Resolving conflicting obligations in Mīmāṃsā: a sequent-based approach

Agata Ciabattoni
TU Wien, Austria
agata@logic.at

Francesca Gulisano
Scuola Normale Superiore Pisa, Italy
francesca.gulisano@sns.it

Björn Lellmann
TU Wien, Austria
lellmann@logic.at

Abstract

The Philosophical School of Mīmāṃsā provides a treasure trove of more than 2000 years worth of deontic investigations. In this paper we formalize the Mīmāṃsā approach of resolving conflicting obligations by giving preference to the more specific ones. From a technical point of view we provide a method to close a set of prima-facie obligations under a restricted form of monotonicity, using specificity to avoid conflicting obligations in a dyadic non-normal deontic logic. A sequent-based decision procedure for the resulting logic is also provided.

1 Introduction

The Mīmāṃsā is a philosophical school which originated in ancient India in the last centuries BCE and whose main focus was the exegesis of the prescriptive portions of the Indian Sacred Texts (the *Veda*s). To this aim over the course of more than two millennia, Mīmāṃsā authors have analyzed normative statements, resulting in theories considered early deontic logic [14]. Despite the undeniable importance of Mīmāṃsā in Indian philosophy, theology and law, and despite the rigorous structure

We thank E. Freschi for her invaluable help with the Indological aspects of this work, and the referees for their insightful comments concerning related literature. Funded by WWTF project MA16-28.

of its texts lending themselves to formal analysis [6], virtually no logical formalization of the deontic concepts in Mīmāṃsā has been carried out so far. The main reason for this is that most Sanskritists are not trained in mathematical logic, and the untranslated or unanalyzed texts are inaccessible to logicians.

In order to enable readers to understand the Vedas independently of any authorial intention, and explain "what has to be done" in presence of seemingly[1] conflicting obligations, Mīmāṃsā authors have proposed a rich body of deontic, hermeneutical and linguistic principles (metarules), called *nyāya*s. Those principles are so modern, rational, scientific, and systematic [1] that they are still applied in Indian jurisprudence to decide court cases, e.g. [15].

To formalize Mīmāṃsā reasoning in a step-by-step bottom-up approach, we have transformed in [5] some of the deontic *nyāya*s into Hilbert axioms. This led to the introduction of the non-normal dyadic deontic logic bMDL, whose proof-calculus and semantics were successfully used there to analyze the seemingly conflicting obligations in the Vedas concerning the *Śyena* sacrifice. However, bMDL is only a first step towards the formalization of Mīmāṃsā reasoning. In particular, many *nyāya*s are still waiting to be found, translated from Sanskrit, and interpreted, which is the subject of ongoing work. Notice also that not all the *nyāya*s can be simply converted into Hilbert axioms. Some of these indeed offer more general interpretative principles to resolve apparent contradictions in the Vedas; prominent examples of such *nyāya*s are *Guṇapradhāna* and *Vikalpa*, which are investigated in this paper. The Vikalpa principle states that when there is a real conflict between obligations, any of the conflicting injunctions may be adopted as option: this principle is known in deontic logic as *disjunctive response* [10] and corresponds to the phenomenon of *floating conclusions* in nonmonotonic reasoning [17]. The Guṇapradhāna principle states that more specific rules override more generic ones. Already introduced by Śabara (3rd-5th c. CE), Guṇapradhāna is widely used, e.g., in Artificial Intelligence, where it was formulated much later and where it is known as *specificity principle*. These principles are also used to capture *defeasible* reasoning in the context of *nonmonotonic logics* [7, 18, 21, 23]. In particular, the specificity principle may lead to the loss of the monotonicity of the consequence relation: an obligation "α should be the case" could follow from a set of premises Γ, but it might be overruled by a more specific obligation β, so that it does not follow from the set $\Gamma \cup \{\beta\}$ anymore.

In this paper we further pursue the proof-theoretic approach to approximate Mīmāṃsā deontic reasoning by extending the deontic part of bMDL with a mechanism to capture the Guṇapradhāna principle. We provide a sequent calculus that

[1] The Vedas are assumed to be not contradictory and Mīmāṃsā authors invested all their efforts in creating a consistent deontic system.

derives what "has to be done" from the explicit prescriptions contained in the Vedas (*Śrauta* in Sanskrit), and a finite set of propositional facts by resolving conflicts using Guṇapradhāna (specificity), and which satisfies Vikalpa (disjunctive response).

Examples of sequent calculi for defeasible reasoning in normative contexts include [3, 12, 24] (the latter is applied in the context of an argument-based system).

As, e.g., in [10, 27] here we interpret the notion of a conditional obligation being more specific than another one as the conditions of the former implying those of the latter. Our calculus is built on the sequent calculus for the $\Box$-free fragment of bMDL, which turns out to be the dyadic version of non-normal deontic logic MD [4] (cf. Prop. 2.2 and [8]). Additional rules to derive all possible prescriptions are defined using limited monotonicity on the conditions of the (non-nested) prescriptions in the Vedas (prima-facie obligations) "up to conflicting obligations" relative to the given set of facts. These additional rules are motivated by the interpretation given by the Mīmāṃsā author Madhatīthi (9-10th c. CE) that more specific *Śrauta* provide exceptions to more general ones and that the latter apply to all circumstances but those indicated in the exceptions (or implied by them). Apart from this non-monotonic inference from prima facie to actual obligations, all inferences use the monotonic system bMDL. Thus we restrict nonmonotonic reasoning using the specificity principle to resolving possible conflicts between prima-facie obligations, but keep the inferences of the logic for arbitrary formulae deductive (i.e., monotone). This is inspired by [25] which states that Indian philosophers – in particular the Mīmāṃsā author Kumārila – tried to keep their arguments not defeasible "as much as possible". From a technical point of view the advantage is that the consequences of a set of prima-facie obligations can be constructed iteratively instead of by a fixed-point construction as e.g. in [13]. Moreover the system does not use key properties of non-monotonic logics (as, e.g., in [20]) which seem not to hold in Mīmāṃsā reasoning (e.g. cautious monotony, see Example 3.1). Finally, we show that the introduced system provides a decision procedure and satisfies the disjunctive response.

2 Basic Mīmāṃsā Deontic Logic

Basic Mīmāṃsā Deontic Logic bMDL was introduced in [5] as a first step towards the formalization of Mīmāṃsā reasoning. The idea was to define a logical system following a bottom-up approach of extracting deontic principles from the Mīmāṃsā texts. The resulting logic extends the alethic system S4 with the following axiom schemata for the deontic operator $\mathcal{O}(A/B)$, which intuitively reads as "A is obligatory under the condition B":

1. $(\Box(A \to B) \wedge \mathcal{O}(A/C)) \to \mathcal{O}(B/C)$

$$\text{(M)}\ \mathcal{O}(A \wedge B/C) \to \mathcal{O}(A/C) \qquad \text{(D)}\ \neg(\mathcal{O}(A/B) \wedge \mathcal{O}(\neg A/B)) \qquad \frac{A \leftrightarrow C \quad B \leftrightarrow D}{\mathcal{O}(A/B) \to \mathcal{O}(C/D)}\ \mathsf{Cg}$$

Figure 1: The modal part of a Hilbert-style system for dyadic MD.

$$\frac{}{p \Rightarrow p}\ \mathsf{init} \qquad \frac{}{\bot \Rightarrow}\ \bot_L \qquad \frac{\Gamma, B \Rightarrow \Delta \quad \Gamma \Rightarrow A, \Delta}{\Gamma, A \to B \Rightarrow \Delta}\ \to_L \qquad \frac{\Gamma, A \Rightarrow B, \Delta}{\Gamma \Rightarrow A \to B, \Delta}\ \to_R$$

$$\frac{A \Rightarrow C \quad B \Rightarrow D \quad D \Rightarrow B}{\mathcal{O}(A/B) \Rightarrow \mathcal{O}(C/D)}\ \mathsf{Mon} \qquad \frac{A, C \Rightarrow \quad B \Rightarrow D \quad D \Rightarrow B}{\mathcal{O}(A/B), \mathcal{O}(C/D) \Rightarrow}\ \mathsf{D} \qquad \frac{A \Rightarrow}{\mathcal{O}(A/B) \Rightarrow}\ \mathsf{P}$$

$$\frac{\Gamma, A, A \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\ \mathsf{Con}_L \qquad \frac{\Gamma \Rightarrow A, A, \Delta}{\Gamma \Rightarrow A, \Delta}\ \mathsf{Con}_R \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}\ \mathsf{W}_L \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow A, \Delta}\ \mathsf{W}_R$$

Figure 2: The sequent calculus $\mathsf{G_{MD}}$ for dyadic MD.

2. $\Box(B \to \neg A) \to \neg(\mathcal{O}(A/C) \wedge \mathcal{O}(B/C))$
3. $(\Box((B \to C) \wedge (C \to B)) \wedge \mathcal{O}(A/B)) \to \mathcal{O}(A/C)$

Axioms (1)-(3) arise by rewriting some of the Mīmāṃsā deontic interpretative principles (*nyāya*s) as logic formulas. E.g., (1) formalizes three different principles; among them the following reformulation of a Sanskrit *nyāya* in the *Tantrarahasya* (15th-17th c. CE) that can be abstracted as (See [6] for details)

> If the accomplishment of X presupposes the accomplishment of Y, the obligation to perform X prescribes also Y.

Remark 2.1. bMDL is weaker than most known deontic logics, e.g., those in [19]; in particular it has neither any deontic aggregation principles nor any form of factual or deontic detachment. In part this is due to our step-by-step methodology: so far indeed we have not found any mention of corresponding principles in the texts. However, the absence of (factual) detachment principles is also in line with the statement by one of the main authors of Mīmāṃsā, Prabhākara, that "*A prescription regards what has to be done. But it does not say that it has to be done*" (*Bṛhatī* I, 7th c. CE).

Here for simplicity we only consider the box-free fragment of bMDL, which coincides with the dyadic version of the logic MD [4] axiomatized as in Fig. 1 (Prop. 2.2). For space reasons we treat the propositional connectives $\wedge, \vee, \neg$ as defined by $\bot, \to$ in the usual way. In the following we will consider an extension of a *sequent calculus*

$$\frac{\Gamma^{\Box} \Rightarrow \varphi}{\Gamma \Rightarrow \Box\varphi, \Delta}\ \mathsf{4} \qquad \frac{\Gamma, \Box\varphi, \varphi \Rightarrow \Delta}{\Gamma, \Box\varphi \Rightarrow \Delta}\ \mathsf{T} \qquad \frac{\Gamma^{\Box}, \varphi \Rightarrow \theta \quad \Gamma^{\Box}, \psi \Rightarrow \chi \quad \Gamma^{\Box}, \chi \Rightarrow \psi}{\Gamma, \mathcal{O}(\varphi/\psi) \Rightarrow \mathcal{O}(\theta/\chi), \Delta}\ \mathsf{Mon}'$$

$$\frac{\Gamma^{\Box}, \varphi \Rightarrow}{\Gamma, \mathcal{O}(\varphi/\psi) \Rightarrow \Delta}\ \mathsf{D}_1 \qquad \frac{\Gamma^{\Box}, \varphi, \theta \Rightarrow \quad \Gamma^{\Box}, \psi \Rightarrow \chi \quad \Gamma^{\Box}, \chi \Rightarrow \psi}{\Gamma, \mathcal{O}(\varphi/\psi), \mathcal{O}(\theta/\chi) \Rightarrow \Delta}\ \mathsf{D}_2$$

Figure 3: The modal part of the sequent calculus $\mathsf{G_{bMDL}}$ for bMDL from [5].

for this logic, where a *sequent* is a tuple of multisets of formulas, written as $\Gamma \Rightarrow \Delta$. The rules of the sequent calculus $\mathsf{G_{MD}}$ are given in Fig. 2, those of the calculus $\mathsf{G_{bMDL}}$ for bMDL from [5] in Fig. 3, where $\Gamma^{\Box}$ denotes Γ in which all formulas not of the form $\Box\varphi$ are deleted. Note that the usual sequent rules for $\wedge, \vee, \neg$ are derivable using the definitions in terms of $\bot, \to$. As usual, a *derivation* is a finite labelled tree where every node is labelled with a sequent such that the labels of a node follow from the labels of its children using the rules of the calculus. In particular, the leaves are labelled with conclusions of the zero-premise rules init or $\bot_L$, see also [26]. For G one of $\mathsf{G_{MD}}, \mathsf{G_{bMDL}}$ we write $\vdash_{\mathsf{G}} \Gamma \Rightarrow \Delta$ if there is a derivation of $\Gamma \Rightarrow \Delta$ in G. For the original semantic equivalent of the following proposition, see [8].

Proposition 2.2. *If* $\Gamma \Rightarrow \Delta$ *does not contain* $\Box$*, then* $\vdash_{\mathsf{G_{MD}}} \Gamma \Rightarrow \Delta$ *iff* $\vdash_{\mathsf{G_{bMDL}}} \Gamma \Rightarrow \Delta$*. Hence the box-free fragment of* bMDL *is* MD*.*

Proof. One direction of the equivalence follows from changing the rules of $\mathsf{G_{bMDL}}$ into the corresponding rules of $\mathsf{G_{MD}}$ possibly followed by the weakening rules $\mathsf{W}_L, \mathsf{W}_R$. The other direction follows since a derivation in $\mathsf{G_{MD}}$ is a derivation in $\mathsf{G_{bMDL}}$ with the addition of the structural rules of weakening $\mathsf{W}_L, \mathsf{W}_R$ and contraction $\mathsf{Con}_L, \mathsf{Con}_R$, which are admissible in $\mathsf{G_{bMDL}}$ [5, Lem. 1]. Completeness and soundness of $\mathsf{G_{MD}}$ for MD follow from general methods for constructing sequent calculi from axioms and proving cut elimination such as [16]. □

Remark 2.3. The mechanism for handling propositional facts employed in this paper differs from that in [5]: whereas there we encoded such assumptions as boxed formulas in the conclusion of a derivation, here we treat them as leaves. This has the welcome consequence that we can avoid the alethic modality $\Box$ including any question about its axiomatisation, in line with the view that Mīmāṃsā authors did not distinguish between necessity and epistemic certainty.

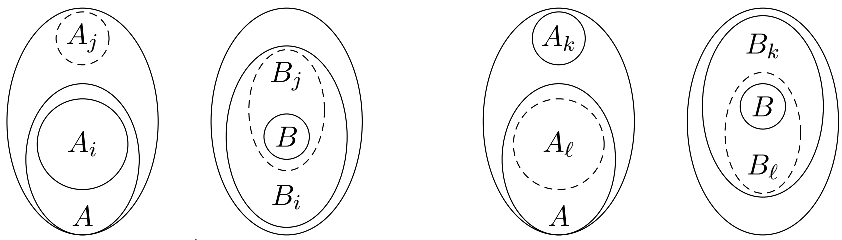
3 Defeasible reasoning in Mīmāṃsā

The specificity principle (*Guṇapradhāna*) is used in Mīmāṃsā to resolve apparent contradictions; these may occur in the set of Vedic (Śrauta) prescriptions or can be derived via the facts. For example, consider the Śrauta prescriptions: (a) A *Śūdra* (i.e., a member of the lower class) should not engage with the Veda, (b) Knowledge of the Vedas is a prerequisite for sacrificing and (c) A chariot maker should sacrifice. The additional fact (d) A chariot maker is a Śūdra, leads to (apparently) conflicting obligations, as extensively discussed by Mīmāṃsā author Jaimini (2nd c. BCE). The following example illustrates the kind of reasoning Mīmāṃsā authors employed to solve such kinds of conflicting obligations.

Example 3.1. Consider the obligations (a) $\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\top)$ ("You ought to perform the ritual offering called Agnihotra") and (b) $\mathcal{O}_{\mathsf{pf}}(\neg\mathtt{agn}/\mathtt{sdr})$ ("You ought not to perform the Agnihotra if you are a Śūdra"). By an implicit deduction from those two premises we could obtain two obligations: (c) $\mathcal{O}(\mathtt{agn}/\gamma)$ and (d) $\mathcal{O}(\neg\mathtt{agn}/\gamma \wedge \mathtt{sdr})$. Now, let us interpret γ as being more specific than $\mathtt{sdr}$, e.g. as "being a chariot maker" ($\mathtt{chmk}$). Since chariot makers are Śūdra, the formulas $\mathtt{chmk}$ and $\mathtt{chmk} \wedge \mathtt{sdr}$ are equivalent, and thus the obligations (c) and (d) give an apparent conflict. One of the solutions to this employed by Mīmāṃsā authors it to interpret (c) as an explicit Vedic (*Śrauta*) prescription, i.e. as $\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\mathtt{chmk})$. In this case, using the specificity principle, the Mīmāṃsā authors derive the opposite of (d), (d') $\mathcal{O}(\mathtt{agn}/\mathtt{chmk} \wedge \mathtt{sdr})$.

However, also a state such that none of γ and $\mathtt{sdr}$ is more specific than the other is compatible with the Mīmāṃsā reasoning; for instance we can imagine a situation where you are asked to decide what to do if you are a Śūdra but you became a school teacher ($\mathtt{sch}$). Also in this case, if (c) is interpreted as a *Śrauta* injunction $\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\mathtt{sch})$, (d) should not follow anymore. In the latter case, writing $\mid\!\sim$ for the consequence relation given by the implicit deduction from Śrauta to actual obligations, we have $\{\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\top), \mathcal{O}_{\mathsf{pf}}(\neg\mathtt{agn}/\mathtt{sdr})\} \mid\!\sim \mathcal{O}(\mathtt{agn}/\mathtt{sch})$ and also $\{\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\top), \mathcal{O}_{\mathsf{pf}}(\neg\mathtt{agn}/\mathtt{sdr})\} \mid\!\sim \mathcal{O}(\neg\mathtt{agn}/\mathtt{sch} \wedge \mathtt{sdr})$, but in contrast also $\{\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\top), \mathcal{O}_{\mathsf{pf}}(\neg\mathtt{agn}/\mathtt{sdr}), \mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\mathtt{sch})\} \not\mid\!\sim \mathcal{O}(\neg\mathtt{agn}/\mathtt{sch} \wedge \mathtt{sdr})$. Hence the Mīmāṃsākas' reasoning can provide a counterexample for *Cautious Monotony* – one of the classical principles of non-monotonic logics [9].

Here we continue the proof-theoretic approach initiated in [5] to reproduce Mīmāṃsā reasoning in a formal framework. We extend the sequent calculus G_{MD} for the logic MD with special rules $\mathsf{ga}_L, \mathsf{ga}_R$ to derive conditional obligations of the form $\mathcal{O}(A/B)$ from prima-facie obligations (i.e. Śrauta prescriptions) written as $\mathcal{O}_{\mathsf{pf}}(C/D)$, adopting limited forms of monotonicity (Sec. 3.1). The resulting calculus is shown to be decidable (Sec. 3.2), applies the specificity principle, and turns



$\exists\, \mathcal{O}_{\mathsf{pf}}(A_i/B_i) \in \mathfrak{L}$ s.t. $\nexists\, \mathcal{O}_{\mathsf{pf}}(A_j/B_j) \in \mathfrak{L}$ $\quad \forall\, \mathcal{O}_{\mathsf{pf}}(A_k/B_k) \in \mathfrak{L}\ \exists\, \mathcal{O}_{\mathsf{pf}}(A_\ell/B_\ell) \in \mathfrak{L}$

Figure 4: A graphical representation of the conditions for $\mathcal{O}(A/B)$ being derivable. Areas can be taken as formulas with containment representing entailment, i.e., more specific formulas are contained in less specific ones.

out to satisfy the disjunctive response (Sec. 3.3).

3.1 Sequent calculus for Specificity/Guṇapradhāna

In order to extend the sequent calculus for MD to capture the specificity principle, loosely following [10, p.281], we interpret the notion of *specificity* as entailment in the presence of (global) propositional assumptions. I.e., given a set $\mathfrak{F}$ of propositional *facts* about the world we say that proposition A is *at least as specific* as proposition B, if $\mathfrak{F}$ entails $A \to B$. Given this interpretation, the *specificity principle* can be understood as limiting monotonicity (the inference by ārtha above) of the operator $\mathcal{O}$ in the second argument in the following sense. Given a list $\mathfrak{L}$ of non-nested *prima facie obligations*, e.g., *Śrauta* prescriptions, and a proposition B, we should be licensed to infer the actual obligation $\mathcal{O}(A/B)$ if there is an injunction $\mathcal{O}_{\mathsf{pf}}(A/C)$ in $\mathfrak{L}$ such that B is at least as specific as C, i.e., we can infer using $\mathfrak{F}$ that $B \to C$, and there is no $\mathcal{O}_{\mathsf{pf}}(D/E)$ in $\mathfrak{L}$ such that B is at least as specific as E and E is at least as specific as C, and further the formulas A and E are inconsistent, i.e., we can infer $\neg(A \wedge E)$. However, while this implements the notion that more specific Śrauta obligations overrule less specific conflicting ones, this only resolves conflicts between propositions $\mathcal{O}_{\mathsf{pf}}(A_i/B_i)$ and $\mathcal{O}_{\mathsf{pf}}(A_j/B_j)$ in $\mathfrak{L}$ for which the conditions are comparable in the sense that either B_i implies B_j or B_j implies B_i. Hence, to make the resulting theory consistent with MD, following the Mīmāṃsā reasoning in Ex. 3.1 we add a further condition stating that there is no obligation $\mathcal{O}_{\mathsf{pf}}(A_k/B_k) \in \mathfrak{L}$ such that B is at least as specific as B_k, the enjoined A and A_k are inconsistent, and which is not overruled by a more specific obligation $\mathcal{O}_{\mathsf{pf}}(A_\ell/B_\ell)$ from $\mathfrak{L}$. Graphically, these two conditions can be visualised as in Fig. 4. In the following we make this formally precise, and prove a cut elimination theorem for the resulting system.

In the remainder of this paper we assume that $\mathfrak{F}$ is a finite set of sequents containing only propositional variables, which is *closed under cuts*, i.e., whenever $\Gamma \Rightarrow \Delta, p$ and $p, \Sigma \Rightarrow \Pi$ are in $\mathfrak{F}$, then so is $\Gamma, \Sigma \Rightarrow \Delta, \Pi$, and *closed under contractions*, i.e., whenever $\Gamma, p, p \Rightarrow \Delta$ or $\Gamma \Rightarrow p, p, \Delta$ are in $\mathfrak{F}$, then so are $\Gamma, p \Rightarrow \Delta$ and $\Gamma \Rightarrow p, \Delta$ respectively. We call $\mathfrak{F}$ the set of *(propositional) facts.* Note that, since every propositional formula is equivalent to a formula in conjunctive normal form, using this definition we can stipulate arbitrary propositional formulas as facts. We further assume a finite set $\mathfrak{L}$ of formulas of the form $\mathcal{O}_{\mathsf{pf}}(A_m/B_m)$ where A_m and B_m do not contain the $\mathcal{O}$-operator. We call these formulas *prima facie obligations.*

To capture the intuition for the specificity principle given above in a well-behaved sequent system, we first need to make the notion of implication used there formally precise. In particular, we would like to define a notion of inference $\vdash$ from the facts in $\mathfrak{F}$ depending on the set $\mathfrak{L}$, such that we can derive a formula $\mathcal{O}(A/B)$ if and only if both of the following hold:

- there is $\mathcal{O}_{\mathsf{pf}}(A_i/B_i) \in \mathfrak{L}$ such that $\mathfrak{F} \vdash B \Rightarrow B_i$ and $\mathfrak{F} \vdash A_i \Rightarrow A$ and for all $\mathcal{O}_{\mathsf{pf}}(A_j/B_j) \in \mathfrak{L}$ we have: ($\mathfrak{F} \nvdash B \Rightarrow B_j$ or $\mathfrak{F} \nvdash B_j \Rightarrow B_i$ or $\mathfrak{F} \nvdash A_j, A \Rightarrow$)
- for all $\mathcal{O}_{\mathsf{pf}}(A_k/B_k) \in \mathfrak{L}$ we have: $\mathfrak{F} \nvdash B \Rightarrow B_k$ or $\mathfrak{F} \nvdash A_k, A \Rightarrow$ or there is a $\mathcal{O}_{\mathsf{pf}}(A_\ell/B_\ell) \in \mathfrak{L}$ such that: ($\mathfrak{F} \vdash B \Rightarrow B_\ell$ and $\mathfrak{F} \vdash B_\ell \Rightarrow B_k$ and $\mathfrak{F} \vdash A_\ell \Rightarrow A$).

Remark 3.2. The formulas we want to infer might have nested deontic operators. Indeed, they should capture key prescriptions like "under the condition of having to perform sacrifice α under the conditions β, you ought to do γ".

To turn this into sequent rules (the rules $\mathsf{ga}_L, \mathsf{ga}_R$ in Def. 3.3 below), we convert every (meta-)conjunction and universal quantifier in this characterization into different premises, while (meta-)disjunctions and existential quantifiers yield a split into different rules. To write the rules in an economic way, for sets $\mathcal{P}, \mathcal{Q}_i$ of premises we use the notation

$$\frac{\mathcal{P} \cup \begin{pmatrix} \mathcal{Q}_1 \\ \vdots \\ \mathcal{Q}_n \end{pmatrix}}{\Gamma \Rightarrow \Delta} \quad \text{for} \quad \left\{ \frac{\mathcal{P} \cup \mathcal{Q}_1}{\Gamma \Rightarrow \Delta}, \quad \dots, \quad \frac{\mathcal{P} \cup \mathcal{Q}_n}{\Gamma \Rightarrow \Delta} \right\}$$

In case $\mathcal{Q}_i$ is a singleton we also omit the braces.

Since the rules now also will mention *underivability*, we further need to add a judgment for this to some of the sequents, written as $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}}$, with the intended meaning that the sequent is not derivable from the facts $\mathfrak{F}$ in the system $\mathsf{G_{MD}ga_{\mathfrak{L}}cut}$ in the sense defined below (Def. 3.4). Thus we will obtain a set of rules

ga_R introducing a formula of the form $\mathcal{O}(A/B)$ on the right hand side of the sequent. For technical reasons we will also add rules ga_L introducing such a formula on the left hand side – these essentially follow from absorbing inferences using the axiom D into the previous rule, and we will show below (Lem. 3.7) that they do not change the set of derivable sequents.

Definition 3.3. Let $\mathfrak{L} = \{\mathcal{O}_{\mathsf{pf}}(A_1/B_1), \ldots, \mathcal{O}_{\mathsf{pf}}(A_n/B_n)\}$ be a finite set of non-nested prima facie obligation formulas and let $\mathfrak{F}$ be a set of propositional sequents. The rules of $\mathsf{ga}_{\mathfrak{L}}$ are given in Fig. 5. A *proto-derivation* with conclusion $\Gamma \Rightarrow \Delta$ in the system $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}$ from assumptions $\mathfrak{F}$ is a finite labelled tree, where each internal node is labelled with a sequent, each leaf is labelled with an initial sequent, a sequent from $\mathfrak{F}$, or an *underivability statement* $\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \Sigma \Rightarrow \Pi$, such that the label of every internal node is obtained from the labels of its children using the rules of G_{MD} or $\mathsf{ga}_{\mathfrak{L}}$. The notion of a *proto-derivation in the system* $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}$ is defined analogously, but also permitting applications of the *cut rule*

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

The *depth* of a proto-derivation is the depth of the underlying tree, i.e., the maximal length of a branch in the tree plus one.

Definition 3.4. A proto-derivation in $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}$ (in $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}$) from $\mathfrak{F}$ is *valid* if for each of the underivability statements $\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \Sigma \Rightarrow \Pi$ occurring as one of the leafs of that derivation there is no valid proto-derivation of $\Sigma \Rightarrow \Pi$ in $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}$ from $\mathfrak{F}$. In case there is such a valid proto-derivation we also write $\mathfrak{F} \vdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}} \Gamma \Rightarrow \Delta$ and $\mathfrak{F} \vdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \Gamma \Rightarrow \Delta$ respectively.

Note that underivability statements are always evaluated in the system with the cut rule. Since the definition of a valid proto-derivation involves the notion of a valid proto-derivation itself, it is not immediately clear that this notion is well-defined. We will show in Thm. 3.10 below that this is indeed the case.

Example 3.5. Consider the prima-facie obligations given by $\mathfrak{L} = \{\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\top), \mathcal{O}_{\mathsf{pf}}(\neg\mathtt{agn}/\mathtt{sdr})\}$ (cf. Ex. 3.1) and the set $\mathfrak{F} = \emptyset$ of facts. Taking the formula $\mathcal{O}_{\mathsf{pf}}(\mathtt{agn}/\top)$ as the formula $\mathcal{O}_{\mathsf{pf}}(A_i/B_i)$ in the general scheme of Fig. 5, we obtain the rules in Fig. 6. In particular, the sequent $\Rightarrow \mathcal{O}(\mathtt{agn}/\mathtt{sch})$ would be derivable using, e.g., an instance of the rule

$$\frac{\begin{array}{c} B \Rightarrow \top \quad \mathtt{agn} \Rightarrow A \quad \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \mathtt{agn}, A \Rightarrow \quad \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} B \Rightarrow \mathtt{sdr} \\ B \Rightarrow \top \quad \top \Rightarrow \top \quad \mathtt{agn} \Rightarrow A \quad \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} B \Rightarrow \mathtt{sdr} \end{array}}{\Rightarrow \mathcal{O}(A/B)}$$

$$
\dfrac{
\begin{array}{ll}
\{B \Rightarrow B_i\} \quad \cup \quad \{A_i \Rightarrow A\} & \\
\cup & \left\{ \left(\begin{array}{l} \{\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} B \Rightarrow B_j\} \\ \{\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} B_j \Rightarrow B_i\} \\ \{\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} A_j, A \Rightarrow \ \} \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_j/B_j) \in \mathfrak{L} \right\} \\
\cup & \left\{ \left(\begin{array}{l} \{\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} B \Rightarrow B_k\} \\ \{\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} A_k, A \Rightarrow \ \} \\ \{B \Rightarrow B_1\} \cup \{B_1 \Rightarrow B_k\} \cup \{A_1 \Rightarrow A\} \\ \vdots \\ \{B \Rightarrow B_n\} \cup \{B_n \Rightarrow B_k\} \cup \{A_n \Rightarrow A\} \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_k/B_k) \in \mathfrak{L} \right\}
\end{array}
}{\Rightarrow \mathcal{O}(A/B)} \ \mathsf{ga}_R
$$

$$
\dfrac{
\begin{array}{ll}
\{D \Rightarrow B_i\} \quad \cup \quad \{A_i, C \Rightarrow \ \} & \\
\cup & \left\{ \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} D \Rightarrow B_j \\ \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} B_j \Rightarrow B_i \\ \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} A_j \Rightarrow C) \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_j/B_j) \in \mathfrak{L} \right\} \\
\cup & \left\{ \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} D \Rightarrow B_j \\ \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} A_j \Rightarrow C \\ \{D \Rightarrow B_1\} \cup \{B_1 \Rightarrow B_j\} \cup \{A_1, C \Rightarrow \ \} \\ \vdots \\ \{D \Rightarrow B_n\} \cup \{B_n \Rightarrow B_j\} \cup \{A_n, C \Rightarrow \ \} \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_j/B_j) \in \mathfrak{L} \right\}
\end{array}
}{\mathcal{O}(C/D) \Rightarrow} \ \mathsf{ga}_L
$$

Figure 5: The rules of $\mathsf{ga}_{\mathfrak{L}}$ for $\mathfrak{L} = \{\mathcal{O}_{\mathsf{pf}}(A_1/B_1), \ldots, \mathcal{O}_{\mathsf{pf}}(A_n/B_n)\}$, with $i = 1, \ldots, n$.

Similarly, taking the formula $\mathcal{O}_{\mathsf{pf}}(A_i/B_i)$ to be $\mathcal{O}_{\mathsf{pf}}(\neg\texttt{agn}/\texttt{sdr})$ we obtain, e.g.

$$
\dfrac{
\begin{array}{c}
B \Rightarrow \texttt{sdr} \quad \neg\texttt{agn} \Rightarrow A \quad \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \top \Rightarrow \texttt{sdr} \quad \mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \neg\texttt{agn}, A \Rightarrow \\
B \Rightarrow \texttt{sdr} \quad \texttt{sdr} \Rightarrow \top \quad \neg\texttt{agn} \Rightarrow A \quad B \Rightarrow \texttt{sdr} \quad \texttt{sdr} \Rightarrow \texttt{sdr} \quad \neg\texttt{agn} \Rightarrow A
\end{array}
}{\Rightarrow \mathcal{O}(A/B)}
$$

which serves to derive the sequent $\Rightarrow \mathcal{O}(\neg\texttt{agn}/\texttt{sch} \wedge \texttt{sdr})$. Finally, using ga_L with $\mathcal{O}_{\mathsf{pf}}(\neg\texttt{agn}/\texttt{sdr})$ for the formula $\mathcal{O}_{\mathsf{pf}}(A_i/B_i)$ yields a derivation of $\mathcal{O}(\texttt{agn}/\texttt{sch} \wedge \texttt{sdr}) \Rightarrow$ and thus $\Rightarrow \neg\mathcal{O}(\texttt{agn}/\texttt{sch} \wedge \texttt{sdr})$. Note that even for just two prima-facie obligations we obtain many (often redundant) rules.

The following lemma is useful to shorten derivations.

$$\frac{\begin{array}{l} \begin{array}{l}\{B \Rightarrow \top\} \\ \cup \{\mathtt{agn} \Rightarrow A\}\end{array} \cup \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow \top \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \top \Rightarrow \top \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \mathtt{agn}, A \Rightarrow \end{array}\right) \cup \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow \mathtt{sdr} \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \mathtt{sdr} \Rightarrow \top \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \neg\mathtt{agn}, A \Rightarrow \end{array}\right) \\ \cup \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow \top \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \mathtt{agn}, A \Rightarrow \\ \{B \Rightarrow \top\} \cup \{\top \Rightarrow \top\} \cup \{\mathtt{agn} \Rightarrow A\} \\ \{B \Rightarrow \mathtt{sdr}\} \cup \{\mathtt{sdr} \Rightarrow \top\} \cup \{\neg\mathtt{agn} \Rightarrow A\} \end{array}\right) \\ \cup \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow \mathtt{sdr} \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \neg\mathtt{agn}, A \Rightarrow \\ \{B \Rightarrow \top\} \cup \{\top \Rightarrow \mathtt{sdr}\} \cup \{\mathtt{agn} \Rightarrow A\} \\ \{B \Rightarrow \mathtt{sdr}\} \cup \{\mathtt{sdr} \Rightarrow \mathtt{sdr}\} \cup \{\neg\mathtt{agn} \Rightarrow A\} \end{array}\right) \end{array}}{\Rightarrow \mathcal{O}(A/B)}$$

Figure 6: The rules from Ex. 3.5.

Lemma 3.6. *For every formula A we have $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \Gamma, A \Rightarrow A, \Delta$.*

Proof. By straightforward induction on the complexity of A, using the rule Mon in the modal case. □

We now show that the rule ga_L indeed is a mere technical convenience.

Lemma 3.7. *If there is a valid proto-derivation of $\Gamma \Rightarrow \Delta$ in $\mathsf{G_{MD}ga_{\mathfrak{L}}cut}$ from $\mathfrak{F}$, then there is a valid proto-derivation of $\Gamma \Rightarrow \Delta$ from $\mathfrak{F}$ in the same system without the rule ga_L.*

Proof. We show how to replace every application a rule ga_L by an application of ga_R and cut. Suppose we have an application of ga_L as given in Fig. 5. From the premises $A_j, C \Rightarrow$ (if any) using weakening and the $\to_R$ rule we obtain $A_j \Rightarrow C \to \bot$. Further, from every underivability statement $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j \Rightarrow C$ we obtain $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j, C \to \bot \Rightarrow$, since, if for the latter there were a valid proto-derivation, we could extend it to one of the former via

$$\dfrac{\dfrac{\dfrac{}{C \Rightarrow C, \bot}\ \text{Lem. 3.6}}{\Rightarrow C, C \to \bot}\ \to_R \qquad A_j, C \to \bot \Rightarrow}{A_j \Rightarrow C}\ \mathsf{cut}$$

But then we have all the premises necessary to apply the rule ga_R with conclusion $\Rightarrow \mathcal{O}(C \to \bot/D)$. From this we obtain the conclusion of the application of ga_L as follows:

$$\dfrac{\Rightarrow \mathcal{O}(C \to \bot/D) \qquad \dfrac{\dfrac{\dfrac{}{\bot, C \Rightarrow}\,\bot_L \qquad \dfrac{}{C \Rightarrow C}\,\text{Lem. 3.6}}{C \to \bot, C \Rightarrow}\,\to_L \qquad \dfrac{}{D \Rightarrow D}\,\text{Lem. 3.6}}{\mathcal{O}(C \to \bot/D), \mathcal{O}(C/D) \Rightarrow}\,\mathsf{D}}{\mathcal{O}(C/D) \Rightarrow}\,\mathsf{cut}$$

□

In order to unravel the definition of valid proto-derivations and to be able to provide a decision procedure, we show the redundancy (actually the eliminability) of the cut rule in valid proto-derivations. First we obtain:

Lemma 3.8. *If* $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}}} \Gamma \Rightarrow \Delta$*, then* $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \Gamma \Rightarrow \Delta$*.*

Proof. Straightforward since every rule in $\mathsf{G_{MD}}$ is a rule in $\mathsf{G_{MD}cut}$, and since the underivability statements range over the same system for valid proto-derivations in both $\mathsf{G_{MD}ga_{\mathfrak{L}}}$ and $\mathsf{G_{MD}ga_{\mathfrak{L}}cut}$ □

In order to fully control the underivability statements involved in the notion of a valid derivation, we further need to show the converse of this statement.

Theorem 3.9 (Partial cut elimination)**.** *If* $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \Gamma \Rightarrow \Delta$*, then* $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}}} \Gamma \Rightarrow \Delta$*.*

Proof. We show how to eliminate topmost applications of the *multicut rule*

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

from a proto-derivation, preserving validity (here A^n is the multiset containing n copies of A). Since cut is a case of mcut and mcut is derivable using $\mathsf{Con}_L, \mathsf{Con}_R$ and cut, this suffices. The proof is by double induction on the complexity of the cut formula A and the sum of the depths of the derivations of the two premises of the application of mcut (see [26, Sec. 4.1.9] for the classical case without underivability statements).

If the complexity of the cut formula is 0, then it is a propositional variable, and hence not principal in a modal or propositional rule or a rule from $\mathsf{ga}_{\mathfrak{L}}$. Thus, as usual, we permute mcut into the premises of the last applied rules using the inner induction on the depths of the derivations, until it is absorbed by an application of weakening, or reaches the leaves of the proto-derivation. In this case the premises of the multicut are initial sequents or elements of $\mathfrak{F}$. If at least one of these is an initial sequent, the multicut is eliminated as usual, if both sequents are elements of $\mathfrak{F}$ we use that $\mathfrak{F}$ is closed under contraction and cuts and replace the multicut with the corresponding element of $\mathfrak{F}$.

So assume that the complexity of the cut formula is $n+1$. Again, using the inner induction on the depth of the proto-derivation we permute the multicut into the premise(s) of the last applied rules, until it is principal in the last rules of the derivations of both premises of the multicut. In case the cut formula is propositional we use the standard transformation, see [26].

The only interesting case is where the cut formula is a deontic formula. If the last applied rules both are among $\mathsf{P}, \mathsf{D}, \mathsf{Mon}$, then the transformation is essentially as for the system $\mathsf{G_{MD}}$. E.g., if the last applied rules were Mon and D, the multicut has the following form:

$$\dfrac{\dfrac{C \Rightarrow A \quad D \Rightarrow B \quad B \Rightarrow D}{\mathcal{O}(C/D) \Rightarrow \mathcal{O}(A/B)}\,\mathsf{Mon} \qquad \dfrac{A, E \Rightarrow \quad B \Rightarrow F \quad F \Rightarrow B}{\mathcal{O}(A/B), \mathcal{O}(E/F) \Rightarrow}\,\mathsf{D}}{\mathcal{O}(C/D), \mathcal{O}(E/F) \Rightarrow}\,\mathsf{mcut}$$

Using the induction hypothesis on the complexity of the cut formula we obtain valid proto-derivations of the conclusions of

$$\dfrac{C \Rightarrow A \quad A, E \Rightarrow}{C, E \Rightarrow}\,\mathsf{mcut} \qquad \dfrac{D \Rightarrow B \quad B \Rightarrow F}{D \Rightarrow F}\,\mathsf{mcut} \qquad \dfrac{F \Rightarrow B \quad B \Rightarrow D}{F \Rightarrow D}\,\mathsf{mcut}$$

Now an application of the rule D yields the sequent $\Gamma, \mathcal{O}(C/D), \Sigma, \mathcal{O}(E/F) \Rightarrow \Delta, \Pi$. In case both principal formulas of the application of D are cut formulas, we proceed similarly, only using the rule P in the last step. The other cases of the modal rules are similar.

In the most interesting cases at least one of the premises of the cut was derived using a rule from $\mathsf{ga}_{\mathfrak{L}}$. We consider all the different cases.

Suppose that the two last applied rules were ga_R and Mon. Then the two derivations end in an instance of a rule from

$$\dfrac{\begin{array}{ll}\{B \Rightarrow B_i\} \quad \cup \quad \{A_i \Rightarrow A\} & \\ \cup \left\{ \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_j \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B_j \Rightarrow B_i \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j, A \Rightarrow \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_j/B_j) \in \mathfrak{L} \right\} & \\ \cup \left\{ \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_k \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_k, A \Rightarrow \\ \{B \Rightarrow B_1\} \cup \{B_1 \Rightarrow B_k\} \cup \{A_1 \Rightarrow A\} \\ \vdots \\ \{B \Rightarrow B_n\} \cup \{B_n \Rightarrow B_k\} \cup \{A_n \Rightarrow A\} \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_k/B_k) \in \mathfrak{L} \right\} & \end{array}}{\Rightarrow \mathcal{O}(A/B)} \,\mathsf{ga}_R \tag{1}$$

and

$$\dfrac{A \Rightarrow C \quad B \Rightarrow D \quad D \Rightarrow B}{\mathcal{O}(A/B) \Rightarrow \mathcal{O}(C/D)} \,\mathsf{Mon}$$

respectively. By induction hypothesis on the complexity of the cut formula we obtain valid proto-derivations of $D \Rightarrow B_i$ and $A_i \Rightarrow C$, as well as for $1 \leq \ell \leq n$ the sequents $D \Rightarrow B_\ell$ and $B_\ell \Rightarrow B_k$ and $A_\ell \Rightarrow C$ whenever the corresponding sequents occur in the application of ga_R. Further, for every underivability statement $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_j$ together with derivability of $B \Rightarrow D$ we obtain the underivability statement $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} D \Rightarrow B_j$ by contradiction: assuming there is a valid proto-derivation of $D \Rightarrow B_j$ in $\mathsf{G_{MD}ga_{\mathfrak{L}}cut}$ from $\mathfrak{F}$ we could apply cut to this and $B \Rightarrow D$ to obtain $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_j$, in contradiction to $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_j$. Similarly, for every underivability statement $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j, A \Rightarrow$ using derivability of $A \Rightarrow C$ we obtain the underivability statement $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j, C \Rightarrow$. Hence we can apply the rule ga_R to obtain a proto-derivation of $\Rightarrow \mathcal{O}(C/D)$. By the reasoning above, all the underivability statements hold, hence the proto-derivation is valid.

The cases where the two last applied rules were ga_R and D with only one of the principal formulas a cut formula or Mon and ga_L are similar, in each case finishing with an application of ga_L.

For the case where the last rules were ga_R and P, we claim that it actually cannot occur. For otherwise the derivations end in an instance of (1) and

$$\dfrac{A \Rightarrow}{\mathcal{O}(A/B) \Rightarrow} \,\mathsf{P} \;.$$

However, then for $i = j$ we have valid proto-derivations for all three of $B \Rightarrow B_j$ and $B_j \Rightarrow B_i$ and $A_j, A \Rightarrow$. The first one is the first premise of the application of ga_R, the second one follows from Lem. 3.6 since $i = j$, and the last one follows from the premise of P using W_L. But then the proto-derivation of $\Rightarrow \mathcal{O}(A/B)$ cannot have been valid since for some of the underivability statements in the premises of the rule ga_R there is a valid proto-derivation.

The case where the last rules were ga_R and D with both principal formulas of the latter cut formulas is analogous to the previous case.

This leaves the case where the last rules were ga_R and ga_L, which likewise cannot happen. For suppose it did, then the derivations would end in (1) and

$$\dfrac{\begin{array}{ll} \{B \Rightarrow B_k\} \quad \cup \quad \{A_k, A \Rightarrow \ \} & \\ \cup & \left\{ \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_\ell \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B_\ell \Rightarrow B_k \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_\ell \Rightarrow A) \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_\ell/B_\ell) \in \mathfrak{L} \right\} \\ \cup & \left\{ \left(\begin{array}{l} \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_\ell \\ \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_\ell \Rightarrow A \\ \{B \Rightarrow B_1\} \cup \{B_1 \Rightarrow B_\ell\} \cup \{A_1, A \Rightarrow \ \} \\ \vdots \\ \{B \Rightarrow B_n\} \cup \{B_n \Rightarrow B_\ell\} \cup \{A_n, A \Rightarrow \ \} \end{array} \right) \mid \mathcal{O}_{\mathsf{pf}}(A_\ell/B_\ell) \in \mathfrak{L} \right\} \end{array}}{\mathcal{O}(A/B) \Rightarrow} \ \mathsf{ga}_L$$

But then in particular the application of the rule ga_R has one of the premises $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_k$ and $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_k, A \Rightarrow$ or all of the three premises

$$B \Rightarrow B_m \qquad B_m \Rightarrow B_k \qquad A_m \Rightarrow A$$

for some $m \leq n$. However, the first case gives a contradiction with the premise $B \Rightarrow B_k$ of the application of ga_L using validity of the proto-derivation. The second case gives a contradiction with the premise $A_k, A \Rightarrow$ of ga_L, again using validity of the proto-derivation. Finally, the third case gives a contradiction because the application of ga_L contains one of the premises

$$\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B \Rightarrow B_m \qquad \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} B_m \Rightarrow B_k \qquad \mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_m \Rightarrow A$$

and the proto-derivation is valid. Hence this case also cannot occur. □

3.2 Applications of cut elimination

Thm. 3.9 is the basis for a number of important results. First and foremost we obtain that the notion of valid proto-derivations (Def. 3.4) actually makes sense.

Theorem 3.10. *The notion of a valid proto-derivation is well-defined.*

Proof. Cut elimination together with Lem. 3.8 shows that we can replace every underivability statement in the rules $\mathsf{ga}_L, \mathsf{ga}_R$ by a statement of the form $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga}_{\mathfrak{L}}} \Sigma \Rightarrow \Pi$ and define valid proto-derivations in terms of cut-free derivations. Since the modal nesting depth properly decreases in the modal rules, the definition of a valid proto-derivation is hence equivalent to a *stratified definition*, where we define the notion of a valid proto-derivation of *rank* n such that every sequent occurring in the derivation has modal nesting depth at most n, and all the underivability statements refer to underivability using valid proto-derivations of rank smaller than n. $\square$

The proof of the previous theorem serves to illustrate one of the main differences between the approach followed here and approaches which model conditional obligations using non-monotone inference or defeasible rules such as [13, 18]: since the underivability statements in the premisses of a rule can be restricted to sequents of smaller modal nesting depth, we can avoid having to perform a fixed-point computation. In particular, for checking whether a non-nested conditional obligation formula is derivable we only need to check (classical) derivability for purely propositional sequents. This is possible because the underivability statements in the rules $\mathsf{ga}_L, \mathsf{ga}_R$ only depend on the list of prima-facie obligations, and not on obligations which themselves are derivable. Consequently, we obtain decidability of the logic:

Theorem 3.11 (Decidability)**.** *The set of all sequents for which there is a valid proto-derivation in* $\mathsf{G_{MD}ga}_{\mathfrak{L}}\mathsf{cut}$ *from* $\mathfrak{F}$ *is decidable.*

Proof. First we show that a sequent has a valid proto-derivation in $\mathsf{G_{MD}ga}_{\mathfrak{L}}$ if and only if it has a valid proto-derivation using rules of the system $\mathsf{G^*_{MD}ga}_{\mathfrak{L}}$, which is obtained from $\mathsf{G_{MD}ga}_{\mathfrak{L}}$ by dropping the contraction rules $\mathsf{Con}_L, \mathsf{Con}_R$ and replacing the propositional rules $\to_L, \to_R$ with their invertible versions where the principal formula is copied into the premises:

$$\frac{\Gamma, A \to B, B \Rightarrow \Delta \quad \Gamma, A \to B \Rightarrow A, \Delta}{\Gamma, A \to B \Rightarrow \Delta} \to^*_L \qquad \frac{\Gamma, A \Rightarrow B, A \to B, \Delta}{\Gamma \Rightarrow A \to B, \Delta} \to^*_R$$

Equivalence of the systems is obtained by first showing equivalence of the propositional rules and their versions above in the presence of weakening and contraction, and then showing that the contraction rules are admissible in $\mathsf{G^*_{MD}ga}_{\mathfrak{L}}$.

The proof of the latter is, as usual, by induction on the depth of the proto-derivation, using that the set $\mathfrak{F}$ of facts is closed under contraction: If the depth is 1, then, the proto-derivation consists of an initial sequent or a sequent from $\mathfrak{F}$. Since $\mathfrak{F}$ is assumed to be closed under contractions, the admissibility follows. If the depth is $n > 1$ we distinguish cases according to the last applied rule in the proto-derivation. The only non-trivial case is when that rule was one of the modal rules, i.e. the rule D, since none of the other rules has two formulae on the same side of the conclusion. In this case, the proto-derivation ends in

$$\frac{A, A \Rightarrow \quad B \Rightarrow B \quad B \Rightarrow B}{\mathcal{O}(A/B), \mathcal{O}(A/B) \Rightarrow}\ \mathsf{D}$$

Applying the induction hypothesis on the proto-derivation of the first premise we obtain a proto-derivation of the sequent $A \Rightarrow$, and an application of the rule P yields the desired sequent $\mathcal{O}(A/B) \Rightarrow$.

To check whether for a given sequent there is a valid proto-derivation in the system $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}$ from $\mathfrak{F}$, by Thm. 3.9 it is enough to search through the possible proto-derivations in $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}$, which by the previous considerations is equivalent to searching through proto-derivations in $\mathsf{G}^{*}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}$. For the latter we perform (depth-first) *backwards proof search*, following a *local loop checking* strategy to prevent rule applications where every formula of a premise already occurs in the conclusion. Upon encountering an underivability statement the decision procedure calls itself recursively and simply flips the answer. The procedure terminates, because the modal nesting depth of the underivability statements in the premises of $\mathsf{ga}_L, \mathsf{ga}_R$ is lower than that of the conclusion, the rules have bounded branching factor and the local loop checking strategy implies that the proto-derivations themselves have bounded depth. □

Furthermore, we obtain that the rules $\mathsf{ga}_{\mathfrak{L}}$ are compatible with deontic logic MD in the sense that they do not yield any conflicting obligations:

Theorem 3.12 (Consistency). *For any $\mathfrak{L}$ and $\mathfrak{F}$ not containing the empty sequent, the consequences of $\mathfrak{L}$ under $\mathfrak{F}$ are consistent over* MD*, i.e., $\mathfrak{F} \nvdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \Rightarrow \bot$. Hence there is no $\mathcal{O}(A/B)$ with $\mathfrak{F} \vdash_{\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}\mathsf{cut}} \Rightarrow \mathcal{O}(A/B) \wedge \mathcal{O}(\neg A/B)$.*

Proof. By inspection it is clear that all the rules in the calculus $\mathsf{G}_{\mathsf{MD}}\mathsf{ga}_{\mathfrak{L}}$ have the *subformula property relative to* $\mathfrak{L}$ in the sense that every formula occurring in a premise of a rule, including the underivability statements, is a subformula of a formula occurring in its conclusion or in $\mathfrak{L}$. Since the empty sequent is not in $\mathfrak{F}$, and apart from W_R there is no rule introducing $\bot$ on the right hand side of a sequent, we cannot derive $\Rightarrow \bot$. The second statement follows from derivability of $\mathcal{O}(A/B) \wedge \mathcal{O}(\neg A/B) \Rightarrow$ and cut. □

3.3 The disjunctive response/Vikalpa

The described system rejects any inferences which would result in conflicting obligations. In particular, for a set $\mathfrak{L} = \{\mathcal{O}_{\mathsf{pf}}(a/b), \mathcal{O}_{\mathsf{pf}}(c/d)\}$ and a set $\mathfrak{F} = \{a, c \Rightarrow\ \}$

establishing that a and c are not jointly possible, neither of the formulas $\mathcal{O}(a/b \wedge d)$ and $\mathcal{O}(c/b \wedge d)$ will be derivable. While this is as intended, intuitively still the disjunction of a and c should be obligatory, i.e., the formula $\mathcal{O}(a \vee c/b \wedge d)$ should be derivable, and similarly for sets of formulas $\{\mathcal{O}_{\mathsf{pf}}(a_1/b_1), \ldots, \mathcal{O}_{\mathsf{pf}}(a_n/b_n)\}$ where all the a_i are not jointly possible. Amazingly, this principle, which is called the *disjunctive response* in [10], was formulated already more than two millennia ago in one of the founding texts of the Mīmāṃsā school, the *Pūrva Mīmāṃsā Sūtras* of Jaimini under the name of *vikalpa*, whose English translation (and reformulation) is

> If a prescription enjoins X and a prohibition forbids one to perform the same act X, and no other interpretation is possible, the act X should be considered optional (*vikalpa*), although this leads to the problematic situation that either the one or the other is transgressed.

Thus, checking whether our rendering of the specificity principle satisfies this principle is a good test for checking suitability to capture both the intuitive and the Mīmāṃsā notion of obligation. Indeed, generalizing the above to sets of obligations, and adding that all the enjoined acts should be possible we have:

Theorem 3.13. *Let $X = \{\mathcal{O}_{\mathsf{pf}}(A_1/B_1), \ldots, \mathcal{O}_{\mathsf{pf}}(A_n/B_n)\} \subseteq \mathfrak{L}$ be a set such that $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_i \Rightarrow$ for every $i \leq n$, and for every $\mathcal{O}_{\mathsf{pf}}(C/D) \in \mathfrak{L} \setminus X$ with $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \bigwedge_{i\leq n} B_i \Rightarrow D$ we have $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \bigvee_{i\leq n} A_i, C \Rightarrow$. Then $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \Rightarrow \mathcal{O}(\bigvee_{i\leq n} A_i / \bigwedge_{i\leq n} B_i)$.*

Proof. We show that we have all the premises to apply the rule ga_R. From the propositional rules we obtain $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_1 \Rightarrow \bigvee_{i\leq n} A_i$ and $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \bigwedge_{i\leq n} B_i \Rightarrow B_1$. Moreover, for every $j \leq n$ we obtain $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j, \bigvee_{i\leq n} A_i \Rightarrow$, since otherwise in particular we would have $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j, A_j \Rightarrow$, and hence $\mathfrak{F} \vdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} A_j \Rightarrow$. Moreover, by assumption, for every $\mathcal{O}_{\mathsf{pf}}(C/D) \in \mathfrak{L} \setminus X$ we have either $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} \bigwedge_{i\leq n} B_i \Rightarrow D$ or $\mathfrak{F} \nvdash_{\mathsf{G_{MD}ga_{\mathfrak{L}}cut}} C, \bigvee_{i\leq n} A_i \Rightarrow$. Now applying ga_R yields $\Rightarrow \mathcal{O}(\bigvee_{i\leq n} A_i / \bigwedge_{i\leq n} B_i)$. □

It should be noted that for the statement of the theorem it is not relevant whether the A_i from the set X are jointly possible or not, only that their disjunction $\bigvee_{i\leq m} A_i$ is not blocked by any C from outside that set. In particular, it also applies to the case where the A_i are not jointly possible. Thus, our system as described indeed satisfies the disjunctive response resp. vikalpa.

4 Conclusion

We have explored connections between the Mīmāṃsā school of Indian philosophy and symbolic deontic logic concerning the *specificity principle.* We investigated a notion of specificity based on a sequent calculus for MD and explored some of its properties. Apart from the technical content, this paper illustrates some of the vast potential for cross-fertilisation between Mīmāṃsā and deontic logic. Of course, many aspects both of the Mīmāṃsā philosophy and of the proposed formal system are still waiting to be unearthed; among them, how to modify our rules $\mathsf{ga}_{\mathfrak{L}}$ to accommodate for new deontic axioms extracted from *nyāya*s that might be added to bMDL. We further conjecture that it is possible to show a completeness result with respect to the class of neighbourhood models under a certain set of global assumptions along the lines of [11]; due to the infinite nature of the set of global assumptions this is not entirely straightforward. In view of Remark 2.3 it would also be interesting to see whether our approach can be generalized to handle nested prima-facie obligations. We also plan to implement the introduced calculus and try to use it to prove Mīmāṃsā conjectures about Vedic sacrifices; e.g., whether the so-called Full and New Moon sacrifice is the archetype of all vegetable sacrifices.

From a more philosophical point of view, we recall the extensive discussion in the deontic literature about the difference between Contrary-To-Duty obligations and instances of the specificity principle [28] and about the connections of this topic with the problem of factual detachment [2, 22]. It would be interesting to see whether our proposed methods and calculi can be used to analyze more closely where the Mīmāṃsā authors would be situated in this debate.

References

[1] K. L. Bathia. *Legal Language and Legal Writing.* Universal Law Publishing Co, 2010.

[2] M. Beirlaen, J. Heyninck, and C. Straßer. Structured argumentation with prioritized conditional obligations and permissions. *Journal of Logic and Computation*, 2018.

[3] P. A. Bonatti and N. Olivetti. Sequent calculi for propositional nonmonotonic logics. *ACM Transactions on Computational Logic*, 3(2):226–278, 2002.

[4] B. F. Chellas. *Modal Logic.* Cambridge University Press, 1980.

[5] A. Ciabattoni, E. Freschi, F. A. Genco, and B. Lellmann. Mīmāṃsā deontic logic: proof theory and applications. In *TABLEAUX 2015*, volume 9323 of *LNCS*, pages 323–338. Springer, 2015.

[6] A. Ciabattoni, E. Freschi, F. A. Genco, and B. Lellmann. Understanding prescriptive texts: Rules and logic as elaborated by the Mīmāṃsā school. *Online Journal of World Philosophies*, 2:47–66, 2017.

[7] J. P. Delgrande and T. H. Schaub. Compiling specificity into approaches to nonmonotonic reasoning. *Artificial Intelligence*, 90(1):301 – 348, 1997.

[8] E. Freschi, A. Ollett, and M. Pascucci. Duty and sacrifice. a logical analysis of the Mīmāṃsā theory of Vedic injunctions. Submitted, 2018.

[9] D. M. Gabbay. Theoretical foundations for non-monotonic reasoning in expert systems. In *Logics and models of concurrent systems*, pages 439–457. Springer, 1985.

[10] L. Goble. Prima facie norms, normative conflicts, and dilemmas. In *Handbook of Deontic Logic and Normative Systems*, pages 241–351. College Publications, 2013.

[11] R. Goré, C. Kupke, and D. Pattinson. Optimal tableau algorithms for coalgebraic logics. In *TACAS 2010*, LNCS, pages 114–128. Springer, 2010.

[12] G. Governatori and A. Rotolo. Logic of violations: A Gentzen system for reasoning with contrary-to-duty obligations. *The Australasian Journal of Logic*, 4:193–215, 2006.

[13] J. F. Horty. *Reasons as Defaults.* Oxford University Press, 2012.

[14] C. H. Huisjes. *Norms and logic.* Thesis, University of Groningen, 1981.

[15] D. Kapoor. `www.patheos.com/blogs/drishtikone/2008/03/indian-supreme-court-uses-mimamsa-school-legal-reasoning-decide-case/`, 2008. Accessed 2018-03-12.

[16] B. Lellmann and D. Pattinson. Constructing cut free sequent systems with context restrictions based on classical or intuitionistic logic. In *ICLA 2013*, volume 7750 of *LNCS*, pages 148–160. Springer, 2013.

[17] D. Makinson and K. Schlechta. Floating conclusions and zombie paths: Two deep difficulties in the 'directly skeptical' approach to inheritance nets. *Artificial Intelligence*, 48:199–209, 1991.

[18] D. Nute. Defeasible logic. In *INAP 2001*, volume 2543 of *LNCS*, pages 151–169. Springer, 2003.

[19] E. Orlandelli. Proof analysis in deontic logics. In *DEON 2014*, volume 8554 of *LNCS*, pages 139–148. Springer, 2014.

[20] X. Parent. Maximality vs. optimality in dyadic deontic logic. *Journal of Philosophical Logic*, 43:1101–1128, 2013.

[21] H. Prakken and G. Sartor. A system for defeasible argumentation, with defeasible priorities. In *FAPR'96*, volume 1085 of *LNCS*, pages 510–524. Springer, 1996.

[22] C. Straßer. A deontic logic framework allowing for factual detachment. *Journal of Applied Logic*, 9(1):61–80, 2011.

[23] C. Straßer and A. Antonelli. Non-monotonic logic. In *The Stanford Encyclopedia of Philosophy.* Stanford University, 2016.

[24] C. Straßer and O. Arieli. Sequent-based argumentation for normative reasoning. In *DEON 2014*, volume 8554 of *LNCS*, pages 224–240. Springer, 2014.

[25] J. Taber. Is Indian logic nonmonotonic? *Philosophy East and West*, 54:143–170, 2004.

[26] A. S. Troelstra and H. Schwichtenberg. *Basic Proof Theory.* Cambridge University Press, 2nd edition, 2000.

[27] L. W. N. van der Torre. Violated obligations in a defeasible deontic logic. In *ECAI 94*, pages 371–375. Wiley, 1994.

[28] L. W. N. van der Torre and Y. Tan. The many faces of defeasibility in defeasible deontic logic. In *Defeasible Deontic Logic*, pages 79–121. Kluwer, 1997.

Received \jreceived