

The Gentle Murder Paradox in Sanskrit Philosophy

Kees van Berkel^{*1} Agata Ciabattoni^{*1} Elisa Freschi^{**2}
Francesca Gulisano^{*3} Maya Olszewski^{*1}

Institute of Logic and Computation, TU Wien^{}
Vienna, Austria*

*Department of Philosophy, University of Toronto^{**}
Toronto, Canada*

Abstract

For decades, the gentle murder paradox has been a central challenge for deontic logic. This article investigates its millennia-old counterpart in Sanskrit philosophy: the *śyena* controversy. We analyze three solutions provided by Mīmāṃsā, the Sanskrit philosophical school devoted to the analysis of normative reasoning in the Vedas, in which the controversy originated. We introduce axiomatizations and semantics for the modal logics formalizing the deontic theories of the main Mīmāṃsā philosophers Prabhākara, Kumāṛila, and Maṇḍana. The resulting logics are used to analyze their distinct solutions to the *śyena* controversy, which we compare with formal approaches developed within the contemporary field of deontic logic.

Keywords: Mīmāṃsā, Dyadic Deontic Logic, Instruments, Gentle Murder Paradox

1 Introduction

Introduced by Forrester [9], the *Gentle Murder Paradox* (GMP) is a well-known problem for monadic deontic logic [13,28], motivating the use of alternative systems employing dyadic deontic operators, e.g., [16,21]. The GMP in a nutshell: (i) x is obliged not to kill, (ii) if x kills, x is obliged to kill gently, (iii) gentle killing implies killing, and (iv) x will kill. Although intuitively consistent, the sentences (i)-(iv) lead to a contradiction in Standard Deontic Logic, implying x 's obligation to kill. Originally, the GMP was introduced as a stronger Good Samaritan Paradox [24], but it is commonly taken as a variant of Chisholm's Paradox [6]. Under the former reading, (i)-(iv) imply conflicting obligations (i.e., a dilemma), inconsistent under normality of deontic operators. Under the latter reading, the GMP relates to challenges of reasoning with violations and

¹ {kees, agata, maya}@logic.at

² elisa.freschi@utoronto.ca

³ f.gulisano2201@gmail.com

contrary-to-duty (CTD) obligations (i.e., the obligation (ii) is only in force if (i) is violated). In fact, the GMP has features of both paradoxes [17].

While the GMP was introduced to the deontic logic community only a few decades ago, a similar example has been thoroughly investigated in Sanskrit philosophy for more than two millennia. This is the renowned *śyena* controversy. The *śyena* is a one-day long ritual in which the Soma beverage is offered. Its putative result is the death of the sacrificer’s enemy. Unlike animal sacrifices it does not involve violence in its performance, violence is only found in its result. The controversy is due to the fact that the *śyena* appears to be prescribed in the Vedas—the sacred texts of what is now known as “Hinduism”—, which also prohibits the performance of violence. The *śyena* controversy in short⁴:

- (A) The one who desires to kill their enemy should sacrifice with the *śyena*
- (B) One should not harm any living being
- (C) Performing *śyena* implies causing someone’s death
- (D) Causing someone’s death implies harming

With (A)-(D), the Vedas seem to provide contradicting commands concerning the performance of violence, a possibility which is ruled out by the (indisputable) claim that the Vedas are consistent.

The Sanskrit philosophical school of Mīmāṃsā—which flourished between the last centuries BCE and the 20th c. CE—paid exceptional attention to the controversy, explaining why the *śyena* should not be performed and why the sacred texts prescribing it are not contradictory. In general, the Mīmāṃsā school focused on the rational interpretation and systematization of the prescriptive portions of the Vedas. To reason with Vedic commands, and resolve seeming conflicts, the Mīmāṃsā developed a vast system of theories containing rigorous analyses of deontic concepts. Key to their enterprise was the formulation of general reasoning principles called *nyāyas*, and the distinction among elective duties (to be performed only if one wishes their specific result), fixed duties (to be performed no matter what), and prohibitions. The resulting theories, which have been extremely influential in Sanskrit philosophy, theology and law, provide an inexhaustible resource for deontic investigation, largely still unexplored.

Although all Mīmāṃsā authors agree that *śyena* should not be performed, they disagree on the reasons underlying it. In this article, we focus on the three main Mīmāṃsā authors: Kumāṛila, Prabhākara (both ca. 7th c. CE), and Maṇḍana (ca. 8th c. CE). They are known for their distinctive deontic theories, which give rise to different interpretations of Vedic commands. Likewise, their solutions to the *śyena* controversy are markedly distinct.

We provide three modal logics⁵ describing the deontic theories of these

⁴ (A) and (B) are direct translations from Sanskrit, whereas (C) and (D) are derived from Mīmāṃsā arguments about the *śyena*.

⁵ The logics in this paper are intended to reason about commands as interpreted by the Mīmāṃsā. Since the Vedas are self-contained and immutable, new Vedic commands cannot

authors, whose rational, structured approach makes their accounts particularly suitable for formalization. The resulting logics are obtained by “extracting” Hilbert axioms out of translated and parsed Mīmāṃsā *nyāyas* and additional passages by the three authors. While the logics for Prabhākara and Kumārila are a modification of those presented in [7] and [20], Maṇḍana’s logic is novel.

The main contributions of this paper are threefold: First, we develop a logic formalizing Maṇḍana’s deontic theory. His account is particularly noteworthy due to its *deontic reduction*: the reduction of all commands of the Vedas (i.e., fixed and elective obligations, as well as prohibitions) to mere descriptive statements of instrumentality. For instance, according to Maṇḍana, an obligation to perform an action means the action is an instrument for attaining a certain result. The introduced logic reproduces Maṇḍana’s reduction by adopting a PDL-like language [8,23], together with a modified Andersonian reduction of deontic modalities [1]. Second, we offer a consistent formalization of the *śyena* controversy as interpreted by Kumārila and Maṇḍana, faithful to the explanations found in Mīmāṃsā texts. Kumārila’s formalization is achieved by introducing a neighbourhood semantics for its logic. Prabhākara’s solution was formally analyzed in [7], however, the logic presented there contained only obligations. In [12,20] it was shown that obligations and prohibitions in Mīmāṃsā are not inter-definable and, hence, we extend Prabhākara’s logic (and solution) with a prohibition operator. Third, we analyze and compare the three formal solutions to the *śyena* controversy and discuss their relations to approaches in contemporary deontic logic. In particular, the dual reading of the GMP is reflected in the different approaches to the *śyena* controversy: As for the Chisholm paradox, Prabhākara takes the *śyena* prescription as a contrary-to-duty obligation. Kumārila addresses the controversy by interpreting *śyena* as an elective sacrifice, to which he assigns no deontic force. As for the Good Samaritan Paradox, Maṇḍana endorses the view that there is a proper dilemma in the controversy, but addresses it through his reduction, arguing for a pragmatic rational-choice solution based on a cost-benefit analysis of (un)desirable outcomes.

Our work is the first study of the *śyena* controversy in Mīmāṃsā, the school in which the controversy originated. The interest in this controversy from the point of view of modern deontic logic is also testified by the recent work [14], where the *śyena* is analyzed from the perspective of the Navya-Nyāya, a different school of Sanskrit philosophy.⁶

2 Prabhākara and Kumārila

Prabhākara and Kumārila interpret Vedic prescriptions as proper commands with deontic force. Despite their shared view on fixed duties (to be read as obligations) and on prohibitions, Prabhākara and Kumārila disagree on the

be derived through Logic. Accordingly, our logics deal with commands on the derived level.

⁶ The work [14], also relating the *śyena* to the GMP, was published while the present paper was under review.

reading of elective sacrifices, which are always conditioned on a desire.

Prabhākara’s system is eminently deontic: Vedic statements are binding, independently from their conditions; hence, an elective sacrifice is also a type of obligation. The desire for a specific worldly result, necessarily mentioned as the condition of an elective ritual, only represents the requirement through which the eligible agents are identified, but it does not weaken the deontic force of the injunction. By contrast, for Kumārila, elective sacrifices are of a different type, not enjoining any deontic force, and can be omitted without risk. Still, an eligible agent—i.e., an agent who desires the expected result of the sacrifice—feels prompted to undertake the sacrifice due to its presence in the Vedas: such sacrifices represent a “guaranteed” method for obtaining the desired results. Hence, whereas Prabhākara sees elective sacrifices as conditional obligations, Kumārila sees them as a different type of Vedic command.

The two logics presented in this section will reflect this distinction. Since their only difference is the presence of elective sacrifices as a distinct deontic concept, the logic for Prabhākara will be a proper subset of Kumārila’s. However, the distinction causes wholly different solutions to the *śyena* controversy. The logics are variants of the formalism introduced in [20], whose properties were extracted from a collection of general Mīmāṃsā reasoning principles (*nyāyas*, see Sect. 1). By adding the deontic operators for prohibitions $\mathcal{F}(\cdot/\cdot)$ and for injunctions prescribing elective duties $\mathcal{E}(\cdot/\cdot)$, the resulting logics extend the non-normal dyadic deontic logic **bMDL**. Introduced in [7] to formalize the deontic theory of Prabhākara, **bMDL** only contained a single deontic operator $\mathcal{O}(\cdot/\cdot)$ for obligations.

2.1 Deontic logics for Kumārila and Prabhākara

The languages \mathcal{L}_{LP_r} for Prabhākara and \mathcal{L}_{LK_u} for Kumārila are defined through the following BNF (with $\mathcal{X} \in \{\mathcal{O}, \mathcal{F}\}$ for \mathcal{L}_{LP_r} and $\mathcal{X} \in \{\mathcal{O}, \mathcal{F}, \mathcal{E}\}$ for \mathcal{L}_{LK_u}):

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \boxed{\varphi} \mid \mathcal{X}(\varphi/\varphi) \quad \text{with } p \in \text{Atom}$$

Atom is the set of atomic propositions, \neg and \vee are primitive connectives, the others defined as usual. $\boxed{\varphi}$ reads “it is universally necessary that φ ”. The operators $\mathcal{O}(\varphi/\psi)/\mathcal{F}(\varphi/\psi)/\mathcal{E}(\varphi/\psi)$ read as “ φ is obligatory/forbidden/enjoined by an injunction prescribing an elective ritual, given ψ ”.

Axiomatization. The properties of the operators \mathcal{O} , \mathcal{F} , and \mathcal{E} are extracted from the following Mīmāṃsā principles (see [11] for details on how these principles were transformed into axioms for \mathcal{O} , and [20] for the remaining axioms):

- (P1) “If the accomplishment of a task presupposes the accomplishment of another connected but different task, the obligation to perform the first task prescribes also the second one”.
- (P2) “Two actions that exclude each other cannot be prescribed simultaneously to the same group of eligible people under the same conditions”.
- (P3) “If two sets of conditions always identify the same group of eligible agents, then a command valid under the conditions in one of those sets is also enforceable under the conditions in the other set”.

The two logics are described in Definition 2.1. In contrast with bMDL [7], we use S5 to characterize necessity \Box , instead of S4: Note that the concept of necessity is not explicitly defined by Mīmāṃsā authors in the context of deontic reasoning, and the choice of S4 in [7] was motivated by the simpler proof theory of this logic, with respect to S5. In this paper we use necessary statements mainly as global assumptions (assertions commonly recognised as describing “facts”); hence, any assumption defines an equivalence class of states sharing the same truths.

Using the corresponding universal modality of S5 makes the bMDL axiom $\Box((\psi \rightarrow \theta) \wedge (\theta \rightarrow \psi)) \wedge \mathcal{O}(\varphi/\psi) \rightarrow \mathcal{O}(\varphi/\theta)$ redundant (it is derivable by using axiom T and the congruence rule of S5), also in the versions for \mathcal{E} and \mathcal{F} .

Definition 2.1 Prabhākara’s logic LPr extends S5 with the following axioms:

- A_{LKu}1. $(\Box(\varphi \rightarrow \psi) \wedge \mathcal{O}(\varphi/\theta)) \rightarrow \mathcal{O}(\psi/\theta)$
- A_{LKu}2. $(\Box(\varphi \rightarrow \psi) \wedge \mathcal{F}(\psi/\theta)) \rightarrow \mathcal{F}(\varphi/\theta)$
- A_{LKu}3. $\Box(\psi \rightarrow \neg\varphi) \rightarrow \neg(\mathcal{X}(\varphi/\theta) \wedge \mathcal{X}(\psi/\theta))$ for $\mathcal{X} \in \{\mathcal{O}, \mathcal{F}\}$
- A_{LKu}4. $\Box(\varphi \rightarrow \psi) \rightarrow \neg(\mathcal{O}(\varphi/\theta) \wedge \mathcal{F}(\psi/\theta))$
- A_{LKu}5. $(\Box((\psi \rightarrow \theta) \wedge (\theta \rightarrow \psi)) \wedge \mathcal{X}(\varphi/\psi)) \rightarrow \mathcal{X}(\varphi/\theta)$ for $\mathcal{X} \in \{\mathcal{O}, \mathcal{F}\}$

Kumārila’s deontic logic LKu, extends LPr with the following axioms:

- A_{LKu}6. $(\Box(\varphi \rightarrow \psi) \wedge \mathcal{E}(\varphi/\theta)) \rightarrow \mathcal{E}(\psi/\theta)$
- A_{LKu}7. $\Box(\neg\varphi) \rightarrow \neg\mathcal{E}(\varphi/\psi)$
- A_{LKu}8. $(\Box((\psi \rightarrow \theta) \wedge (\theta \rightarrow \psi)) \wedge \mathcal{E}(\varphi/\psi)) \rightarrow \mathcal{E}(\varphi/\theta)$

A derivation of φ in LKu (i.e., $\vdash_{\text{LKu}} \varphi$) is defined as usual [3] (similarly for LPr).

Axioms A_{LKu}1, A_{LKu}2, A_{LKu}6 are based on (P1) and correspond to the property of monotonicity. Axioms A_{LKu}3, A_{LKu}4 formally represent (P2) (found in Kumārila’s *Tantravārtika* ad 1.3.3 [27]). Last, the Mīmāṃsā property (P3) is ensured by A_{LKu}5, A_{LKu}8.

Semantics We present a neighbourhood semantics (e.g., see [5]) for LPr and LKu (resp.), as defined along the lines of the one for bMDL in [7]:

Definition 2.2 An LPr-frame $\mathfrak{F}_{\text{LPr}} = \langle W, R_{\Box}, \mathcal{N}_{\mathcal{O}}, \mathcal{N}_{\mathcal{F}} \rangle$ is a tuple where $W \neq \emptyset$ is a set of worlds, $R_{\Box} = W \times W$ is the universal relation, and $\mathcal{N}_{\mathcal{X}} : W \mapsto \wp(\wp(W) \times \wp(W))$ is a neighborhood function (for $\mathcal{X} \in \{\mathcal{O}, \mathcal{F}\}$). $\mathfrak{F}_{\text{LPr}}$ satisfies:

- (i) If $(X, Z) \in \mathcal{N}_{\mathcal{O}}(w)$ and $X \subseteq Y$, then $(Y, Z) \in \mathcal{N}_{\mathcal{O}}(w)$;
- (ii) If $(X, Y) \in \mathcal{N}_{\mathcal{X}}(w)$, then $(\overline{X}, Y) \notin \mathcal{N}_{\mathcal{X}}(w)$ for $\mathcal{X} \in \{\mathcal{O}, \mathcal{F}\}$;
- (iii) If $(X, Z) \in \mathcal{N}_{\mathcal{F}}(w)$ and $Y \subseteq X$, then $(Y, Z) \in \mathcal{N}_{\mathcal{F}}(w)$;
- (iv) It cannot be the case that $(X, Z) \in \mathcal{N}_{\mathcal{O}}(w)$ and $(X, Z) \in \mathcal{N}_{\mathcal{F}}(w)$.

An LPr-model $\mathfrak{M}_{\text{LPr}} = \langle W, R_{\Box}, \mathcal{N}_{\mathcal{O}}, \mathcal{N}_{\mathcal{F}}, V \rangle$ extends the LPr-frame by a *valuation* function V which maps propositional variables to subsets of W .

Definition 2.3 An LKu-frame $\mathfrak{F}_{\text{LKu}} = \langle W, R_{\Box}, \mathcal{N}_{\mathcal{O}}, \mathcal{N}_{\mathcal{F}}, \mathcal{N}_{\mathcal{E}} \rangle$ is an LPr-frame extended with a neighbourhood function $\mathcal{N}_{\mathcal{E}} : W \mapsto \wp(\wp(W) \times \wp(W))$ s.t.:

- (v) If $(X, Z) \in \mathcal{N}_{\mathcal{E}}(w)$ and $X \subseteq Y$, then $(Y, Z) \in \mathcal{N}_{\mathcal{E}}(w)$;
 (vi) If $(X, Y) \in \mathcal{N}_{\mathcal{E}}(w)$, then $X \neq \emptyset$.

An LKu-model $\mathfrak{M}_{\text{LKu}} = \langle \mathfrak{F}_{\text{LKu}}, V \rangle$ is an LKu-frame with a valuation function V .

Note that (i), (iii), (v) express the property of monotonicity in the first argument of the deontic operators (cf. $A_{\text{LKu}}1, A_{\text{LKu}}2, A_{\text{LKu}}5$); (ii), (iv) correspond to the principle (P2) (cf. $A_{\text{LKu}}3, A_{\text{LKu}}4$), and (vi) expresses the self consistency of statements prescribing elective sacrifices (cf. $A_{\text{LKu}}6$).

Definition 2.4 Let $\mathfrak{M}_{\text{LKu}}$ be an LKu-model and $\|\theta\| = \{w \in W \mid \mathfrak{M}_{\text{LPr}}, w \models \theta\}$. We define the *satisfaction* of a formula $\varphi \in \mathcal{L}_{\text{LKu}}$ at any w of $\mathfrak{M}_{\text{LKu}}$ inductively:

$$\begin{array}{ll} \mathfrak{M}_{\text{LKu}}, w \models p & \text{iff } w \in V_{\text{LPr}}(p), \text{ for any } p \in \text{Atom} \\ \mathfrak{M}_{\text{LKu}}, w \models \varphi \rightarrow \psi & \text{iff } \mathfrak{M}_{\text{LKu}}, w \not\models \varphi \text{ or } \mathfrak{M}_{\text{LKu}}, w \models \psi \\ \mathfrak{M}_{\text{LKu}}, w \models \neg\varphi & \text{iff } \mathfrak{M}_{\text{LKu}}, w \not\models \varphi \\ \mathfrak{M}_{\text{LKu}}, w \models \boxplus\varphi & \text{iff for all } w_i \in W \text{ s.t. } (w, w_i) \in R_{\boxplus}, \mathfrak{M}_{\text{LKu}}, w_i \models \varphi \\ \mathfrak{M}_{\text{LPr}}, w \models \mathcal{X}(\varphi/\psi) & \text{iff } (\|\varphi\|, \|\psi\|) \in \mathcal{N}_{\mathcal{X}}(w) \text{ for } \mathcal{X} \in \{\mathcal{O}, \mathcal{F}, \mathcal{E}\} \end{array}$$

Global truth and validity are defined as usual [3]. Note that *satisfaction* for $\mathfrak{M}_{\text{LPr}}$ -models is defined as for $\mathfrak{M}_{\text{LKu}}$, without the clause for $\mathcal{N}_{\mathcal{E}}(w)$.

Theorem 2.5 (Soundness and completeness) *The logic LKu (LPr) is sound and complete with respect to the class of LKu-frames (LPr – frames).*

Soundness and completeness are proven as usual. The latter is shown using the method of canonical models [5], generalized to the dyadic setting.

2.2 The solutions of Prabhākara and Kumārila

The sentences (A)-(D) comprising the *śyena* controversy (Sect. 1) are formalized in a similar way by the two authors. The only difference is their interpretation of (A), prescribing the *śyena* sacrifice: for Prabhākara this is a conditional obligation (A_P), whereas Kumārila interprets it as an elective sacrifice (A_K). Their formalization:

$$\begin{array}{ll} (A_P) \mathcal{O}(\mathfrak{S}y/\text{des_kill}) & (B) \mathcal{F}(\text{harm}/\top) \\ & (C) \boxplus(\mathfrak{S}y \rightarrow \text{death}) \\ (A_K) \mathcal{E}(\mathfrak{S}y/\text{des_kill}) & (D) \boxplus(\text{death} \rightarrow \text{harm}) \end{array}$$

Fig. 1 shows the models \mathfrak{M}_P and \mathfrak{M}_K demonstrating the mutual satisfiability of (A_P), (B), (C), (D) in LPr and of (A_K), (B), (C), (D) in LKu, respectively, and hence the consistency of the *śyena* controversy for both authors. That is, there is always at least one world in which no command is violated. (A command $\mathcal{O}(\phi/\psi)$ or $\mathcal{E}(\phi/\psi)$ is violated if ψ is satisfied, but ϕ is not. $\mathcal{F}(\phi/\psi)$ is violated when both ϕ and ψ are satisfied.)

The models \mathfrak{M}_P and \mathfrak{M}_K , satisfying Def. 2.2 and 2.3, are defined as: $W^P = W^K = \{w_i \mid 1 \leq i \leq 8\}$ s.t. $\|\text{harm}\| = V(\text{harm}) = \{w_2, w_3, w_4, w_6, w_7, w_8\}$, $\|\text{kill}\| = V(\text{kill}) = \{w_2, w_3, w_6, w_7\}$, $\|\mathfrak{S}y\| = V(\mathfrak{S}y) = \{w_4, w_8\}$, $\|\text{des_kill}\| = V(\text{des_kill}) = \{w_5, w_6, w_7, w_8\}$ (with $V^P = V^K = V$), $\mathcal{N}_{\mathcal{F}}^P(w_i) = \mathcal{N}_{\mathcal{F}}^K(w_i) = \{(X, Y) \mid X \subseteq \{w_2, w_3, w_4, w_6, w_7, w_8\}, Y = W\}$, $\mathcal{N}_{\mathcal{O}}^P(w_i) = \mathcal{N}_{\mathcal{O}}^K(w_i) = \{(V, Z) \mid \{w_2, w_6\} \subseteq V, Z = \{w_5, w_6, w_7, w_8\}\}$ and $\mathcal{N}_{\mathcal{E}}^K(w_i) = \emptyset$.

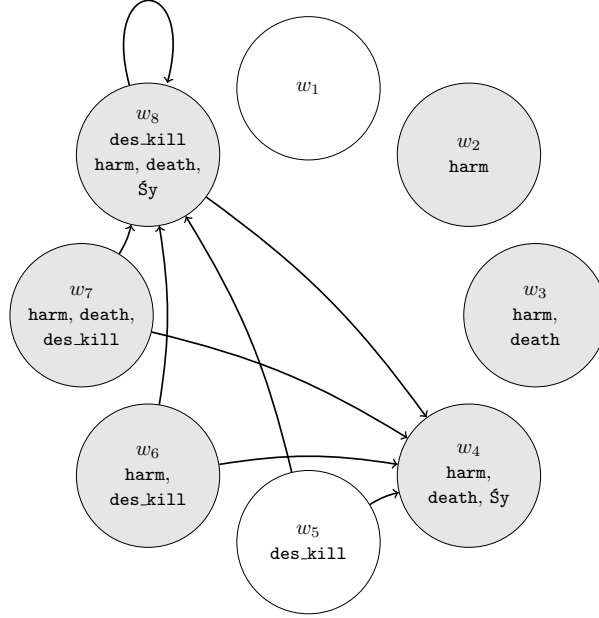


Fig. 1. depicts the models \mathfrak{M}_P and \mathfrak{M}_K satisfying the Śyena controversy (A_i)-(D) ($i \in \{P, K\}$): for $(\|\mathbf{harm}\|, \|\top\|) \in \mathcal{N}_F^i$ ($i \in \{P, K\}$). The worlds $w_i \in \|\mathbf{harm}\|$ are coloured grey and for $(\|\mathbf{Śy}\|, \|\mathbf{des_kill}\|) \in \mathcal{N}_E^P(w_i) = \mathcal{N}_O^K(w_i)$ (expressing A_P and A_K, resp.) the elements are indicated by arrows from each $w_i \in \|\mathbf{des_kill}\|$ to each $w_j \in \|\mathbf{Śy}\|$. For Prabhākara w_1 is the only deontically acceptable world, while Kumāriḷa also accepts w_5 , as (A_K) has no deontic force.

From Kumāriḷa’s perspective, all the worlds that are not coloured grey—i.e., worlds where the prohibition (B) is not violated—are deontically acceptable, namely, no command with deontic force is violated. Kumāriḷa’s answer relies on the distinction between obligations and statements prescribing elective sacrifices, which are mutually independent: i.e., in case they conflict with a prohibition, elective sacrifices can be omitted without risk, thus avoiding to violate the prohibition. In contrast, in Prabhākara’s logic the two neighbourhoods associated with (A_P) and (B) are not independent: i.e., condition (iv) of Def. 2.2 excludes the possibility that the same neighbourhood of a world represents both a prohibition and an obligation. However, since the eligibility conditions of the two commands do not exactly coincide, there is at least one world—i.e., w_1 —in which no command is violated. That is, w_1 is the only deontically acceptable world from Prabhākara’s point of view, the state in which one does not desire to kill one’s enemy. Since desires are interpreted by Prabhākara as irreversible decisions—i.e., for Prabhākara the desire to kill amounts to a decision to kill—his solution is a case of CTD reasoning: the injunction to perform the śyena represents an obligation taking effect when a violation (the decision to cause a death) has occurred.

Following [7], the model \mathfrak{M}_P also explains Prabhākara’s claim that the Vedas do not impel one to perform the malevolent sacrifice *śyena*, they only say that it is obligatory. This claim that a Vedic obligation does not necessarily impel was wrongly considered meaningless, for instance, in [25].

3 Maṇḍana

Maṇḍana is a key Sanskrit philosopher, also known for his revolutionary approach to deontic modals. He reduces commands to descriptions of states of affairs, that is, to *instrumentality* relations holding between actions and desired results. For instance, “you are obliged to perform the *kārīri* ritual, if you desire rain” is reduced to “the *kārīri* is an instrument for attaining rain”. Presently we interpret ‘instrument’ as an action sufficient to guarantee its result. Despite his revolutionary approach, Maṇḍana did not wish to break with the Mīmāṃsā tradition and its distinction among different types of duties. Still, his reduction may suggest that since all command-types are mere instrument statements there is also no difference in degrees of commands. To retain the distinction, Maṇḍana adopts two constraints involving *pāpa*, i.e., *bad karma*.

First, to individuate fixed duties, Maṇḍana argues for the universal desirability of their coveted result: the *reduction of bad karma*. For Maṇḍana, the reduction of bad karma is a desire shared by every agent. The introduction of this fixed desire, preserves the distinction between obligations (instruments that reduce bad karma) and elective duties (instruments serving specific desires). Second, to ensure the prohibitive strength of actions leading to undesirable outcomes, Maṇḍana argues that prohibited actions are instruments to outcomes whose undesirability is incommensurably greater than any desirable result. This universally undesired result is the *accumulation of bad karma*.

As will be shown at the end of this section, Maṇḍana’s solution to the *śyena* controversy centers on the rationality of the agents involved. No rational agent would desire the small benefit of performing *śyena* in exchange for its accessory negative result, the accumulation of bad karma.

Related work. As Maṇḍana reasons about actions and outcomes, a PDL-like language [8,23] seems adequate. Actually, a minimal action-language suffices: i.e., negation and combination. Hence, we base our logic on [2]: a basic PDL-like language reducing action-modalities to action constants. The formalism in [2] is aimed at representing Von Wright’s theory of instrumentality and hence appears particularly suitable. An alternative approach may be BDI logics [22], due to its connection to means-end reasoning (cf. [18]). However, they do not accommodate the required distinction between actions and outcomes. (Due to the role of desires in Maṇḍana’s account, BDI-like extensions of our logic will be reserved for future work.)

To reason about bad karma, we adopt and enhance an Andersonian reduction to deontic logic [1]: φ is obligatory iff $\neg\varphi$ necessarily implies a sanction. The reduction was adapted by Meyer [23] to the deontic action setting: action Δ is obligatory iff all performances of its complement $\bar{\Delta}$ lead to a violation. Similarly, Maṇḍana can be seen as a reductionist of deontic reasoning: every

Vedic command is an instrumentality statement about actions leading to states of affairs, sanctions, and rewards. In deontic logic, the use of positive constants was introduced by Kanger [19]: φ is obligatory iff in the good state φ holds. However Kanger’s approach takes φ as a necessary condition for the ‘good state’ whereas Maṇḍana takes φ as sufficient condition for ‘reducing bad karma’.

3.1 The logic LMa: Language, Axioms and Semantics

We introduce the normal modal logic LMa equipped with action constants and karma constants. LMa captures Maṇḍana’s intended reduction of norms to claims of instrumentality. We start by introducing an algebra of action \mathcal{L}_{Act} and the logical language \mathcal{L}_{LMa} into which these actions will be translated. Presently, a single-agent setting suffices. Let Act be a set of atomic actions δ (such as ‘opening the window’). The action language \mathcal{L}_{Act} is defined as

$$\Delta ::= \delta \mid \overline{\Delta} \mid \Delta \cup \Delta \quad \text{with } \delta \in \text{Act}$$

The operator $\overline{}$ denotes the complement of an action, whereas \cup is read as a disjunctive action. We use uppercase Greek letters Δ, Γ, \dots to denote arbitrary actions. We define $\Delta \cap \Gamma = \overline{\overline{\Delta} \cup \overline{\Gamma}}$ as the the joint performance of actions.

The language \mathcal{L}_{LMa} for Maṇḍana is defined through the following BNF:

$$\varphi ::= p \mid \mathbf{d}_\delta \mid \mathbf{P} \mid \mathbf{R} \mid \neg\varphi \mid \varphi \vee \varphi \mid \boxed{\varphi} \mid \sqcup\varphi$$

with $p \in \text{Atom}$ and $\mathbf{d}_\delta \in \text{Wit}_{\text{Act}}$, where Atom is the set of atomic propositions, and Wit_{Act} the set of atomic *constants* called ‘action-witnesses’. The constant \mathbf{d}_δ is to be read as a witness stating that ‘the action δ has just been successfully performed’.⁷ \mathbf{P} is a constant reading ‘bad karma is accumulated’ and the constant \mathbf{R} reads ‘bad karma is reduced’. The unary operators $\boxed{}$ and \sqcup are interpreted as ‘in all succeeding states it holds that’ and ‘it is universally necessary that’, respectively. Their respective duals \diamond and \heartsuit are defined as usual.

The translation between \mathcal{L}_{Act} and action formulae in our object language \mathcal{L}_{LMa} is established through the following recursive definition:

- For all $\delta \in \text{Act}$, $t(\delta) = \mathbf{d}_\delta$
- For all $\Delta \in \mathcal{L}_{\text{Act}}$, $t(\overline{\Delta}) = \neg t(\Delta)$
- For all $\Delta, \Gamma \in \mathcal{L}_{\text{Act}}$, $t(\Delta \cup \Gamma) = t(\Delta) \vee t(\Gamma)$

The upshot of the above translation is that it enables us to reason with actions on the object language level. The resulting versatility will prove useful in (i) defining a variety of modal operators (including instruments and commands) and (ii) axiomatizing action-properties. For instance, $\boxed{t(\Delta) \rightarrow \varphi}$ reads “at every successor state witnessing the successful performance of action Δ , the

⁷ The logic LMa does not allow to keep track of action histories, only the last executed actions are known (cf. the presence of action witnesses). This is due to the absence of modalities referring to the past, which are not required in our present analysis of instruments.

state-of-affairs φ holds”. When taken together with action, the modality $\boxed{\square}$ can be seen as an indeterministic *execution operator*, in the spirit of propositional dynamic logic (PDL): “every successful execution of Δ , guarantees φ ”. See [2] for a discussion of this basic PDL-reductionist approach, and for a formal analysis of different notions of instrumentality.

Axiomatization. As for the previously introduced logics, also for Maṇḍana we want to avoid imposing any property that cannot be traced back to the Mīmāṃsā in general, and Maṇḍana in particular. For this reason, the proposed logic will be rather minimal. The Hilbert-style axiomatization of LMa is presented in Def. 3.1 (below) and justified accordingly: Both $\boxed{\square}$ and $\boxed{\square}$ are normal modal operators due to the S5 characterization of the former and A_{LMa1} for the latter. A_{LMa2} expresses a bridge axiom, stating that what holds universally, must also hold at any successor state. A_{LMa3} conveys the Mīmāṃsā principle that whenever bad karma is attainable, it is also avoidable. (This principle is based on the Mīmāṃsā meta-rule according to which all commands need to be non-trivial and to prescribe something new, see [10].) A_{LMa4} captures the same property for the reduction of bad karma, and A_{LMa5} gives a central Mīmāṃsā principle: “if an action is executable, then it is executable in such a way that it does not trigger both the reduction and the increase of bad karma” ([29] ad 1.1.2), see Remark 3.9.

Definition 3.1 Maṇḍana logic LMa extends $\boxed{\square}$ -S5 with the following axioms:

$$\begin{array}{ll} A_{LMa1}. \boxed{\square}(\varphi \rightarrow \psi) & A_{LMa3}. \diamondsuit P \rightarrow \diamondsuit \neg P \\ \rightarrow (\boxed{\square}\varphi \rightarrow \boxed{\square}\psi) & A_{LMa4}. \diamondsuit R \rightarrow \diamondsuit \neg R \\ A_{LMa2}. \boxed{\square}\varphi \rightarrow \boxed{\square}\varphi & A_{LMa5}. \diamondsuit t(\Delta) \rightarrow \diamondsuit (t(\Delta) \wedge (\neg P \vee \neg R)) \end{array}$$

A derivation of $\varphi \in \mathcal{L}_{LMa}$ in LMa from a set $\Sigma \subseteq \mathcal{L}_{LMa}$ (i.e., $\Sigma \vdash_{LMa} \varphi$) is defined as usual [3]. When $\Sigma = \emptyset$, we say φ is an LMa-theorem, and write $\vdash_{LMa} \varphi$.

Semantics. We introduce a relational semantics for the logic LMa:

Definition 3.2 An LMa-frame $\mathfrak{F}_{LMa} = \langle W, \{W_\delta : \delta \in \text{Act}\}, W_P, W_R, R_{\boxed{\square}}, R_{\boxed{\square}} \rangle$ is a tuple with $W \neq \emptyset$ a set of worlds w, v, u, \dots etc. For every $d_\delta \in \text{Wit}_{\text{Act}}$, let $W_\delta \subseteq W$ be the set of worlds witnessing the successful performance of δ . Let $W_{\overline{\Delta}} = W \setminus W_\Delta$, and $W_{\Delta \cup \Gamma} = W_\Delta \cup W_\Gamma$. $W_P \subseteq W$ and $W_R \subseteq W$, are sets of worlds witnessing the accumulation and reduction of bad karma, resp. Last, $R_{\boxed{\square}} \subseteq W \times W$ and $R_{\boxed{\square}} = W \times W$ are binary relations s.t. the following holds:

- (i) $R_{\boxed{\square}} \subseteq R_{\boxed{\square}}$;
- (ii) $\forall w, v \in W ((w, v) \in R_{\boxed{\square}} \text{ and } v \in W_P) \text{ implies } \exists u ((w, u) \in R_{\boxed{\square}} \text{ and } u \notin W_P)$
- (iii) $\forall w, v \in W ((w, v) \in R_{\boxed{\square}} \text{ and } v \in W_R) \text{ implies } \exists u ((w, u) \in R_{\boxed{\square}} \text{ and } u \notin W_R)$
- (iv) $\forall w, v \in W ((w, v) \in R_{\boxed{\square}} \text{ and } v \in W_\Delta) \text{ implies } \exists u ((w, u) \in R_{\boxed{\square}} \text{ and } u \in W_\Delta \setminus W_R \cap W_P)$

An LMa-model is a tuple $\mathfrak{M}_{LMa} = \langle \mathfrak{F}_{LMa}, V \rangle$ where \mathfrak{F}_{LMa} is an LMa-frame and V is a valuation function mapping atomic propositional symbols from $\text{Atom} \cup \text{Wit}_{\text{Act}} \cup \{P\} \cup \{R\}$ to sets of worlds, such that the following conditions are

satisfied: $V(d_\delta) = W_\delta$ for every $d_\delta \in \text{Wit}_{\text{Act}}$, $V(\mathbf{P}) = W_{\mathbf{P}}$, and $V(\mathbf{R}) = W_{\mathbf{R}}$. (n.b. constants \mathbf{P} , \mathbf{R} and those from Wit_{Act} have a fixed evaluation over frames). We use $\mathcal{C}_{\text{LMa}}^f$ to refer to the entire class of LMa-frames.

The $\boxed{\square}$ -modality behaves as a universal modality, hence its corresponding accessibility relation $R_{\boxed{\square}}$ is reflexive, symmetric and transitive (cf. Sect. 2). The purpose of $\boxed{\square}$ is to represent universally true statements, which should hold ‘at every world’. The intended use of the $\boxed{\square}$ -modality is to represent the possible outcomes of transitions triggered by actions. We have adopted a very general notion of the ‘immediate successor’ relation, by imposing no additional properties on this relation (cf. the absence of irreflexivity and asymmetry). We point out that there is no Mīmāṃsā characterization of time available to justify such properties. However, we do realize that, in general, these properties may be desirable in a temporal logic of action. Following [2], one can show that LMa can likewise be characterized by a subclass of LMa-frames including only asymmetric and intransitive *tree-like* frames (this is due to the fact that languages such as \mathcal{L}_{LMa} cannot force these additional frame properties; cf. [3]). For the purpose of this paper, a general notion of the immediate successor relation suffices.

Semantic evaluation of formulae φ from \mathcal{L}_{LMa} is defined accordingly:

Definition 3.3 Let $\mathfrak{M}_{\text{LMa}}$ be an LMa-model and $w \in W$ of $\mathfrak{M}_{\text{LMa}}$. We define the *satisfaction* of a formula $\varphi \in \mathcal{L}_{\text{LMa}}$ in $\mathfrak{M}_{\text{LMa}}$ at w inductively:

$\mathfrak{M}_{\text{LMa}}, w \models \chi$	iff	$w \in V(\chi)$, for any $\chi \in \text{Atom} \cup \text{Wit}_{\text{Act}} \cup \{\mathbf{P}\} \cup \{\mathbf{R}\}$
$\mathfrak{M}_{\text{LMa}}, w \models \neg\varphi$	iff	$\mathfrak{M}_{\text{LMa}}, w \not\models \varphi$
$\mathfrak{M}_{\text{LMa}}, w \models \varphi \vee \psi$	iff	$\mathfrak{M}_{\text{LMa}}, w \models \varphi$ or $\mathfrak{M}_{\text{LMa}}, w \models \psi$
$\mathfrak{M}_{\text{LMa}}, w \models \boxed{\square}\varphi$	iff	for all $v \in W$ s.t. $(w, v) \in R_{\boxed{\square}}$, $\mathfrak{M}_{\text{LMa}}, v \models \varphi$
$\mathfrak{M}_{\text{LMa}}, w \models \boxed{\square}\varphi$	iff	for all $v \in W$ s.t. $(w, v) \in R_{\boxed{\square}}$, $\mathfrak{M}_{\text{LMa}}, v \models \varphi$

The semantic clauses for the dual operators \diamond and \heartsuit as well as global truth, validity and semantic entailment are defined as usual (see [3]).

Theorem 3.4 (*Soundness*) For all $\varphi \in \mathcal{L}_{\text{LMa}}$ and $\Gamma \subseteq \mathcal{L}_{\text{LMa}}$, if $\Gamma \vdash_{\text{LMa}} \varphi$, then $\mathcal{C}_{\text{LMa}}^f, \Gamma \models_{\text{LMa}} \varphi$

Proof. Soundness is proven as usual. Explicating the use of constants we prove axiom $\text{ALMa}5$. Let $\mathfrak{M}_{\text{LMa}}$ be an LMa-model with $w \in W$. Suppose $\mathfrak{M}_{\text{LMa}}, w \models \diamond t(\Delta)$. Then $\exists v \in W$ s.t. $(w, v) \in R_{\diamond}$ and $\mathfrak{M}_{\text{LMa}}, v \models t(\Delta)$. So $v \in W_\Delta$. By (iv) of Def. 3.2, $\exists u \in W$ s.t. $(w, u) \in R_{\heartsuit}$ and $u \in W_\Delta \setminus W_{\mathbf{R}} \cap W_{\mathbf{P}}$. So $\mathfrak{M}_{\text{LMa}}, u \models t(\Delta)$ and $\mathfrak{M}_{\text{LMa}}, u \not\models \mathbf{R} \wedge \mathbf{P}$. Which gives $\mathfrak{M}_{\text{LMa}}, w \models \heartsuit(t(\Delta) \wedge (\neg\mathbf{R} \vee \neg\mathbf{P}))$. \square

Strong completeness is proven via canonical model construction, adjusted to the inclusion of constants. LMa-maximal consistent sets (MCS) are defined as usual, enjoying the usual properties. We define the following canonical model:

Definition 3.5 Let $M^c = \langle W^c, \{W_{d_\delta}^c \mid d_\delta \in \mathcal{L}_{\text{LMa}}\}, W_{\mathbf{P}}^c, W_{\mathbf{R}}^c, R_{\boxed{\square}}^c, R_{\boxed{\square}}^c, V^c \rangle$ be a canonical model, where W^c is the set of all LMa-MCSs ($\Gamma, \Sigma, \Phi, \dots$) and:

- For all $d_\delta \in \mathcal{L}_{\text{LMa}}$ and $\Sigma \in W^c$, $\Sigma \in W_{d_\delta}^c$ iff $d_\delta \in \Sigma$

- For $\alpha \in \{\mathbf{P}, \mathbf{R}\}$, and all $\Sigma \in W^c$, $\Sigma \in W_\alpha^c$ iff $\alpha \in \Sigma$
- For $\alpha \in \{\mathbf{S}, \mathbf{R}\}$, and all $\Sigma, \Gamma \in W^c$, $(\Sigma, \Gamma) \in R_\alpha^c$ iff $\{\phi \mid [\alpha]\phi \in \Sigma\} \subseteq \Gamma$
- For all $\chi \in \text{Atom} \cup \text{ActWit} \cup \{\mathbf{P}\} \cup \{\mathbf{R}\}$, $V^c(\chi) = \{\Sigma \mid \chi \in \Sigma \in W^c\}$

The existence lemma and truth lemma are proven in [3, Sect. 4.2] (nb. LMa is a normal modal logic). We show that M^c belongs to the class of LMa-models, i.e., satisfying the properties of Def. 3.2.

Theorem 3.6 M^c is an LMa-model.

Proof. We demonstrate the LMa specific properties (ii) and (iv) (Def. 3.2). The proofs of the remaining properties are similar.

(ii) For all $\Sigma, \Gamma \in W^c$, $(\Sigma, \Gamma) \in R_{\mathbf{S}}^c$ with $\Gamma \in W_{\mathbf{P}}^c$, there exists a $\Theta \in W^c$ s.t. $(\Sigma, \Theta) \in R_{\mathbf{S}}^c$ and $\Theta \notin W_{\mathbf{P}}^c$. Assume the antecedent, we construct the set Θ . Let $\Theta^- = \{\neg\mathbf{P}\} \cup \{\phi \mid \mathbf{S}\phi \in \Sigma\}$. Suppose Θ^- is not LMa-consistent. Hence for some $\phi_1, \dots, \phi_n \in \Theta^-$, we have $\vdash_{\text{LMa}} \phi \wedge \dots \wedge \phi_n \rightarrow \mathbf{P}$. By LMa we have $\vdash_{\text{LMa}} \mathbf{S}(\phi \wedge \dots \wedge \phi_n \rightarrow \mathbf{P})$, which implies $\vdash_{\text{LMa}} \mathbf{S}(\phi \wedge \dots \wedge \phi_n) \rightarrow \mathbf{S}\mathbf{P}$, and so $\vdash_{\text{LMa}} \mathbf{S}\phi \wedge \dots \wedge \mathbf{S}\phi_n \rightarrow \neg\mathbf{S}\neg\mathbf{P}$. By monotonicity of LMa, $\vdash_{\text{LMa}} \mathbf{S}\phi \wedge \dots \wedge \mathbf{S}\phi_n \wedge \mathbf{S}\mathbf{P} \rightarrow \neg\mathbf{S}\neg\mathbf{P}$. By assumption $\mathbf{S}\phi_1, \dots, \mathbf{S}\phi_n, \mathbf{S}\mathbf{P} \in \Sigma$ and MCS properties, we have $\neg\mathbf{S}\neg\mathbf{P} \in \Sigma$. However, since Σ is a LMa-MCS we have $\mathbf{S}\mathbf{P} \rightarrow \mathbf{S}\neg\mathbf{P} \in \Sigma$, and thus $\mathbf{S}\neg\mathbf{P} \in \Sigma$. Contradiction. Hence, Θ^- is LMa-consistent. Let Θ be the LMa-MCS extending Θ^- (Lindenbaum's lemma). By construction of M^c we obtain $(\Sigma, \Theta) \in R_{\mathbf{S}}^c$ and since $\neg\mathbf{P} \in \Theta^- \subseteq \Theta$ we have $\Theta \notin W_{\mathbf{P}}^c$.

(iv) For all $\Sigma, \Gamma \in W^c$, if $(\Sigma, \Gamma) \in R_{\mathbf{S}}^c$ with $\Gamma \in W_{t(\Delta)}^c$, then there exists a $\Theta \in W^c$ s.t. $(\Sigma, \Theta) \in R_{\mathbf{S}}^c$ and $\Theta \in W_{t(\Delta)}^c \setminus W_{\mathbf{R}}^c \cap W_{\mathbf{P}}^c$. Assume the antecedents, we construct such a Θ . Let $\Theta^- = \{t(\Delta)\} \cup \{\neg\mathbf{R} \vee \neg\mathbf{P}\} \cup \{\phi \mid \mathbf{S}\phi \in \Sigma\}$. Suppose Θ^- is LMa-inconsistent. Then there are $\phi_1, \dots, \phi_n \in \Theta^-$ s.t. $\vdash_{\text{LMa}} \phi_1 \wedge \dots \wedge \phi_n \rightarrow \neg(t(\Delta) \wedge (\neg\mathbf{R} \vee \mathbf{P}))$. Hence, we have $\vdash_{\text{LMa}} \mathbf{S}\phi_1 \wedge \dots \wedge \mathbf{S}\phi_n \rightarrow \neg\mathbf{S}(t(\Delta) \wedge (\neg\mathbf{R} \vee \mathbf{P}))$. By monotonicity, $\vdash_{\text{LMa}} \mathbf{S}\phi_1 \wedge \dots \wedge \mathbf{S}\phi_n \wedge \mathbf{S}(t(\Delta)) \rightarrow \neg\mathbf{S}(t(\Delta) \wedge (\neg\mathbf{R} \vee \mathbf{P}))$. Since $\mathbf{S}\phi_1, \dots, \mathbf{S}\phi_n, \mathbf{S}(t(\Delta)) \in \Sigma$, we get $\neg\mathbf{S}(t(\Delta) \wedge (\neg\mathbf{R} \vee \mathbf{P})) \in \Sigma$. By inclusion of axiom $\mathbf{S}t(\Delta) \rightarrow \mathbf{S}(t(\Delta) \wedge (\neg\mathbf{R} \vee \mathbf{P})) \in \Sigma$, we get a contradiction. Hence, Θ^- is consistent. Let Θ be the LMa-MCS extending Θ^- . By construction of Θ we get $(\Sigma, \Theta) \in R_{\mathbf{S}}^c$. Since $t(\Delta) \in \Theta$ we have $\Theta \in W_{t(\Delta)}^c$. Last, since $\neg\mathbf{R} \vee \mathbf{P} \in \Theta$ we get $\Theta \notin W_{\mathbf{R}}^c \cap W_{\mathbf{P}}^c$, hence $\Theta \in W_{t(\Delta)}^c \setminus W_{\mathbf{R}}^c \cap W_{\mathbf{P}}^c$. \square

Corollary 3.7 (Strong Completeness for LMa) For all $\phi \in \mathcal{L}_{\text{LMa}}$ and $\Gamma \subseteq \mathcal{L}_{\text{LMa}}$, we have: if $\mathcal{C}_{\text{LMa}}^f, \Gamma \models \phi$, then $\Gamma \vdash_{\text{LMa}} \phi$.

3.2 Instrumentality and Mīmāṃsā properties

We introduce Maṇḍana's notion of instruments, his deontic reduction, and discuss important Mīmāṃsā properties and their rendering in Maṇḍana's logic.

Instruments and Maṇḍana's deontic reduction. Maṇḍana's program consists in reducing all deontic modalities to a uniform notion of instrumentality. Our uniform definition of instrumentality must satisfy the following Maṇḍana-criteria: First, (i) the instrument relation contains three components: (a) an action Δ , serving as the instrument; (b) a state-of-affairs φ , represent-

ing the outcome of Δ ; and (c) a state-of-affairs χ defining the circumstances in which Δ functions as an instrument for bringing about φ . Second, (ii) the circumstances χ must be meaningful, that is, χ must be possible in the broadest sense. Last, the agent must have a *choice* to perform the action Δ when circumstances χ occur; (iii) Δ can be performed and (iv) Δ can be refrained from (for (ii–iv) see Śābara on *Mīmāṃsāsūtra* 6.1 in [27]). In short, we take $\mathcal{I}(\Delta/\varphi/\chi)$ to read “ Δ is an instrument for guaranteeing φ in circumstances χ ”, which amounts to:

“(i) Whenever circumstances χ hold, performing Δ guarantees φ , (ii) χ is a possible circumstance, (iii) at χ , Δ is possible, and (iv) at χ , $\bar{\Delta}$ is possible.”

The formal definition of instrumentality, based on (i)–(iv), is given in Def. 3.8.

Maṇḍana’s reduction, that is, the reduction of commands to statements of instrumentality, is then obtained accordingly: prohibited and obligatory actions are defined in terms of those actions being instrumental to the outcome of bad karma (i.e., P) and the reduction of bad karma (i.e., R), respectively. Electives are those actions instrumental to outcomes that are neither P nor R.

Definition 3.8 Maṇḍana’s notion of instruments in the logic LMa:

$$\begin{array}{lll} \mathcal{I}(\Delta/\varphi/\chi) & := & \begin{array}{ll} \text{(i)} & \boxed{\chi \rightarrow \boxed{t(\Delta) \rightarrow \varphi}} \quad \wedge \\ \text{(ii)} & \diamond \chi \quad \wedge \\ \text{(iii)} & \boxed{\chi \rightarrow \diamond t(\Delta)} \quad \wedge \\ \text{(iv)} & \boxed{\chi \rightarrow \diamond \neg t(\Delta)} \end{array} \end{array}$$

Maṇḍana’s reduction of obligations, prohibitions and elective sacrifices in LMa:

$$\begin{array}{lll} \mathcal{O}(\Delta/\chi) & := & \mathcal{I}(\Delta/R/\chi) \\ \mathcal{F}(\Delta/\chi) & := & \mathcal{I}(\Delta/P/\chi) \\ \mathcal{E}(\Delta/\varphi/\chi) & := & \mathcal{I}(\Delta/\varphi/\chi) \quad \text{with } \varphi \not\vdash_{\text{LMa}} P \text{ and } \varphi \not\vdash_{\text{LMa}} R \end{array}$$

(n.b. we differentiate actions (kill, sacrifice), from results (death, P, R), with a contrasting font style.)

Recall that for the Mīmāṃsā school, obligations, prohibitions and electives cannot be expressed in terms of one another [12,20]. Similarly, Maṇḍana adopts this irreducibility by limiting the result of the instruments corresponding to the three norm types. This property is preserved in Def. 3.8. In particular, we note that the result of an elective sacrifice cannot entail either of the results identified with obligations or prohibitions. We come back to this in Sect. 3.3, when we discuss Maṇḍana’s solution to the *śyena* controversy. Furthermore, in LMa, the elective operator \mathcal{E} has one additional argument. This is because the instrument notation is more expressive, and there are variables for both the eligibility condition (the desire) and the purpose (the object of the desire).

Remark 3.9 Now that we have defined instruments, let us go back to axiom $A_{\text{LMa}5}$ in Def. 3.1. Observe that it limits the interaction between actions and karma constants. In essence, $A_{\text{LMa}5}$ ensures that an action Δ cannot be both obligatory and prohibited, i.e., Δ cannot at the same time be an *instrument*

for the reduction and accumulation of bad karma (n.b. the inconsistent action $\Delta \cap \bar{\Delta}$ is excluded from being an instrument by Def. 3.8). Nevertheless, we still allow for singular situations where we end up with both **P** and **R** after executing Δ , however $A_{\text{LMa}5}$ guarantees that this must be the result of some other action Γ executed alongside Δ (one being a prohibition, the other an obligation).

We show below that the logic **LMa** is expressive enough to entail other principles that can be found in *Mīmāṃsā* (as **LMa**-theorems).

Contingency. For the *Mīmāṃsā* it is essential that actions in commands are meaningful (see Śābara on *Mīmāṃsāsūtra* 6.1, [27]). For an action to be meaningful, an agent must have the choice to perform it as well as refrain from performing it:

$$\mathcal{I}(\Delta/\varphi/\chi) \rightarrow \boxed{\chi} \rightarrow (\diamond t(\Delta) \wedge \diamond \neg t(\Delta)) \text{ for } \varphi \in \{\mathbf{P}, \mathbf{R}\} \text{ or } (\varphi \not\vdash \mathbf{P} \text{ and } \varphi \not\vdash \mathbf{R})$$

In deontic logic this property is known as the *contingency principle* [30, p. 11][1]. The above formula is an **LMa**-theorem, guaranteed solely by our definition of instruments ((iii) and (iv) in Def. 3.8). However, for obligations and prohibitions the property is also implied in association with axioms $A_{\text{LMa}3}$ and $A_{\text{LMa}4}$. That is,

$$\mathcal{I}(\Delta/\mathbf{R}/\chi) \equiv \mathcal{O}(\Delta/\chi) \equiv (\diamond \chi \wedge \boxed{\chi} \rightarrow \boxed{\mathbf{S}}(t(\Delta) \rightarrow \mathbf{R})) \wedge \boxed{\chi} \rightarrow \diamond t(\Delta)$$

and,

$$\mathcal{I}(\Delta/\mathbf{P}/\chi) \equiv \mathcal{F}(\Delta/\chi) \equiv (\diamond \chi \wedge \boxed{\chi} \rightarrow \boxed{\mathbf{S}}(t(\Delta) \rightarrow \mathbf{P})) \wedge \boxed{\chi} \rightarrow \diamond t(\Delta)$$

are **LMa**-theorems. These theorems demonstrate that condition (iv) of instruments (Def. 3.8) is *admissible* in the light of Maṇḍana's analysis. However, (iv) is still necessary to ensure meaningfulness of actions for elective duties; i.e., $\not\vdash_{\text{LMa}} \mathcal{E}(\Delta/\varphi/\chi) \equiv (\diamond \chi \wedge \boxed{\chi} \rightarrow \boxed{\mathbf{S}}(t(\Delta) \rightarrow \varphi)) \wedge \boxed{\chi} \rightarrow \diamond t(\Delta)$.

No impossible commands. Although the logic **LMa** does not adopt a D-axiom for deontic consistency, the following formula is in fact an **LMa**-theorem:

$$\vdash_{\text{LMa}} \neg(\mathcal{F}(\Delta/\chi) \wedge \mathcal{F}(\bar{\Delta}/\chi))$$

The theorem corresponds to the *Mīmāṃsā* principle: “It is impossible that the Vedas tell you that you’ll fall (i.e., be reborn in hell) both if you do something and if you don’t do it” ([29, p. 32]). The quote refers to the impossibility of the Vedas giving contradictory instruments. The theorem is a direct consequence of the definition of instrumentality together with axiom $A_{\text{LMa}3}$. In fact, we obtain a similar theorem for obligations from axiom $A_{\text{LMa}4}$. Clearly, the scheme does not hold for elective duties (cf. Def. 3.8). Last, the logic **LMa** satisfies the *Mīmāṃsā* principle that obligations and prohibitions are strictly incompatible (even on the derived level). That is, the following formula is **LMa**-valid:

$$\neg(\mathcal{O}(\Delta/\chi) \wedge \mathcal{F}(\Delta/\chi))$$

Mīmāṃsā principles. The logics for Prabhākara and Kumārila are built upon principles (P1)-(P3) recalled in Sect. 2. A natural question to ask is whether these principles are preserved in Maṇḍana's reduction logic. Their reformulation in LMa is as follows (notice that (P1)-(P3) were postulated for commands in particular, not instruments in general):

$$(p1) \ (\mathcal{I}(\Delta/\varphi/\chi) \wedge \Box(t(\Delta) \rightarrow t(\Gamma))) \rightarrow \mathcal{I}(\Gamma/\varphi/\chi) \text{ such that } (\star)$$

$$(p2) \ (\mathcal{I}(\Delta/\varphi/\chi) \wedge \Box(\varphi \rightarrow \neg\psi)) \rightarrow \neg\mathcal{I}(\Delta/\psi/\chi) \text{ such that } (\star)$$

$$(p3) \ (\mathcal{I}(\Delta/\varphi/\chi) \wedge \Box(\chi' \equiv \chi)) \rightarrow \mathcal{I}(\Delta/\varphi/\chi') \text{ such that } (\star)$$

$$(\star) \ \varphi \in \{\mathbf{P}, \mathbf{R}\} \text{ or } (\varphi \not\vdash \mathbf{P} \text{ and } \varphi \not\vdash \mathbf{R})$$

(p1)-(p3) deal with instruments that are obligations, prohibitions and electives.

Principle (p1) is not an LMa-valid formula, (a counter-model is easily obtained), and for good reasons: instrumentality is a notion of sufficient cause. Maṇḍana knew this, but he had to somehow preserve the property expressed in (P1). He achieved this by explaining necessity as external to instruments: that is, Maṇḍana's account of *universally desirable outcomes* (i.e., R and P) ensures that no agent would, from a rational point of view, transgress such commands. Hence, although necessary conditions of instruments leading to reducing bad karma are themselves not recognized as instruments, from a meta point of view, no rational agent would refrain from performing them.

Principle (p2) is LMa-valid and it follows from Maṇḍana property (cf. Def. 3.8) that actions must be meaningful (thus leading to meaningful outcomes).

Last, (p3) is LMa-valid and follows from the fact that universal necessity is a normal modal operator.

3.3 Maṇḍana's solution

We utilize Maṇḍana's reduction and demonstrate that, when formalized in terms of *instrumentality*, the sentences (A)-(D) from Sect. 1 are satisfiable. That is, we show the consistency of Maṇḍana's solution to the *śyena* controversy by providing an LMa-model satisfying the following:

$$(A_M) \ \mathcal{E}(\acute{S}y/\mathbf{death}/\mathbf{des_kill}) \equiv \mathcal{I}(\acute{S}y/\mathbf{death}/\mathbf{des_kill}) \\ \equiv \Box(\mathbf{des_kill} \rightarrow \Box(t(\acute{S}y) \rightarrow \mathbf{death})) \wedge \Diamond \mathbf{des_kill} \wedge \\ \Box(\mathbf{des_kill} \rightarrow \Diamond t(\acute{S}y)) \wedge \Box(\mathbf{des_kill} \rightarrow \Diamond \neg t(\acute{S}y))$$

$$(B_M) \ \mathcal{F}(\mathbf{harm}/\mathbf{T}) \equiv \mathcal{I}(\mathbf{harm}/\mathbf{P}/\mathbf{T}) \equiv \Box \Diamond t(\mathbf{harm}) \wedge \Box \Box(t(\mathbf{harm}) \rightarrow \mathbf{P})$$

$$(C_M) \ \Box(t(\acute{S}y) \rightarrow \mathbf{death})$$

$$(D_M) \ \Box(\mathbf{death} \rightarrow t(\mathbf{harm}))$$

A model satisfying (A_M)-(D_M) is defined as follows: $\mathfrak{M}_{\text{LMa}} = \langle \mathfrak{F}_{\text{LMa}}, V \rangle$ with $\mathfrak{F}_{\text{LMa}} = \langle W, W_{\acute{S}y}, W_{\mathbf{harm}}, W_{\mathbf{P}}, W_{\mathbf{R}}, R_{\Box}, R_{\Box} \rangle$ s.t.: $W = \{w_1, w_2, w_3\}$, $W_{\acute{S}y} = W_{\mathbf{harm}} = W_{\mathbf{P}} = \{w_2\}$, $W_{\mathbf{R}} = \emptyset$, $V(\mathbf{des_kill}) = \{w_0\}$, $V(\mathbf{death}) = \{w_2\}$, $R_{\Box} = W \times W$, $R_{\Box} = \{(w_1, w_2), (w_1, w_3), (w_2, w_2), (w_2, w_3), (w_3, w_2), (w_3, w_3)\}$. Note that $\mathfrak{M}_{\text{LMa}}$ satisfies the properties in Def. 3.2. The model

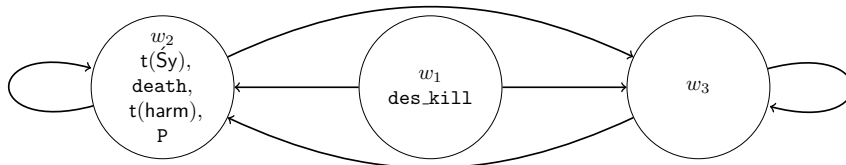


Fig. 2. *śyena* model in LMa, with the arrows representing the relation $R_{\mathbb{S}}$.

is represented in Fig. 2.

For all $w \in \{w_0, w_1, w_2\}$, we have $\mathfrak{M}, w \models \mathcal{I}(\text{harm}/\text{P}/\text{T}) \wedge \mathcal{I}(\dot{\text{Śy}}/\text{death}/\text{des_kill}) \wedge \Box(t(\dot{\text{Śy}}) \rightarrow \text{death}) \wedge \Box(\text{death} \rightarrow t(\text{harm}))$. The model shows that the *śyena* example is consistent. Furthermore, it illustrates that, given our assumptions, one cannot perform the *śyena* without accumulating bad karma P. It can easily be verified that $\mathcal{I}(\dot{\text{Śy}}/\text{P}/\text{des_kill})$ is the case (conditions (ii)-(iv) of the definition of instrumentality (Def. 3.8) follow from A_M , and condition (i) follows from A_M , B_M and D_M). So, the *śyena* is an instrument for bad karma. In fact, in the logic LMa, that is on a *derived level* (see footnote 5), the *śyena* sacrifice is prohibited $\mathcal{F}(\dot{\text{Śy}}/\text{des_kill})$. By contrast, on the Vedic level *śyena* is not prohibited. Following Maṇḍana, in LMa something can be prohibited and elective at the same time, without it being inconsistent. Maṇḍana's reasoning for the *śyena* sacrifice is the following: from a state where one desires to kill their enemy, it is rationally preferable not to perform the *śyena*. Performing it would transgress the prohibition of harming a living being, with the result of accumulating bad karma. This is necessarily undesirable for anyone, as discussed in Maṇḍana:

When it comes to pain and its cause, the one who is afflicted by them will always desire their removal. And the one who desires well-being desires to destroy the obstacle (bad karma) towards it. Therefore, the destruction of bad karma, a destruction which is the cause of what is desired, is always desired. (*Vidhiviveka* ad 2.8 [26])

4 Discussion of the three *śyena* solutions

We presented formal models that capture Prabhākara, Kumārila and Maṇḍana's responses to the *śyena* case. Here we compare the different solutions relating them to the history of deontic logic. Recall the main challenge facing the three authors: how to deal with seemingly conflicting prescriptions coming from a source that is assumed to be consistent. Prabhākara's solution is akin to CTD reasoning in deontic logic, which introduces (sub-ideality) levels to a normative system, not treating every norm on equal footing. We distinguish norms that hold primarily (possibly conditioned on circumstances) from norms that only arise in case of a norm violation, the latter being CTD obligations. In this case, the prohibition to commit violence is a primary norm, whereas the prescription of the *śyena* is an obligation that only comes into

force once a violation has occurred: for Prabhākara, the intention to kill one’s enemy amounts to violence. Here, we see a striking similarity with the most common interpretation of the GMP, namely Chisholm’s paradox [6].

Although Prabhākara and Kumārila agree that the *śyena* case does not constitute a dilemma, they argue so on different grounds. For Kumārila, prohibitions do not interact with electives in a mutually conflicting way. In particular, as an elective sacrifice the *śyena* has no deontic force and is thus overturned by the Vedic prohibition to commit violence. Despite some shallow similarities with the deontic logic literature on priority orderings (e.g., [15])—i.e., obligations and prohibitions being of highest priority for Kumārila—and hierarchies of different norm systems (e.g., [4])—i.e., obligations and prohibitions forming a proper norm system in contrast to electives—we note that Kumārila’s approach is different, in the sense that he assigns no deontic force to elective sacrifices whatsoever. They are mere sacrificial ways to attain one’s end, without being compelling, eliminating the controversy altogether.

Maṇḍana preserves Kumārila’s distinction between obligatory and elective sacrifices but offers a different solution: deontic modalities are just variations of a shared underlying structure, namely, instrument relations. In order to preserve the appealing distinction between the three norm types, Maṇḍana relates obligations and prohibitions to the reduction and accumulation of bad karma. Elective sacrifices are karma-independent. They might have indirect consequences on the reduction/accumulation of bad karma (e.g., the *śyena*), but their direct results are not karma-results. Maṇḍana argues that avoiding the accumulation of bad karma is *a priori* desired by all human beings, similarly its reduction. By reducing the Vedic norm system to notions of instruments and desires Maṇḍana does not yet resolve the problem, but transforms the seeming problem of conflicting norms to a problem of conflicting desires instead. What remains is a conflict between one’s desire to kill an enemy and one’s “rational” desire to avoid accumulating bad karma. In other words, Maṇḍana does not address the problem on the command level, but on the instrument level and, subsequently, solves it on the desire level. Interestingly, Maṇḍana is the only author that endorses the view that there is a real dilemma or conflict at stake in the *śyena* case. Nevertheless, he resolves the dilemma by arguing that avoiding the accumulation of bad karma is the highest of desires, which implies that no rational agent would ever perform the *śyena*. We find a priority ordering on the level of desires, consequently resolving the implied commands. There are striking similarities between Maṇḍana’s approach and the Kanger-Anderson approach to deontic logic [1,19], by opting for a unifying approach reducing a variety of modalities to a single (alethic) modality together with the notions of sanction (accruing bad karma) and goodness (reducing bad karma). However, Maṇḍana’s final solution to the *śyena* controversy is a decision-making problem that occurs on the meta-level, by making an appeal to rationality and undesirability.

Future work. Our interdisciplinary work only scratches the surface of the research opportunities offered by formal approaches to Mimāṃsā reasoning. As

illustrated in this work, these approaches can provide a better understanding of Mīmāṃsā texts, and may offer new stimuli for the deontic logic community.

Since the logics of the first two authors, Prabhākara and Kumārila, have been extensively studied elsewhere [7,20], further investigation of the logic of Maṇḍana and his reduction is planned. For instance, to simplify matters, in this work we took desires as regular terms of our object language. We plan to investigate the logical behaviour of desires as an intentional modality interacting with instruments and norms.

Acknowledgements. Work funded by the projects WWTF MA16-028 and FWF W1255-N23.

References

- [1] Anderson, A. R. and O. K. Moore, *The formal analysis of normative concepts*, American Sociological Review **22** (1957), pp. 9–17.
- [2] van Berkel, K. and M. Pascucci, *Notions of instrumentality in agency logic*, in: *International Conference on Principles and Practice of Multi-Agent Systems*, Springer, 2018, pp. 403–419.
- [3] Blackburn, P., M. de Rijke and Y. Venema, *Modal logic*, Cambridge tracts in theoretical computer science (2004).
- [4] Boella, G. and L. van der Torre, *Institutions with a hierarchy of authorities in distributed dynamic environments*, Artificial Intelligence and Law **16** (2008), pp. 53–71.
- [5] Chellas, B. F., “Modal Logic,” Cambridge University Press, Cambridge, 1980.
- [6] Chisholm, R. M., *Contrary-to-duty imperatives and deontic logic*, Analysis **24** (1963), pp. 33–36.
- [7] Ciabattoni, A., E. Freschi, F. A. Genco and B. Lellmann, *Mīmāṃsā deontic logic: Proof theory and applications*, in: H. De Nivelle, editor, *Automated Reasoning with Analytic Tableaux and Related Methods* (2015), pp. 323–338.
- [8] Fischer, M. J. and R. E. Ladner, *Propositional dynamic logic of regular programs*, Journal of computer and system sciences **18** (1979), pp. 194–211.
- [9] Forrester, J., *Gentle murder, or the adverbial samaritan*, J. of Philosophy (1984), pp. 193–197.
- [10] Freschi, E., *The role of paribhāṣās in Mīmāṃsā: rational rules of textual exegesis*, Asiatische Studien/Études Asiatiques **72** (2018), pp. 567–595.
- [11] Freschi, E., A. Ciabattoni, F. A. Genco and B. Lellmann, *Understanding prescriptive texts: Rules and logic as elaborated by the Mīmāṃsā school*, Journal of World Philosophies **2** (2017).
- [12] Freschi, E. and M. Pascucci, *Deontic concepts and their clash in Mīmāṃsā: towards an interpretation*, Theoria (2021).
- [13] Goble, L., *Murder most gentle: The paradox deepens*, Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition **64** (1991), pp. 217–227.
- [14] Guhe, E., *Killing gently by means of the śyena: The navya-nyāya analysis of vedic and secular injunctions (vidhi) and prohibitions (niṣedha) from the perspective of dynamic deontic logic*, Journal of Indian Philosophy (2021).
- [15] Hansen, J., *Prioritized conditional imperatives: problems and a new proposal*, Autonomous Agents and Multi-Agent Systems **17** (2008), pp. 11–35.
- [16] Hansson, B., *An analysis of some deontic logics*, in: *Deontic Logic: Introductory and Systematic Readings* (1971), pp. 121–147.
- [17] Hilpinen, R. and P. McNamara, *Deontic logic: A historical survey and introduction*, Handbook of deontic logic and normative systems **1** (2013), pp. 3–136.
- [18] Hughes, J., P. Kroes and S. Zwart, *A semantics for means-end relations*, Synthese **158** (2007), pp. 207–231.

- [19] Kanger, S., *New foundations for ethical theory*, in: *Deontic logic: Introductory and systematic readings*, Springer, 1971 pp. 36–58.
- [20] Lellmann, B., F. Gulisano and A. Ciabattani, *Mīmāṃsā deontic reasoning using specificity: a proof theoretic approach*, *Artificial Intelligence and Law* (2021).
- [21] Lewis, D., *Semantic analysis for dyadic deontic logics*, in: *Logical Theory and Semantical Analysis* (1974), pp. 1–14.
- [22] Meyer, J., J. Broersen and A. Herzig, *Bdi logics*, *Handbook of Logics of Knowledge and Belief* (2015), p. 453.
- [23] Meyer, J.-J. C. et al., *A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic.*, *Notre Dame J. Formal Log.* **29** (1988), pp. 109–136.
- [24] Prior, A. N., *Escapism: The logical basis of ethics*, in: A. Melden, editor, *Essays in Moral Philosophy* (1958).
- [25] Stcherbatsky, T. I., *Über die Nyāyakaṇikā des Vācaspatimiśra und die indische Lehre des kategorischen Imperativ*, in: W. Kierfel, editor, *Beiträge zur Literaturwissenschaft und Geistesgeschichte Indiens. Festgabe Hermann Jacobi zum 75. Geburtstag*, Kommissionsverlag F. Klopp, Bonn, 1926 pp. 369–380.
- [26] Stern, E., editor, “Vidhiviveka of Maṇḍana Miśra with the commentary Nyāyakaṇikā of Vācaspati Miśra,” forthcoming.
- [27] Subbāśāstrī, editor, “Śrīmajjaiminipraṇitaṃ Mīmāṃsādarśanam,” Ānandāśramamudrāṇālaya, Poona, 1929-1934.
- [28] van der Torre, L. and Y. hua Tan, *Deliberate robbery, or the calculating samaritan*, in: *In Proceedings of the ECAI’98 Workshop on Practical Reasoning and Rationality (PRR’98)*, 1998.
- [29] Viraraghavacharya, U. T., editor, “Seśvaramīmāṃsā-Mīmāṃsāpaduke, Seswara Mīmāṃsā and Mīmāṃsā paduka [by Venkaṭanātha],” *Ubhaya Vedanta Granthamala*, Madras, 1971.
- [30] von Wright, G. H., *Deontic logic*, *Mind* **60** (1951), pp. 1–15.