- (ii) You ought to put your napkin on your lap,
- (iii) If you are served asparagus, you ought to eat it with your fingers,

from the asparagus example (e.g. [31,16]) we would like to be able to apply strengthening of the antecedent to (ii) to derive "If you are served asparagus, you ought to put your napkin on your lap". However, as is well-known, adopting an unrestricted form of strengthening of the antecedent would also yield "If you are served asparagus, you ought not to eat with your fingers". Together with (iii) this yields a pair of conflicting obligations, and hence an inconsistency in any logic satisfying the D-axiom for obligations.

Our proposal for dealing with this situation is based on two main aspects: First, it is parametric in the base logic, and second it follows what could be called a generous approach towards applying strengthening of the antecedent. The latter means that given a set of deontic assumptions we apply strengthening of the antecedent whenever this is possible without resulting in inconsistencies over the base logic. In particular, this aims at keeping in force as many prima-facie norms as possible. Conflicts between norms are resolved following the specificity principle, i.e., assuming that conditional norms with more specific conditions like (iii) above overrule those with more general conditions like (i), and an (optional) superiority relation on the deontic assumptions. Note that inconsistencies are always evaluated with respect to the base logic. Hence for logics containing no principles ruling out conflicting or impossible norms there are no conflicts to avoid, and we obtain unrestricted strengthening of the antecedent/downwards monotonicity in the second argument.

In the asparagus example above, given a base logic ruling out conflicting obligations, we thus should derive "If you are served asparagus, you ought to put your napkin on your lap" as well as, e.g., "If you are served asparagus at your grandparents', you ought to eat it with your fingers": For the former, there are no assumptions which could yield a conflict; for the latter, the assumption (i) could be used to derive a conflicting obligation, but this assumption is overruled by the more specific assumption (iii).

The situation becomes more interesting if we consider the following additional deontic assumption: $^{\rm 1}$

(iv) If you are at your grandparents', you ought not to eat with your fingers. Now neither of the two assumptions (iii) and (iv) is more specific than the other. Hence, in order to keep the derived obligations consistent over the base logic we cannot derive the obligation "If you are served asparagus at your

 $^{^{1}}$ We are grateful to the anonymous reviewer for bringing this and the following examples to our attention.

grandparents', you ought to eat it with your fingers" anymore, because then by symmetry we should also be able to derive the conflicting obligation "If you are served asparagus at your grandparents', you ought not to eat it with your fingers" from assumption (iv). Note that this shows the difference to credulous approaches, where both the above statements would be derivable. The situation changes again, however, if we add the permission (see Sect. 6.2):

(v) If you are served asparagus at your grandparents', you may eat it with your fingers.

Assuming that the base logic contains the principle that there are no conflicting pairs of obligations and permissions, this assumption would prevent the derivation of the obligation "If you are served asparagus at your grandparents', you ought not to eat it with your fingers" from assumption (iv), since it is more specific. But then we cannot derive any obligation which would conflict with "If you are served asparagus at your grandparents', you ought to eat it with your fingers". Hence, following the generous approach towards applying strengthening of the antecedent, we don't have any reason to refrain from deriving this obligation from assumption (iii). While it has been argued, e.g., in [29] that it might be undesired if more specific permissions reinstate less specific obligations, this is in line with the idea of preventing the derivation of only those obligations which would result in inconsistencies over the base logic.

The generous approach of deriving every obligation which would not result in an inconsistency over the base logic further motivates the idea that the notion of a conflicting assumption is evaluated with respect to the obligation we want to derive, and not the assumption we want to derive it from. As an example, consider the additional assumption:

(vi) You ought not to eat with your fingers and not to pick your nose.

While the obligation "If you are served asparagus, you ought not to eat it with your fingers and not to pick your nose" is in conflict with the more specific assumption (iii) and hence should not be derivable, the obligation "If you are served asparagus, you ought not to pick your nose" is not. Thus, we don't refrain from deriving the latter, even though the content of the assumption we derived it from is inconsistent with the content of the more specific applicable assumption (iii).

This focus on what we want to derive instead of the assumptions we derive it from has the additional benefit that we do not need to worry about chains of more and more specific assumptions, each in conflict with the previous one: Given that the set of deontic assumptions is finite, such a chain will contain a most specific applicable assumption. To see whether we should refrain from deriving an obligation which would follow from one of the more general ones, we thus only need to check the most specific assumption which is in conflict with what we want to derive. If this one is overruled by an even more specific assumption, then we can use the latter to derive the obligation in question; otherwise we refrain from doing so.

We would like to stress again that the approach is parametric in the base

logic. Hence the resulting systems inherit some of the limitations imposed by the latter, both in terms of what is removed as inconsistent and of what can be derived from the assumptions. In this work we consider only relatively weak base logics. In particular, they neither permit to aggregate obligations, nor rule out conflicts between more than two obligations, where each pair of these is nonconflicting. We are, however, confident that the general method can be extended to stronger base logics as well (see Sec. 7). Note also that since we only aim to consistently close a set of conditional deontic assumptions under strengthening of the antecedent with respect to a base logic, the derivable formulae are still conditional statements, and hence we do not incorporate factual detachment principles.

3 The base system

Formally, the basic logical systems we consider are propositional deontic logics. Our logics extend the language of classical propositional logic consisting of variables (p, q, ...), falsum (\bot) and implication (\to) , with dyadic deontic operators $\heartsuit(./.)$ where the first argument represents the content of a conditional norm, while the second argument represents its condition. We distinguish two kinds of operators, depending on what it takes to comply with the norm:

- An operator \heartsuit is of *obligation-type* if the norm $\heartsuit(A/B)$ is complied with whenever A is true;
- An operator \heartsuit is of *prohibition-type* if the norm $\heartsuit(A/B)$ is complied with whenever A is false.

Note that this makes our operators upwards monotone in the first argument for obligation type operators, and downwards monotone for prohibition type ones. To capture relations between operators and their properties, given a set Op of deontic operators with associated types, we assume a reflexive and transitive inclusion relation \to , a symmetric conflict relation $\not \in$, and a unary nontriviality predicate nt with the following intended meaning:

- If $\heartsuit \to \spadesuit$ for two operators $\heartsuit, \spadesuit \in \mathsf{Op}$ of the same type, then complying with $\heartsuit(A/B)$ implies complying with $\spadesuit(A/B)$.
- If $\heartsuit_{\ell} \spadesuit$ for two operators of the same type $\heartsuit, \spadesuit \in \mathsf{Op}$, then complying with one of $\heartsuit(A/B), \spadesuit(\neg A/B)$ entails violating the other.
- If $\mathsf{nt}(\heartsuit)$ for an operator $\heartsuit \in \mathsf{Op}$, then \heartsuit is *non-trivial*, in that it is logically possible to comply with it.

For operators \heartsuit, \spadesuit of different type we flip the polarity of A in one of the assumptions, i.e., we replace $\spadesuit(A/B)$ with $\spadesuit(\neg A/B)$ and vice versa. We assume that the relations \nleq and nt are closed under preimages of the implication relation, i.e., if $\heartsuit \not \updownarrow \spadesuit$ and $\diamondsuit \to \heartsuit$, then also $\diamondsuit \not \updownarrow \spadesuit$. In the following, an *operator characterisation* is a tuple $\mathfrak{O} = (\mathsf{Op}, \to, \not \downarrow, \mathsf{nt})$ consisting of a set Op of operators with types together with inclusion, conflict, and non-triviality relations.

The base logic we will consider then contains the Hilbert-style rules and axioms of propositional classical logic together with the rules and axioms in

```
 \begin{cases} \frac{A \to C}{\heartsuit(A/B) \to \heartsuit(C/B)} : \heartsuit \text{ obl type} \end{cases} \quad \begin{cases} \frac{C \to A}{\heartsuit(A/B) \to \heartsuit(C/B)} : \heartsuit \text{ proh type} \end{cases} 
 \{ \heartsuit(A/B) \to \spadesuit(A/B) : \heartsuit \to \spadesuit, \text{ same type } \} 
 \cup \quad \{ \heartsuit(A/B) \to \spadesuit(\neg A/B) : \heartsuit \to \spadesuit, \text{ different type } \} 
 \cup \quad \{ \neg(\heartsuit(A/B) \land \spadesuit(\neg A/B)) : \heartsuit \not \downarrow \spadesuit, \text{ same type } \} 
 \cup \quad \{ \neg(\heartsuit(A/B) \land \spadesuit(A/B)) : \heartsuit \not \downarrow \spadesuit, \text{ different type } \} 
 \cup \quad \{ \neg(\heartsuit(\bot/B) \land \heartsuit(\top/B)) : \text{nt}(\heartsuit) \}
```

Fig. 1. The deontic axioms and rules for $\mathfrak{O} = (\mathsf{Op}, \to, \cancel{\xi}, \mathsf{nt})$.

Fig. 1. Note that due to upwards and downwards monotonicity respectively, for operators \heartsuit with $\mathsf{nt}(\heartsuit)$ the axiom $\neg(\heartsuit(\bot/B) \land \heartsuit(\top/B))$ is equivalent to $\neg \heartsuit(\bot/B)$ for \heartsuit of obligation type and to $\neg \heartsuit(\top/B)$ for \heartsuit of prohibition type.

Example 3.1 (i) Setting $Op = \{\mathcal{O}\}$ with \mathcal{O} of obligation type and $nt(\mathcal{O})$ yields the dyadic version of *minimal deontic logic* MP from [6].

- (ii) Replacing $\mathsf{nt}(\mathcal{O})$ with $\mathcal{O} \not\downarrow \mathcal{O}$ in (i) yields the dyadic version of monotone modal logic M extended with the D axiom $\neg(\mathcal{O}(A/B) \land \mathcal{O}(\neg A/B))$.
- (iii) Setting $\operatorname{Op} = \{\mathcal{O}, \mathcal{F}\}$ with \mathcal{O} of obligation type, \mathcal{F} of prohibition type, and $\mathcal{O} \not\leftarrow \mathcal{F}$ yields a logic with upwards monotone obligations \mathcal{O} , downwards monotone prohibitions \mathcal{F} , and no conflicts between obligations and prohibitions, i.e., the axiom $\neg(\mathcal{O}(A/B) \land \mathcal{F}(A/B))$. Note that this does not rule out conflicts between obligations or between prohibitions. This could be added by stipulating $\mathcal{O} \not\leftarrow \mathcal{O}$ and $\mathcal{F} \not\leftarrow \mathcal{F}$, respectively.
- (iv) Let $Op = \{must, ought, should\}$ with all operators of obligation type. Setting $must \to ought$, $must \to should$, $ought \to should$ with $must \not = and nt(ought)$ illustrates the possibility of using different operators for analysing different strengths of obligations. The intuition is that must behaves like an obligation, while the weaker ought behaves more like a recommendation, hence satisfies only the P axiom instead of D. See, e.g., [1].

To facilitate automated reasoning and prove useful meta-logical properties, we switch from Hilbert-style calculi to sequent calculi. As usual, a sequent is a tuple $\Gamma \Rightarrow \Delta$ of multisets of formulae, with formula interpretation $\bigwedge \Gamma \to \bigvee \Delta$, see, e.g., [30]. To write the rules with a coincise notation we introduce the following two abbreviations:

$$\mathsf{Impl}_{\heartsuit, \spadesuit}(A, B) := \begin{cases} A \Rightarrow B & \heartsuit, \spadesuit \text{ obligation type} \\ A, B \Rightarrow & \heartsuit \text{ obligation type }, \spadesuit \text{ prohibition type} \\ B \Rightarrow A & \heartsuit, \spadesuit \text{ prohibition type} \\ \Rightarrow A, B & \heartsuit \text{ prohibition type }, \spadesuit \text{ obligation type} \end{cases}$$

Fig. 2. The base calculus for a given operator characterisation $\mathfrak{O} = (\mathsf{Op}, \to, \cancel{\xi}, \mathsf{nt})$

$$\mathsf{Confl}_{\heartsuit, \spadesuit}(A, B) := \left\{ \begin{array}{l} A, B \Rightarrow \ \heartsuit, \spadesuit \ \mathrm{obligation} \ \mathrm{type} \\ A \Rightarrow B \ \heartsuit \ \mathrm{obligation} \ \mathrm{type} \\ \Rightarrow A, B \ \heartsuit, \spadesuit \ \mathrm{prohibition} \ \mathrm{type} \\ B \Rightarrow A \ \heartsuit \ \mathrm{prohibition} \ \mathrm{type} \end{array} \right.$$

The intuition is that, e.g., for two operators \heartsuit, \spadesuit of obligation type, complying with $\heartsuit(A/C)$ means that A is true, whereas violating $\spadesuit(B/C)$ means that B is false. Hence complying with $\heartsuit(A/C)$ implies violating $\spadesuit(B/C)$ if A implies $\neg B$. This is captured in $\mathsf{Confl}_{\heartsuit, \spadesuit}(A, B)$, i.e., the sequent $A, B \Rightarrow .$ Using these abbreviations, converting the Hilbert-style axioms into sequent rules using the general method from [19] then gives the deontic rules $\mathsf{Mon}_{\heartsuit, \spadesuit}, \mathsf{D}_{\heartsuit, \spadesuit}, \mathsf{P}_{\heartsuit}$ of the base calculus in Fig. 2. Note that since the relation \to is reflexive, we have for every operator \heartsuit either the upwards or downwards monotonicity rule:

$$\frac{A\Rightarrow C\quad B\Rightarrow D\quad D\Rightarrow B}{\Gamma, \heartsuit(A/B)\Rightarrow \heartsuit(C/D), \Delta} \ \operatorname{Mon}\, \uparrow \qquad \frac{C\Rightarrow A\quad B\Rightarrow D\quad D\Rightarrow B}{\Gamma, \heartsuit(A/B)\Rightarrow \heartsuit(C/D), \Delta} \ \operatorname{Mon}\, \downarrow$$

The resulting calculi are equivalent and admit cut-elimination, see [19].

4 Reasoning from assumptions

To reason on the conditions of norms in the above systems we will introduce special sequent rules. These allow us to reason from deontic assumptions, i.e., a finite set $\mathfrak L$ of deontic formulae $\mathfrak D(A/B)$ with $\mathfrak D\in \mathsf{Op}$. To ensure well-definedness and termination of proof search we require that these assumptions are non-nested, i.e., that A,B are purely propositional. The main idea is to make the second argument downwards monotone "up to conflicting assumptions". Since a formula A with $A\to B$ can be seen as "more specific" than B, this captures the specificity principle, that more specific conflicting deontic assumptions overrule less specific ones. Before considering the rules in detail we mention two more features of the system.

To be able to reason with non-deontic propositions as well, we also consider propositional facts as assumptions. W.l.o.g. we assume that these are given in the form of a finite set \mathfrak{F} of atomic sequents, i.e., sequents of the form

 $p_1, \ldots, p_n \Rightarrow q_1, \ldots, q_m$ with the p_i, q_j propositional variables. Since every purely propositional formula is equivalent to a formula in conjunctive normal form, this is equivalent to permitting arbitrary purely propositional formulae as assumptions. The sequent rules for these assumptions then are given by

$$\left\{ \; \overline{\Sigma,\Gamma\Rightarrow\Delta,\Pi} \;\; \mathfrak{F} : \Gamma\Rightarrow\Delta\in\mathfrak{F} \right\} \;.$$

Often obligations and prohibitions further come with a priority order. To capture this, we follow the standard approach and say that a *superiority relation* is a binary relation \succ on the set of deontic assumptions. The intuition is that for two deontic assumptions A,B with $A\succ B$, the former is superior, or has higher authority, than the latter, and hence A cannot be overruled by B, even if the latter is more specific. For technical reasons we impose that for every two assumptions A,B we have $A\not\succeq B$ or $B\not\succeq A$. Note that this rules out cycles of length one or two, but due to lack of transitivity not those of greater length.

Sequent calculus rules: We then extend the base calculus with sequent rules for capturing the specificity principle in presence of prioritized deontic assumptions. The idea is that we can use downwards monotonicity in the second argument to derive, e.g., $\heartsuit(A/B)$ from a deontic assumption $\heartsuit(A/B \lor C)$ unless the latter is overruled by a more specific conflicting deontic assumption or in conflict with the P axiom. In addition, we also need to rule out that there is another conflicting assumption which is not overruled by a more specific one. The crucial feature needed for this is the addition of underivability statements in the premisses of the rules. These are used for stating, e.g., that we cannot derive a conflict between two formulae. The general conditions for deriving an obligation or a prohibition from a deontic assumption then are as follows:

Given a list $\mathfrak L$ of deontic assumptions, we can derive $\mathfrak O(A/B)$ from the assumption $\spadesuit(C/D) \in \mathfrak L$ with $\spadesuit \to \mathfrak O$ if:

- the assumption $\spadesuit(C/D)$ is applicable, i.e., if we can derive that the condition B implies the condition D; AND
- complying with the assumption $\spadesuit(C/D)$ implies complying with $\heartsuit(A/B)$, i.e., if we can derive $\mathsf{Impl}_{\spadesuit, \heartsuit}(C, A)$; AND
- there is no conflict with the non-triviality axiom P for \heartsuit , i.e., we cannot derive $\mathsf{Confl}_{\heartsuit,\heartsuit}(A,A)$ provided that $\mathsf{nt}(\heartsuit)$; AND
- the assumption $\spadesuit(C/D)$ is neither overruled by a more specific one, nor in conflict with another assumption which is not overruled. I.e., for every assumption $\clubsuit(E/F) \in \mathfrak{L}$ with $\clubsuit \not \downarrow \spadesuit$ and $\spadesuit(C/D) \not \succ \clubsuit(E/F)$ we have:
 - · the assumption $\clubsuit(E/F)$ is not applicable, i.e., we cannot derive that the condition B implies F; OR
 - the assumption $\clubsuit(E/F)$ is not in conflict with what we want to derive, i.e., we cannot derive Confl $\blacktriangle_{\infty}(E,A)$; OR
 - the assumption $\clubsuit(E/F)$ is not more specific than $\spadesuit(C/D)$ and it is not overruled by another more specific one, i.e.:
 - * the assumption $\clubsuit(E/F)$ is not more specific than $\spadesuit(C/D)$, i.e., we

```
B\Rightarrow D \\ \begin{aligned} & \underset{\mathbb{R}}{\operatorname{Impl}}_{\boldsymbol{A}, \boldsymbol{\heartsuit}}(C, A) \\ & \{ \boldsymbol{\nvdash} \operatorname{Confl}_{\boldsymbol{A}, \boldsymbol{\diamondsuit}}(A, A) : \operatorname{nt}(\boldsymbol{\clubsuit}), \boldsymbol{\clubsuit} = \boldsymbol{\heartsuit} \} \\ & \left\{ \begin{bmatrix} \boldsymbol{\nvdash} B \Rightarrow F \\ \boldsymbol{\nvdash} \operatorname{Confl}_{\boldsymbol{A}, \boldsymbol{\heartsuit}}(E, A) \\ B \Rightarrow Y, \\ Y \Rightarrow F, \\ \operatorname{Impl}_{\boldsymbol{\diamondsuit}, \boldsymbol{\heartsuit}}(X, A) \end{bmatrix} : & \boldsymbol{\diamondsuit}(X/Y) \in \mathfrak{L}, \\ \boldsymbol{\&}(E/F) \neq \boldsymbol{\diamondsuit}(X/Y) \end{bmatrix} : & \boldsymbol{\clubsuit}(E/F) \in \mathfrak{L}, \\ \boldsymbol{\clubsuit}(C/D) \neq \boldsymbol{\clubsuit}(E/F) \\ \boldsymbol{\clubsuit}(C/D) \neq \boldsymbol{\clubsuit}(E/F) \\ \boldsymbol{\clubsuit}(C/D) \neq \boldsymbol{\clubsuit}(C/D) \\ \boldsymbol{\clubsuit} \Rightarrow D \\ & \operatorname{Confl}_{\boldsymbol{\diamondsuit}, \boldsymbol{\heartsuit}}(C, A) \\ \left\{ \begin{bmatrix} \boldsymbol{\nvdash} B \Rightarrow F \\ \boldsymbol{\nvdash} \operatorname{Impl}_{\boldsymbol{\diamondsuit}, \boldsymbol{\heartsuit}}(E, A) \\ B \Rightarrow Y, \\ \boldsymbol{\vee} B \Rightarrow F, \\ \boldsymbol{\nvdash} \operatorname{Impl}_{\boldsymbol{\diamondsuit}, \boldsymbol{\heartsuit}}(E, A) \\ \boldsymbol{\clubsuit}(E/F) \neq \boldsymbol{\diamondsuit}(X/Y) \end{bmatrix} : & \boldsymbol{\clubsuit}(E/F) \in \mathfrak{L}, \\ \boldsymbol{\clubsuit}(E/F) \in \mathfrak{L}, \\ \boldsymbol{\clubsuit}(E/F) \neq \boldsymbol{\diamondsuit}(E/F) \\ \boldsymbol{\clubsuit}(E/F) \\ \boldsymbol{\clubsuit}(E/F) \neq \boldsymbol{\diamondsuit}(E/F) \\ \boldsymbol{\clubsuit}(E/F) \\ \boldsymbol{\clubsuit}(E/F)
```

Fig. 3. The deontic assumption rules.

cannot derive that the condition F implies D; AND

- * there is another more specific applicable assumption $\Diamond(X/Y)$, complying with which implies complying with $\heartsuit(A/B)$, i.e., for one of $\Diamond(X/Y) \in \mathfrak{L}$ with $\Diamond_{\ell} \clubsuit$ and $\clubsuit(E/F) \not\succ \Diamond(X/Y)$ we have:
- the assumption $\Diamond(X/Y)$ applies, i.e., we can derive that the condition B implies Y; AND
- the condition Y is more specific than the condition of $\clubsuit(E/F)$, i.e., we can derive that the condition Y implies F; AND
- complying with the assumption $\Diamond(X/Y)$ implies complying with $\heartsuit(A/B)$, i.e., we can derive $\mathsf{Impl}_{\Diamond, \heartsuit}(X, A)$.

In order to formalise this as sequent rules we use the following abbreviation. Let $\mathfrak{S} = \{S_1, \dots, S_n\}$ be a finite set of sets of premisses. Then we write

$$\frac{\mathcal{P} \cup [\mathfrak{S}]}{\mathcal{C}} \quad \text{for the set of rules} \quad \left\{ \frac{\mathcal{P} \cup \mathcal{S}_1}{\mathcal{C}}, \quad \dots, \quad \frac{\mathcal{P} \cup \mathcal{S}_n}{\mathcal{C}} \right\} \; .$$

The general assumption right rules $\heartsuit_R^{\spadesuit(C/D)}$ are given in Fig. 3, where we write $\nvdash \Gamma \Rightarrow \Delta$ for an underivability statement. Note that in this notation sets essentially correspond to conjunctive conditions on the premisses and capture the "AND" and "for all" above, while the choice notation [.] essentially corresponds to disjunctive conditions and captures the "OR" and "there is". In particular, the notation $[S_{\diamondsuit(X/Y)}: \diamondsuit(X/Y) \in \mathfrak{L}]$ corresponds to the big disjunction over the $\diamondsuit(X/Y) \in \mathfrak{L}$ of the $S_{\diamondsuit(X/Y)}$ and hence the existential quantification over

the finite set \mathfrak{L} . To abbreviate the notation we equivalently incorporated the premiss $\nvdash F \Rightarrow D$ into the following choice block.

Remark 4.1 The D axiom is equivalent to $\heartsuit(A/B) \to \neg \heartsuit(\neg A/B)$, and hence from an assumption $\heartsuit(A/B)$ we should be able to derive $\neg \heartsuit(\neg A/B)$. The assumption right rules allow us to do that *only if* we use the cut rule, see Sec. 5. As the presence of this rule destroys useful properties of the calculus, we introduce in the system the corresponding left rules $\heartsuit_L^{\spadesuit(C/D)}$ in Fig. 3, obtained by absorbing cuts between the assumption right rules $\heartsuit_R^{\spadesuit(C/D)}$ and the D-rules $D_{\heartsuit, \clubsuit}$. As usual, introducing a formula $\heartsuit(A/B)$ on the left hand side of the sequent, amounts to deriving $\neg \heartsuit(A/B)$.

Note that the nonderivability premiss for removing conflicts with the P axiom is no longer present – if $\mathsf{nt}(\heartsuit)$ and $\mathsf{Confl}_{\heartsuit,\heartsuit}(A,A)$ is derivable, then we immediately obtain the conclusion using the rule P_{\heartsuit} . The full calculus then contains the base rules of Fig. 2 together with the rules:

$$\left\{ \heartsuit_R^{\spadesuit(C/D)} \; : \; \spadesuit \to \heartsuit, \spadesuit(C/D) \in \mathfrak{L} \right\} \; \cup \; \left\{ \heartsuit_L^{\spadesuit(C/D)} \; : \; \spadesuit \not \downarrow \heartsuit, \spadesuit(C/D) \in \mathfrak{L} \right\} \; .$$

4.1 Examples

The examples below can be checked at http://subsell.logic.at/bprover/deonticProver/version2.0/, where also more examples are available.

Example 4.2 Continuing Ex. 3.1.(ii), consider \mathcal{O} of obligation type with $\mathcal{O}_{\xi}\mathcal{O}$ and the deontic assumptions corresponding to the asparagus example [31,16] (see also Sec. 2) given by $\mathfrak{L} = \{\mathcal{O}(\neg \mathsf{fingers}/\top), \mathcal{O}(\mathsf{fingers}/\mathsf{asparagus}), \mathcal{O}(\neg \mathsf{asparagus}/\top)\}$. Since asparagus $\to \top$ and there is no conflicting assumption, we can derive $\mathcal{O}(\neg \mathsf{asparagus}/\mathsf{asparagus})$, hence the contrary-to-duty obligation $\mathcal{O}(\mathsf{fingers}/\mathsf{asparagus})$ does not override the primary obligation $\mathcal{O}(\neg \mathsf{asparagus}/\top)$. However, the more specific obligation (or exception) $\mathcal{O}(\mathsf{fingers}/\mathsf{asparagus})$ overrides $\mathcal{O}(\neg \mathsf{fingers}/\top)$. Moreover, exemplifying Rem. 4.1, since we can derive $\mathcal{O}(\mathsf{fingers}/\mathsf{asparagus})$, due to $\mathcal{O}_{\xi}\mathcal{O}$ and the assumption left rule $\mathcal{O}_L^{\mathcal{O}(\mathsf{fingers}/\mathsf{asparagus})}$ we also derive $\neg \mathcal{O}(\neg \mathsf{fingers}/\mathsf{asparagus})$.

Example 4.3 Consider the classical drowning twins example, for the same operator \mathcal{O} as in the previous example, deontic assumptions $\mathfrak{L} = \{\mathcal{O}(\mathsf{save_twin_1/\top}), \ \mathcal{O}(\mathsf{save_twin_2/\top})\}$ and the propositional fact $\mathsf{save_twin_1}, \mathsf{save_twin_2} \Rightarrow \bot$ which stipulates that saving both twins is impossible. Neither of the two assumptions is derivable because it is in conflict with the other one. However, the formula $\mathsf{save_twin_1} \lor \mathsf{save_twin_2}$ is noncontradictory, hence we can derive $\mathcal{O}(\mathsf{save_twin_1} \lor \mathsf{save_twin_2/\top})$. This shows that norms which are nonderivable can still serve to derive other norms, and in particular that our system satisfies the disjunctive response of [10] for two conflicting deontic assumptions. Adding superiority between the two assumptions, e.g., stipulating $\mathcal{O}(\mathsf{save_twin_1/\top}) \succ \mathcal{O}(\mathsf{save_twin_2/\top})$, would break the tie and make the $\mathcal{O}(\mathsf{save_twin_1/\top})$ derivable.

Fig. 4. The structural rules.

Example 4.4 Continuing Ex. 3.1.(iv), with the operator characterisation given there for the operators must, ought, should and the assumptions $\{ \text{must}(\neg \text{murder}/\top), \text{ought}(\text{help_friend}/\top) \}$ as well as the unfortunate factual assumption help_friend \Rightarrow murder we can derive must($\neg \text{murder}/\top$), ought($\neg \text{murder}/\top$), should($\neg \text{murder}/\top$). We also derive ought(murder/ \top) using ought $_R^{\text{ought}(\text{help_friend}/\top)}$, but since ought behaves like a recommendation and hence doesn't satisfy the D axiom, these two are not in conflict.

Example 4.5 Consider the order puzzle from, e.g., [14], with the operator \mathcal{O} as in Ex. 4.2 and deontic assumptions given by the ordered list $\mathcal{O}(\neg \mathsf{open_window/heating}) \succ \mathcal{O}(\mathsf{open_window/T}) \succ \mathcal{O}(\mathsf{heating/T})$. For the situation where the window is open and the heating is off we can derive $\mathcal{O}(\mathsf{open_window/open_window} \land \neg \mathsf{heating})$ as well as $\mathcal{O}(\mathsf{heating/open_window} \land \neg \mathsf{heating})$, but not $\mathcal{O}(\neg \mathsf{open_window/open_window} \land \neg \mathsf{heating})$, since the assumption $\mathcal{O}(\neg \mathsf{open_window/heating})$ does not apply. This illustrates that deontic detachment/transitivity does not hold (since these principles are not present in the base logic). In particular, there also is no aggregation of priorities along chains of obligations which could make the assumption $\mathcal{O}(\mathsf{open_window/T})$ overrule the inferior $\mathcal{O}(\mathsf{heating/T})$. A similar effect could be achieved, however, by adding the assumption $\mathcal{O}(\neg \mathsf{heating/open_window})$, since by specificity this would block the derivation of $\mathcal{O}(\mathsf{heating/open_window} \land \neg \mathsf{heating})$.

5 Cut-elimination and Consequences

We now consider the formal details of the introduced calculi. Due to the underivability statements in the rules we proceed in two stages.

Definition 5.1 We call deontic assumptions a finite set $\mathfrak L$ of non-nested deontic formulae. We further call propositional facts a finite set $\mathfrak F$ of atomic sequents closed under applications of the cut rule below and the contraction rules $\mathsf{Con}_L, \mathsf{Con}_R$ of Fig. 4.

$$\frac{\Gamma\Rightarrow\Delta,A\quad A,\Sigma\Rightarrow\Pi}{\Gamma,\Sigma\Rightarrow\Delta,\Pi} \text{ cut }$$

A normative basis is a triple $\mathfrak{N}=(\mathfrak{O},\mathfrak{L},\succ,\mathfrak{F})$ consisting of an operator characterisation \mathfrak{O} , deontic assumptions \mathfrak{L} with a superiority relation \succ , and propositional facts \mathfrak{F} . Given a normative basis, the rules of the system $\mathsf{G}_{\mathfrak{N}}$ are those of the base calculus for \mathfrak{O} from Fig. 2, the factual assumption rules \mathfrak{F} , the deontic assumption rules of Fig. 3 and the structural rules of Fig. 4. The system $\mathsf{G}_{\mathfrak{N}}$ cut extends $\mathsf{G}_{\mathfrak{N}}$ with the rule cut.

Definition 5.2 Given a normative basis $\mathfrak{N} = (\mathsf{Op}, \mathfrak{L}, \succ, \mathfrak{F})$, a proto-derivation in $\mathsf{G}_{\mathfrak{N}}$ (or $\mathsf{G}_{\mathfrak{N}}\mathsf{cut}$) is a finite labelled tree, with every internal node labelled

with a sequent which is obtained from the labels of the node's children using a rule of $\mathsf{G}_{\mathfrak{N}}$ (or $\mathsf{G}_{\mathfrak{N}}$ plus cut, respectively), and every leaf node labelled with the conclusion of a zero-premiss rule in $\mathsf{G}_{\mathfrak{N}}$ or an underivability statement $\nvdash \Gamma \Rightarrow \Delta$. The conclusion of a proto-derivation is the label of its root. A proto-derivation of rank n is a proto-derivation where the nesting depth of operators from Op in every formula occurring in it is at most n. A proto-derivation (of rank n) is a derivation (of rank n), if none of the underivability statements occurring in it have a derivation in $\mathsf{G}_{\mathfrak{N}}$ cut (of rank n-1). We write $\vdash_{\mathsf{G}_{\mathfrak{N}}} \Gamma \Rightarrow \Delta$ if there is a derivation of rank n.

Note that underivability statements always range over $G_{\mathfrak{N}}$ cut, i.e., the system with the cut rule. Since the definition of a derivation refers to itself, we need to show that it is well-defined. This follows from the observation that the modal nesting depth of the underivability statements occurring in the premisses of the assumption rules is strictly smaller than that of the conclusion, together with the main result of this section, stating that cut is admissible.

Before proving this theorem (in its stronger version, namely that the cut rule is eliminable) we show some preliminary results:

Proposition 5.3 The following rules are derivable in G_Ncut:

$$\frac{\mathsf{Impl}_{\heartsuit, \spadesuit}(A, B) \quad \mathsf{Impl}_{\spadesuit, \clubsuit}(B, C)}{\mathsf{Impl}_{\heartsuit, \spadesuit}(A, C)} \ \mathsf{cut} \qquad \frac{\mathsf{Impl}_{\heartsuit, \spadesuit}(A, B) \quad \mathsf{Confl}_{\spadesuit, \clubsuit}(B, C)}{\mathsf{Confl}_{\heartsuit, \spadesuit}(A, C)} \ \mathsf{cut}$$

Proof. By applying cut and spelling out the cases for Impl and Imple Confl. \Box

Lemma 5.4 The generalised initial sequents $\Gamma, A \Rightarrow A, \Delta$ are derivable.

Proof. By induction on the depth of the derivation, using
$$\mathsf{Mon}_{\heartsuit,\heartsuit}$$
.

The proof of the cut-elimination theorem generalizes the one in [8], which was tailored to the particular rules of the modalities for the dyadic version of the non-normal deontic logic MD [6] (see Section 6.1).

Theorem 5.5 (Cut elimination) If $\vdash_{\mathsf{G}_{\mathfrak{N}}\mathsf{cut}}\Gamma\Rightarrow\Delta$, then $\vdash_{\mathsf{G}_{\mathfrak{N}}}\Gamma\Rightarrow\Delta$.

Proof. By eliminating topmost applications of multicut, i.e., the rule

$$\frac{\Gamma\Rightarrow\Delta,A^n\quad A^m,\Sigma\Rightarrow\Pi}{\Gamma,\Sigma\Rightarrow\Delta,\Pi}\ \, \mathrm{mcut}$$

using a double induction on the complexity of the cut formula A and the sum of the depths of the two premisses of the application of multicut. The interesting case is for A being a deontic formula, the propositional cases are standard.

The case of the last applied rules being modal is straightforward, e.g., for

$$\frac{\mathsf{Impl}_{\heartsuit, \spadesuit}(A, C) \quad B \Rightarrow D \quad D \Rightarrow B}{\Gamma, \heartsuit(A/B) \Rightarrow \spadesuit(C/D), \Delta} \ \mathsf{Mon}_{\heartsuit, \spadesuit} \quad \frac{\mathsf{Confl}_{\spadesuit, \clubsuit}(C, E) \quad D \Rightarrow F \quad F \Rightarrow D}{\Sigma, \spadesuit(C/D), \clubsuit(E/F) \Rightarrow \Pi} \ \mathsf{D}_{\spadesuit, \clubsuit}$$

we replace the cut on $\spadesuit(C/D)$ by cuts on the premisses and an application of $D_{\heartsuit, \clubsuit}$ using Prop.5.3 and the fact that since $\heartsuit \to \spadesuit$ and $\spadesuit_{\rlap/} \clubsuit$, also $\heartsuit_{\rlap/} \clubsuit$. The

case of both rules being Mon is similar. A multicut between the conclusions of

$$\frac{\mathsf{Impl}_{\heartsuit, \spadesuit}(A, C) \quad B \Rightarrow D \quad D \Rightarrow B}{\Gamma, \heartsuit(A/B) \Rightarrow \spadesuit(C/D), \Delta} \ \mathsf{Mon}_{\heartsuit, \spadesuit} \quad \frac{\mathsf{Confl}_{\spadesuit, \spadesuit}(C, C) \quad D \Rightarrow D \quad D \Rightarrow D}{\Sigma, \spadesuit(C/D), \spadesuit(C/D) \Rightarrow \Pi} \ \mathsf{D}_{\spadesuit, \spadesuit}(C, C) \quad D \Rightarrow D \quad D \Rightarrow D$$

is replaced by two cuts of smaller complexity, obtaining first $\mathsf{Confl}_{\heartsuit, \spadesuit}(A, C)$ and then $\mathsf{Confl}_{\heartsuit,\heartsuit}(A,A)$ using Prop.5.3. Then we apply $\mathsf{D}_{\heartsuit,\heartsuit}$. The case of a cut between the conclusions of the rules $\mathsf{Mon}_{\heartsuit, \spadesuit}$ and P_{\spadesuit} is analogous.

The case involving the right assumption rule and the monotonicity rule is as follows (strictly speaking, the first denotes a set of rules). Suppose we have

$$B \Rightarrow D$$

$$|\operatorname{Impl}_{\Phi,\heartsuit}(C,A) = \{ F \operatorname{Confl}_{\Phi,\Phi}(A,A) : \operatorname{nt}(\Phi), \Phi = \emptyset \}$$

$$\left\{ \begin{bmatrix} F & B \Rightarrow F \\ F & \operatorname{Confl}_{\Phi,\diamondsuit}(E,A) \\ F & F \Rightarrow D \\ F & F & F \end{bmatrix} : \begin{pmatrix} (E/F) \in \mathfrak{L}, \\ (E/F) \notin \Phi = \mathfrak{L}, \\ ($$

and

$$\frac{\mathsf{Impl}_{\heartsuit,\heartsuit'}(A,G) \quad B \Rightarrow H \quad H \Rightarrow B}{\Sigma, \heartsuit(A/B) \Rightarrow \heartsuit'(G/H), \Pi} \ \mathsf{Mon}_{\heartsuit,\heartsuit'}$$

By induction hypothesis on the cut complexity we obtain the premisses of

This uses Prop. 5.3 for obtaining $\mathsf{Impl}_{\spadesuit, \heartsuit'}(C,G)$ and $\mathsf{Impl}_{\diamondsuit, \heartsuit'}(X,G)$, as well as obtaining $\nvdash \mathsf{Confl}_{\clubsuit, \heartsuit'}(E,G)$ from $\nvdash \mathsf{Confl}_{\clubsuit, \heartsuit}(E,A)$ and $\mathsf{Impl}_{\heartsuit, \heartsuit'}(A,G)$. Finally, Prop. 5.3 also yields $\nvdash \mathsf{Confl}_{\heartsuit,\heartsuit'}(G,G)$ from $\nvdash \mathsf{Confl}_{\heartsuit,\heartsuit}(A,A)$ and $\mathsf{Impl}_{\heartsuit,\heartsuit'}(A,A)$ in case we have $\mathsf{nt}(\heartsuit')$ and the premiss needs to be present – in that case we also have $\operatorname{nt}(\heartsuit)$ and the corresponding premiss is in (1) as well. The cases involving $\heartsuit_L^{\spadesuit(C/D)}$ and $\operatorname{\mathsf{Mon}}_{\heartsuit,\heartsuit'}$ or $\heartsuit_R^{\spadesuit(C/D)}$ and $\operatorname{\mathsf{D}}_{\heartsuit,\heartsuit'}$ are similar. For the case of a multicut between $\heartsuit_R^{\spadesuit(C/D)}$ and both principal formulae of

the D rule, we claim that this cannot happen. For suppose we had (1) and

$$\frac{\mathsf{Confl}_{\heartsuit,\heartsuit}(A,A) \quad B \Rightarrow B \quad B \Rightarrow B}{\Sigma, \heartsuit(A,B), \heartsuit(A/B) \Rightarrow \Pi} \ \mathsf{D}_{\heartsuit,\heartsuit}$$

If $\operatorname{nt}(\heartsuit)$ we immediately obtain a contradiction since $\operatorname{Confl}_{\heartsuit,\heartsuit}(A,A)$ is both derivable and not derivable. Otherwise, since $\operatorname{D}_{\heartsuit,\heartsuit}$ is in the system, we have $\heartsuit \not\models \heartsuit$, and since the rule $\heartsuit_R^{\spadesuit(C/D)}$ was used, we have $\spadesuit \to \heartsuit$. Hence we also have $\spadesuit \not\models \spadesuit$. Thus one instance of $\clubsuit(E/F)$ in the set of premisses of $\heartsuit_R^{\spadesuit(C/D)}$ is the assumption $\spadesuit(C/D)$. But for this formula the first premiss gives us $B \Rightarrow D$, hence in the choice block the instantiation $\nvdash B \Rightarrow D$ of the first underivability statement $\nvdash B \Rightarrow F$ does not hold. Further, from Prop. 5.3 with the premisses $\operatorname{Impl}_{\spadesuit,\heartsuit}(C,A)$ and $\operatorname{Confl}_{\heartsuit,\heartsuit}(A,A)$ we get $\operatorname{Confl}_{\spadesuit,\heartsuit}(C,A)$, hence this instantiation of the second underivability statement $\nvdash \operatorname{Confl}_{\spadesuit,\heartsuit}(E,A)$ of the choice block also does not hold. Finally, the instantiation $\nvdash D \Rightarrow D$ of the third underivability statement $\nvdash F \Rightarrow D$ also does not hold due to Lem. 5.4, and hence the proto-derivation ending in (1) cannot have been a derivation. The case involving $\heartsuit_R^{\spadesuit(C/D)}$ and P_{\heartsuit} is completely analogous. Also in the case of the assumption right rule versus the assumption left rule

Also in the case of the assumption right rule versus the assumption left rule we claim that this cannot happen. Suppose we would have (1) and

Since both rules are in the system, we have $\spadesuit \to \heartsuit$ and $\spadesuit' \not\downarrow \heartsuit$, and hence also $\spadesuit' \not\downarrow \spadesuit$. Further, since the superiority relation is acyclic, we have either $\spadesuit(C/D) \not\succ \spadesuit'(C'/D')$ or $\spadesuit'(C'/D') \not\succ \spadesuit(C/D)$. Suppose $\spadesuit(C/D) \not\succ \spadesuit'(C'/D')$. Then instantiating $\clubsuit(E/F)$ in the premisses of $\heartsuit_R^{\spadesuit(C/D)}$ with $\spadesuit'(C'/D')$ we have either $\nvdash B \Rightarrow D'$, or $\nvdash \mathsf{Confl}_{\spadesuit',\heartsuit}(C',A)$ or $\nvdash D' \Rightarrow D$ together with the choice. The first of these cannot be the case, because from $\heartsuit_L^{\spadesuit'(C'/D')}$ we have $B \Rightarrow D'$. The second also cannot be the case because again from $\heartsuit_L^{\spadesuit'(C'/D')}$ we get $\mathsf{Confl}_{\spadesuit',\heartsuit}(C',A)$. So assume that $\nvdash D' \Rightarrow D$ and for some $\diamondsuit(X/Y) \in \mathfrak{L}$ with $\diamondsuit\not\downarrow \spadesuit'$ and $\spadesuit'(C'/D') \not\succ \diamondsuit(X/Y)$ we have all three of

$$B \Rightarrow Y \quad Y \Rightarrow D' \quad \mathsf{Impl}_{\land \circlearrowleft}(X, A)$$
 (2)

But then instantiating this assumption $\diamondsuit(X/Y)$ for $\clubsuit'(E'/F')$ in the premisses of $\heartsuit_L^{\spadesuit'(C'/D')}$ yields that one of $\nvdash B \Rightarrow Y$ or $\nvdash \operatorname{Impl}_{\diamondsuit, \heartsuit}(X, A)$ or $\nvdash Y \Rightarrow D'$ holds. This is clearly in contradiction to (2). Hence every possibility yields a contradiction, and thus one of the two proto-derivations was not a derivation. The case of $\spadesuit'(C'/D') \not\succ \spadesuit(C/D)$ is analogous, starting with instantiating the formula $\clubsuit'(E'/F')$ in the premisses of the rule $\heartsuit_L^{\spadesuit'(C'/D')}$ with the assumption $\spadesuit(C/D)$ and then reasoning as in the first case.

An important corollary of this result is that we can reduce derivability to

derivability of bounded rank, and hence obtain well-definedness of the former notion:

Theorem 5.6 (Derivability is well-defined) Let the maximal nesting depth of operators in $\Gamma \Rightarrow \Delta$ be n. Then we have $\vdash_{\mathsf{G}_{\mathfrak{N}}\mathsf{cut}} \Gamma \Rightarrow \Delta$ iff $\vdash_{\mathsf{G}_{\mathfrak{N}}} \Gamma \Rightarrow \Delta$ iff $\vdash_{\mathsf{G}_{\mathfrak{N}}} \Gamma \Rightarrow \Delta$. Hence derivability in $\mathsf{G}_{\mathfrak{N}}$ is well-defined.

Proof. The first equivalence follows straightforwardly from cut elimination (Thm. 5.5). The proof for the second equivalence is by induction on n. For n=0 the sequent is purely propositional. Hence the derivation cannot contain underivability statements, and the statement is straightforward. Suppose the statement holds for all m < n. Due to the shape of the rules, every sequent in a derivation of $\Gamma \Rightarrow \Delta$ has nesting depth $\leq n$, and the underivability statements mention sequents of depth $\leq n-1$. Thus by induction hypothesis on the underivability statements the derivation is of rank n and we have $\vdash_{\mathsf{G}_{\mathfrak{N}}}^{n} \Gamma \Rightarrow \Delta$. Similarly, if $\vdash_{\mathsf{G}_{\mathfrak{N}}}^{n} \Gamma \Rightarrow \Delta$, then by induction hypothesis on the underivability statements occurring in the derivation we obtain $\vdash_{\mathsf{G}_{\mathfrak{N}}} \Gamma \Rightarrow \Delta$.

As a further corollary we obtain decidability of the system and complexity results. Notably, the complexity of reasoning from assumptions is the same as that of reasoning without assumptions in Standard Deontic Logic [18]:

Theorem 5.7 Given \mathfrak{N} , the problem of deciding whether $\vdash_{\mathsf{G}_{\mathfrak{N}}} \Gamma \Rightarrow \Delta$ is decidable in space polynomial in the size of $\Gamma \Rightarrow \Delta$.

Proof. (Sketch) The idea is to perform backwards proof search to find a proto derivation. For each underivability statement we then recursively call the algorithm and flip the answer. To prevent loops caused by contraction, we copy the principal formula of the implication rules into the premisses and omit the weakening and contraction rules. Standard inductions on the depth of the proto derivation then show admissibility of the contraction and weakening rules. The proof search procedure existentially guesses the last applied rule, checks that its application is non-redundant, i.e., introduces at least one new formula, then universally chooses its premisses and checks derivability. Since each backwards application of a rule adds at least one new subformula of the conclusion or reduces the maximal nesting depth of the sequent, the depth of the search tree is polynomial in the size of the conclusion. Since moreover its branching factor only depends on the number of rules, i.e., deontic assumptions, it is independent of the size of the input. Hence the procedure runs in alternating polynomial time, which is equivalent to polynomial space [5].

6 Applications

We apply our methodology to the case studies of Mīmāṃsā-inspired logic, permissions as exception, and a logic of sanction and violation, showing how contrary-to-duties can be modeled as instance of defeasible reasoning [26].

6.1 Mīmāṃsā-inspired logic

The specificity rules in [8] for the Mīmāṃsā-inspired logic are a particular case of our general rule schemas. Before showing how to model these rules, and

how to extend them with prioritized obligations, we briefly recall the logic in question, introduced to formalize and provide a better understanding of the deontic reasoning of Mīmāṃsā authors. Mīmāṃsā is an ancient influential school of Indian philosophy mainly focusing on the exegesis of the prescriptive portions of the Vedas – the Sacred Texts of Hinduism. In order to explain the deontic content of the Vedas and interpret them in a noncontradictory way, Mīmāṃsā authors proposed a rich body of deontic, hermeneutical and linguistic principles called $ny\bar{a}yas$. In [7] some of the deontic $ny\bar{a}yas$ were transformed into Hilbert axioms for a non-normal dyadic deontic logic, which yielded a formal analysis of a famous deontic controversy contained in the Vedas. Interestingly, this solution coincided with that of Prabhākara, one of the chief Mīmāṃsā authors, which previous approaches failed to make sense of. As shown in [9] the \Box -free fragment of this logic is the dyadic version of the non-normal deontic logic MD [6].

Not all $ny\bar{a}yas$ can be converted into Hilbert axioms. These include more general interpretative principles to resolve apparent contradictions in the Vedas like the specificity principle, discussed already by Mīmāmsā author Sabara (3rd-5th c. CE) under the name $gunapradh\bar{a}na$. Hence the dyadic version of MD was extended in [8] with sequent rules for specificity. These rules can be seen as a particular case of the general scheme described here by considering an operator characterisation with only one obligation type operator \mathcal{O} with $\mathcal{O}_{\frac{1}{2}}\mathcal{O}$ and no superiority relation. Going beyond [8], the superiority relation in the rules of Fig. 3 lets us deal with the Mīmāmsā interpretative principle called hierarchy of sources (śrutismrtyādibādha). This principle states that out of two apparently clashing commands, the one issued by a less authoritative source is to be suspended. Indeed, Mīmāmsā author Kumārila describes four sources of duty, in decreasing order of authority: śruti (the Vedas), smrti (the 'recollected texts', based on the Vedas), $sad\bar{a}c\bar{a}ra$ (the behaviour of good people, who are learned in the Vedas) and ātmatusti (the inner feeling of approval by people who are learned in the Vedas). Hence, the considered norms can be formalized by four obligation type operators $\mathcal{O}_V, \mathcal{O}_{rt}, \mathcal{O}_{gp}, \mathcal{O}_{if}$ with $\emptyset \not \downarrow \spadesuit$ for each $\emptyset, \spadesuit \in$ $\{\mathcal{O}_{\mathsf{V}}, \mathcal{O}_{\mathsf{rt}}, \mathcal{O}_{\mathsf{gp}}, \mathcal{O}_{\mathsf{if}}\}$, with the transitive closure of the priorities $\mathcal{O}_{\mathsf{V}}(A/B) \succ$ $\mathcal{O}_{\mathsf{rt}}(C/D), \ \mathcal{O}_{\mathsf{rt}}(C/D) \succ \mathcal{O}_{\mathsf{gp}}(E/F) \ \text{and} \ \mathcal{O}_{\mathsf{gp}}(E/F) \succ \mathcal{O}_{\mathsf{if}}(G/H) \ \text{between any}$ assumptions using these operators.

6.2 Permissions as exceptions

Considered often as the dual of obligation, permission has been treated as primitive operator as well [22,11]. Here we model the notion of permissions as exceptions to other deontic operators (compare [2] for an analogous treatment in the context of input-output logics). Intuitively, a permission $\mathcal{P}^{\heartsuit}(A/B)$ acts as an exception to deontic assumptions in \heartsuit , in that it blocks the derivation of a formula $\heartsuit(C/D)$ whenever A and C are in conflict. To define what "in conflict" means, we assume that what is permitted is not forbidden, i.e., that given $\mathcal{P}^{\heartsuit}(A/B)$ we have not $\heartsuit(A/B)$ if \heartsuit is of prohibition type and not $\heartsuit(\neg A/B)$ for \heartsuit of obligation type. This suggests that permission operators are of obligation

type, i.e., upwards monotone in the first argument, in line with the standard notion that if something is permitted, everything which follows from this is also permitted. Thus, to model permissions for an operator \heartsuit , we add an obligation type operator \mathcal{P}^{\heartsuit} with $\heartsuit \not\in \mathcal{P}^{\heartsuit}$. Note that \heartsuit could be of obligation or prohibition type, and it can but does not need to satisfy $\heartsuit \not\in \heartsuit$ and $\mathsf{nt}(\heartsuit)$.

Example 6.1 To model the sentence "Parking is forbidden, unless one has a permit" we use a prohibition-type operator \mathcal{F} with $\mathcal{F} \not\downarrow \mathcal{F}$ and the corresponding (obligation type) permission operator $\mathcal{P}^{\mathcal{F}}$ with $\mathcal{F} \not\downarrow \mathcal{P}^{\mathcal{F}}$. The deontic assumptions are $\{\mathcal{F}(\mathsf{parking}/\top), \mathcal{P}^{\mathcal{F}}(\mathsf{parking}/\mathsf{permit})\}$. We can then derive, e.g., $\mathcal{F}(\mathsf{parking}/\top)$ and $\mathcal{F}(\mathsf{parking}/\mathsf{lazy})$, but neither $\mathcal{F}(\mathsf{parking}/\mathsf{permit})$ nor $\mathcal{F}(\mathsf{parking}/\mathsf{permit} \land \mathsf{lazy})$. Hence the permission $\mathcal{P}^{\mathcal{F}}(\mathsf{parking}/\mathsf{permit})$ acts as an explicit exception to the more general prohibition $\mathcal{F}(\mathsf{parking}/\top)$.

Note that adding permission operators also makes permission formulae derivable, e.g., $\mathcal{P}^{\mathcal{F}}(\mathsf{parking/permit} \land \mathsf{lazy})$ in Ex. 6.1. These could be read as "explicit" or "strong" permissions in that they are derived from permissions explicitly mentioned in the assumptions. To keep them implicit, we can consider permissions in the assumptions, but not as derived formulae. Note also that to introduce a more general permission operator \mathcal{P} which acts as exception to several other operators $\heartsuit_1, \ldots, \heartsuit_n$, it is enough to add $\heartsuit_i \not\leftarrow \mathcal{P}$ for every $i \leq n$.

6.3 Sanctions and violations

We can also use our approach to differentiate between exceptions to a primary norm (as above), and secondary norms, which come into effect after a primary one has been violated. The crucial difference is that for exceptions to a more general norm there is no violation, whereas for secondary norms the primary one stays in force, and hence can be violated. This is similar to the distinction between violations of norms and sanctions as a result of violations. We model this using two prohibition type operators S and V with corresponding permission operators P^S and P^V as in Sec. 6.2. The intuitive reading of S(A/B) is that A is forbidden given B, and doing A results in a sanction. For V(A/B) we read that A is forbidden given B, and doing A results in a violation but not necessarily a sanction. Here we assume that there is no sanction without violation, $S \to V$, and that $V \not\in V$, $V \not\in P^V$, $S \not\in P^S$. Closure under \to then yields $S \not\in V$, $S \not\in S$, $S \not\in P^V$. The latter means that exceptions to violations can overrule sanctions, but in absence of $V \not\in P^S$ exceptions to sanctions cannot overrule violations. Hence there might be a violation, even though there is no sanction.

Example 6.2 Consider the assumptions $\{S(\mathsf{parking}/\top), \mathcal{V}(\mathsf{parking}/\top), \mathcal{P}^{\mathcal{V}}(\mathsf{parking}/\mathsf{permit}), \mathcal{P}^{\mathcal{S}}(\mathsf{parking}/\mathsf{fine_paid})\}$, modelling the fact that once a fine for illegal parking has been paid, there is no further sanction. We derive all three of $S(\mathsf{parking}/\top)$, $\mathcal{V}(\mathsf{parking}/\top)$, $\mathcal{V}(\mathsf{parking}/\mathsf{fine_paid})$. However, we cannot derive either of $S(\mathsf{parking}/\mathsf{fine_paid})$, $S(\mathsf{parking}/\mathsf{permit})$, $\mathcal{V}(\mathsf{parking}/\mathsf{permit})$. The first of these is overruled by $\mathcal{P}^{S}(\mathsf{fine_paid})$, the second and third ones by $\mathcal{P}^{\mathcal{V}}(\mathsf{parking}/\mathsf{permit})$. So if there is no permit, but the fine has been paid, there is no further sanction but still a violation of the prohibition to park.

Remark 6.3 Similarly, we can model contrary-to-duty (CTD) obligations while maintaining the distinction between defeasibility and violation of primary obligations. Indeed, borrowing the example from [26], we can model "There must be no fence", as $\mathcal{S}(\mathsf{fence}/\top) \land \mathcal{V}(\mathsf{fence}/\top)$, and "If there is a fence, it must be a white fence" as $\mathcal{P}^{\mathcal{S}}(\mathsf{white_fence}/\mathsf{fence})$. Then we derive that the primary obligation is in force ($\mathcal{S}(\mathsf{fence}/\top) \land \mathcal{V}(\mathsf{fence}/\top)$) and having a white fence results in a violation of the primary norm ($\mathcal{V}(\mathsf{white_fence}/\mathsf{fence})$), but does not violate the secondary norm ($\mathcal{V}(\mathsf{white_fence}/\mathsf{fence})$). This distinction between violations of primary and secondary norms is somewhat similar to the distinction between instrumental/actual and proper/ideal obligations in [28] and [4] respectively: Roughly speaking, proper or ideal obligations, i.e., all obligations that apply to a context, including violated primary ones, correspond to the violation operator, while instrumental or actual ones, i.e., those detailing what to do in a particular situation, correspond to the sanction operator.

In general, CTDs of other CTDs are modeled by as many different operators as nested CTDs +1. A similar approach is in [13], that employs the (n-ary) substructural connective \otimes where $A \otimes B$ stands for "the violation of A can be repaired by B" to reduce CTD to a special kind of normative exception.

7 Conclusions and Related Work

We introduced sequent rules for reasoning with deontic assumptions using specificity in presence of prioritized deontic operators. The method, which relies on cut elimination in presence of underivability premisses, captures systems with an arbitrary finite number of dyadic deontic operators based on M possibly extended with axioms P or D and inclusions among the operators. The method is applied to various case studes and implemented in deonticProver2.0.

Related work. The approaches closest to ours are those in the framework of dyadic deontic logic, e.g., [33,6,32,20,26]. The main difference is that we consider reasoning from deontic assumptions to be inherently nonmonotonic, and hence do not attempt to capture it purely axiomatically. Indeed, while from the assumption $\mathcal{O}(A/\top)$ we derive $\mathcal{O}(A/\top)$, this no longer holds if we add the conflicting assumption $\mathcal{O}(\neg A/\top)$. This aspect cannot be captured in a purely axiomatic setting, since propositional logic already gives $\mathcal{O}(A/\top) \wedge \mathcal{O}(\neg A/\top) \to \mathcal{O}(A/\top)$. Additionally, unlike our system, most dyadic deontic logics derive $\mathcal{O}(A/A)$, which rules out, e.g., the derivation of a formula like $\mathcal{O}(\neg asparagus/asparagus)$ in Ex. 4.2.

In the nonmonotonic setting, different methods have been introduced to deal with conflicts using specificity and/or superiority; these are either logic-tailored, e.g. [29,27], or are handled within general frameworks like the following.

Deontic default logic [15,16] uses semantical extensions to provide a credulous or skeptical approach (an obligation is derivable if it belongs to at least one or all extensions, respectively). While our system is heavily influenced by the notions of specificity and overriding in [15,16], it avoids the fixpoint construction necessary there, accounts for explicit exceptions, and permits nested obligations on the logic level.

Defeasible deontic logic (DDL), introduced in [12], uses facts, strict and defeasible rules, undercutting rules, and a binary superiority relation on the rules to solve conflicts between defeasible rules. The main differences with our approach are that in DDL propositional reasoning is defeasible, tractable complexity is paid for by the omission of binary connectives, specificity is handled "manually" by adding the superiority relation to all rules where it should apply.

A very influential logic expressing conditional norms is $Input-Output\ Logic\ [21,23,24]$. The main difference w.r.t. our approach is that their base logic is based fundamentally on (deontic or factual) detachment principles. Perhaps more in line with the notion of $contextual\ obligations\ [26]$, neither of these holds in our system, nor, e.g., in the Mīmāṃsā-inspired logic (see Section 6.1).

Limitations and future work. An obvious limitation of our proposal is that the underlying non-normal deontic logics are rather weak. In particular, it would be interesting to extend the logic with an aggregation principle $\heartsuit(A/C) \wedge \heartsuit(B/C) \rightarrow \heartsuit(A \wedge B/C)$. We anticipate that this is possible by suitably adjusting the assumption rules, albeit at a severe cost to the complexity. The more interesting question is how to extend the assumption rules to additional axioms in a general way. We'd also like to solve the limitation mentioned in [16] and rule out conflicts between more than two deontic assumptions, i.e., to incorporate the rules $\vdash \neg(A_1 \land \cdots \land A_n/B)/\vdash \neg(\heartsuit(A_1/B) \land \cdots \land \heartsuit(A_n/B))$ in the base logic. This should be possible using methods similar to those for aggregation. A perhaps more challenging extension would be to incorporate principles like deontic detachment / transitivity. It is not entirely clear whether it is possible to avoid a fixpoint construction in this case. Finally, while neighbourhood semantics for the base logics as in [6] are reasonably straightforward, the big challenge is to find a suitable semantic characterisation for the assumption rules. These topics are left for future work.

References

- Björnsson, G. and R. Shanklin, 'must', 'ought' and the structure of standards, in: F. Cariani, D. Grossi, J. Meheus and X. Parent, editors, DEON 2014, LNAI 8554, Springer, 2014 pp. 33–48.
- [2] Boella, G. and L. van der Torre, *Institutions with a hierarchy of authorities in distributed dynamic environments*, Artif. Intell. Law **16** (2008), pp. 53–71.
- [3] Bonatti, P. A. and N. Olivetti, Sequent calculi for propositional nonmonotonic logics, ACM Trans. Comput. Log. $\bf 3$ (2002), pp. 226–278.
- [4] Carmo, J. and A. J. Jones, Deontic logic and contrary-to-duties, in: D. Gabbay and F. Guenthner, editors, Handbook of Philosophical Logic, Volume 8, Kluwer Academic Publishers, 2002, 2 edition pp. 265–343.
- [5] Chandra, A. K., D. C. Kozen and L. J. Stockmeyer, Alternation, J. Assoc. Comput. Mach. 28 (1981), pp. 114–133.
- [6] Chellas, B. F., "Modal Logic," Cambridge University Press, 1980.
- [7] Ciabattoni, A., E. Freschi, F. Genco and B. Lellmann, Mīmāṃsā deontic logic: proof theory and applications, in: H. De Nivelle, editor, TABLEAUX 2015, LNCS 9323, Springer, 2015 pp. 323–338.

- [8] Ciabattoni, A., F. Gulisano and B. Lellmann, Resolving conflicting obligations in Mīmāmsā: A sequent-based approach, in: J. Broersen, C. Condoravdi, S. Nair and G. Pigozzi, editors, DEON 2018 proceedings, College Publications, 2018 pp. 91–109.
- [9] Freschi, E., A. Ollett and M. Pascucci, Duty and sacrifice. a logical analysis of the Mīmāṃsā theory of Vedic injunctions, History and Philosophy of Logic 40(4) (2019), pp. 323-354.
- [10] Goble, L., Prima facie norms, normative conflicts, and dilemmas, in: Handbook of Deontic Logic and Normative Systems, College Publications, 2013 pp. 241–351.
- [11] Governatori, G., F. Olivieri, A. Rotolo and S. Scannapieco, Computing strong and weak permissions in defeasible logic, J. Philos. Logic 42 (2013), pp. 799–829.
- [12] Governatori, G. and A. Rotolo, Defeasible logic: Agency, intention and obligation, in: DEON 2004 proceedings, LNAI 3065, Springer, 2004 pp. 114–128.
- [13] Governatori, G. and A. Rotolo, Logic of violations: A Gentzen system for reasoning with contrary-to-duty obligations, The Australasian Journal of Logic 4 (2006), pp. 193–215.
- [14] Horty, J., Defaults with priorities, J. Philos. Logic 36 (2007), pp. 367–413.
- [15] Horty, J. F., Deontic logic as founded on nonmonotonic logic, Ann. Math. Artif. Intell. 9 (1993), pp. 69–91.
- [16] Horty, J. F., Nonmonotonic foundations for deontic logic, in: D. Nute, editor, Defeasible Deontic Logic, Kluwer, 1997 pp. 17–44.
- [17] Horty, J. F., "Reasons as Defaults," Oxford University Press, 2012.
- [18] Ladner, R. E., The computational complexity of provability in systems of modal logic, SIAM J. Comput. 6 (1977), pp. 467–480.
- [19] Lellmann, B. and D. Pattinson, Constructing cut free sequent systems with context restrictions based on classical or intuitionistic logic, in: K. Lodaya, editor, ICLA 2013, LNCS 7750, Springer, 2013 pp. 148–160.
- [20] Lewis, D., Semantic analyses for dyadic deontic logics, in: S. Stendlund, editor, Logical Theory and Semantic Analysis, Reidel, 1974 pp. 1–14.
- [21] Makinson, D. and L. W. N. van der Torre, Input/output logics, J. Philos. Logic 29 (2000), pp. 383–408.
- [22] Makinson, D. and L. W. N. van der Torre, Permission from an input/output perspective, J. Philos. Logic 32 (2003), pp. 391–416.
- [23] Parent, X. and L. van der Torre, Input/output logic, in: D. Gabbay, J. Horty, X. Parent, R. van der Meyden and L. van der Torre, editors, Handbook of Deontic Logic and Normative Systems, College Publications, 2013 pp. 495–544.
- [24] Parent, X. and L. van der Torre, "Introduction to Deontic Logic and Normative Systems," College Publications, 2018.
- [25] Piazza, M. and G. Pulcini, Unifying logics via context-sensitiveness, Journal of Logic and Computation 1 (2017), pp. 21–40.
- [26] Prakken, H. and M. J. Sergot, Contrary-to-duty obligations, Studia Logica 57 (1996), pp. 91–115.
- [27] Putte, F. V. D. and C. Strasser, A logic for prioritized normative reasoning, J. Log. Comput. 23 (2013), pp. 563—583.
- [28] Straßer, C., A deontic logic framework allowing for factual detachment, J. Appl. Log. 9 (2011), pp. 61–80.
- [29] Straßer, C. and O. Arieli, Normative reasoning by sequent-based argumentation, J. Log. Comput. 29 (2019), pp. 381–415.
- [30] Troelstra, A. S. and H. Schwichtenberg, "Basic Proof Theory," Cambridge Tracts In Theoretical Computer Science 43, Cambridge University Press, 2000, 2 edition.
- [31] van der Torre, L., Violated obligations in a defeasible deontic logic, in: A. Cohn, editor, ECAI 94, Wiley, 1994 pp. 371–375.
- [32] van Fraassen, B. C., The logic of conditional obligation, J. Philos. Logic 1 (1972), pp. 417–438.
- [33] von Wright, G., A note on deontic logic and derived obligation, Mind 65 (1956), pp. 507–509