Normative reasoning in Mīmāmsā: a deontic logic approach

Agata Ciabattoni

Vienna University of Technology agata@logic.at



What I am going to present

interdisciplinar joint work with

- Sanskritists: E. Freschi and colleagues
- Logicians: K. van Berkel, F. Gulisano, B. Lellmann, M. Olszewski, ...
- Computer Scientists: E. Bartocci, G. Governatori, E. Neufeld, ...
- ▶ research project funded by WWTF (2017-2022)

Reasoning tools for Deontic Logic and Applications to Indian Sacred Texts

https://mimamsa.logic.at

The topic in a nutshell



Main idea:

Use *deontic* logic to formally analyse ancient texts of Sanskrit philosophy

Benefits:

- ► Indology:
 - Better understanding of the texts, (clarification through formalisation),
 - Some solutions and many challenging questions
- Logic: New inputs and development of new methods

The topic in a nutshell



Main idea:

Use deontic logic to formally analyse ancient texts of Sanskrit philosophy

Benefits:

- Indology:
 - Better understanding of the texts, (clarification through formalisation),
 - Some solutions and many challenging questions
- Logic: New inputs and development of new methods which might be used in AI



The Mīmāmsā school

 $M\bar{\imath}m\bar{a}ms\bar{a}$ is one of the most important schools of Indian philosophy

 \blacktriangleright Flourished from the last centuries BCE to the beginning of $20^{\rm th}$ c.

▶ Focus on the interpretation of the **prescriptive portions** of the *Vedas*

▶ Rational and a theistic analysis of the Vedas as a corpus of commands that **does not contain contradictions**

... The beginning of the story



বাৎকা হয় বুধাৰু। জুনিৰা এইখ- জুলবুৱে গৈ হৈ হয় হৈ বিশ্ব প্ৰয়ান বাবেৰু নগা এল বিবা স্থান জুৰা বিজে বিজে বিজে বিজে হয় হৈছিল আৰু নামৰা কাৰি নাম কিছে বিজে বিজ কৰি বিজ কৰি বিজ কৰি বিজ কৰি বিজ কৰি বিজ কয়। মধাৰ কৰা নাম কৰি নাম কৰি নাম কৰি বিজ কৰি বাৰে বিজ কৰি বিজ The beginning of the story

The Śyena controversy

श्येनेनाभिचरन् यजेत

1.

2.

न हिंस्यात् सर्वा भूतानि

The beginning of the story The Syena controversy

श्येनेनाभिचरन् यजेत

1.

 $\begin{array}{l} One \ should \ sacrifice \ bewitching \ with \\ \acute{Syena} \end{array}$

2. न हिंस्यात् सर्वा भूतानि

One should not harm any living being

- cannot be contradictory, because the Vedas are not (!).
- ▶ Many different **explanations** proposed all rational

Mīmāmsā & Logic

- Subject: inferencial reasoning, discussing (apparent) contradictions,...
- Method:
 - conclusions derived from reliable premisses via rigorous (and verifiable) reasoning steps
 (a text is not epistemically reliable if its beginning is not reliable, even if its whole chain of transmission is, as e.g. for "a chain of truthful blind people transmitting information concerning colours" (TV on PMS 1.3.27))
 - formulation of interpretative principles nyāyas (literally "rules", or "method") lends itself to formalisation

Mīmāmsā is considered early Deontic Logic

• Obstacles:

 $M\bar{i}m\bar{a}ms\bar{a}$ texts have not been formally analyzed due to:

- Lack of translations
- Highly technical language

Powerful tools: the nyāyas

Our logics/methods are defined formalizing

Mīmāmsā interpretative principles $(ny\bar{a}yas)$

i.e., **rules** formulated for the **interpretation** of Vedic **prescriptions**: hermeneutic, linguistic and **deontic**.

▶ Rational, scientific and systematic

also applied as **general** rules in Indian jurisprudence $(Dharmaś\bar{a}stra)$.

Powerful tools: the nyāyas

Our logics/methods are defined formalizing

Mīmāmsā interpretative principles $(ny\bar{a}yas)$

i.e., **rules** formulated for the **interpretation** of Vedic **prescriptions**: hermeneutic, linguistic and **deontic**.

▶ Rational, scientific and systematic

also applied as **general** rules in Indian juris prudence ($Dharmaś\bar{a}stra$).

...However



introduced in connection with concrete cases.

Examples of nyāyas

Linguistic

• A prescriptive sentence is unitary because of the unity of the purpose [it communicates]; (PMS 2.1.46).

Hermeneutic

a more specific rule overrules a generic one (PMS 3.1.26-27)

Logic

 When there is a contradiction, at the denial of one alternative, the other is known [to be true] (Jayanta's Nyāyamañjarī, 9th c. CE)

Chapter 1: From Mīmāmsā to Logic

A first step: basic Mīmāmsā Deontic Logic (AC, Freschi, Genco and Lellmann, Tableaux 2015) Guiding principle: to formalize and analyze Mīmāmsā reasoning



the logics/methods should only contain properties that can be traced back to $M\bar{i}m\bar{a}m\bar{s}\bar{a}$

A first step: basic Mīmāmsā Deontic Logic

(AC, Freschi, Genco and Lellmann, Tableaux 2015) Ingredients: A first step: basic Mīmāmsā Deontic Logic

(AC, Freschi, Genco and Lellmann, Tableaux 2015) Ingredients:

- ▶ Base Logic: Classical Logic
- Dyadic deontic operator O(/)
 O(φ/ψ) is for "φ is prescribed in case ψ is true".
 nyāya: "Each action is prescribed in relation to a responsible person who is identified because of her desire (adhikāra)"
- Modal operator \square (logic S4)

From the nyāyas to Hilbert axioms: an example

The properties of the **deontic operator** $\mathcal{O}(/)$ are extracted from $ny\bar{a}yas$. An example:

Rāmānujācārya's Tantrarahasya IV.4.3.3 (14th c. CE)

 \rightarrow Axiom (1):

$$\Box(\varphi \to \psi) \land \mathcal{O}(\varphi/\theta) \to \mathcal{O}(\psi/\theta)$$

From the nyāyas to Hilbert axioms: an example

The properties of the **deontic operator** $\mathcal{O}(/)$ are extracted from $ny\bar{a}yas$. An example:

"When, on the other hand, coming into being [of something needed], etc., are not realized by another prescription, [the principal prescription] itself begets the four [stages] of coming into being, etc., [of the prescriptions] connected to itself".

 \rightsquigarrow Axiom (1):

$$\Box(\varphi \to \psi) \land \mathcal{O}(\varphi/\theta) \to \mathcal{O}(\psi/\theta)$$

From the nyāyas to Hilbert axioms: an example

The properties of the **deontic operator** $\mathcal{O}(/)$ are extracted from $ny\bar{a}yas$. An example:

"When, on the other hand, coming into being [of something needed], etc., are not realized by another prescription, [the principal prescription] itself begets the four [stages] of coming into being, etc., [of the prescriptions] connected to itself". \sim (after many interactions with the Sanskritists) \sim

If a prescription enjoins something which has requirements, then it enjoins the requirements as well.

 \rightarrow Axiom (1):

$$\Box(\varphi \to \psi) \land \mathcal{O}(\varphi/\theta) \to \mathcal{O}(\psi/\theta)$$

.. Is it the right formalization?



Basic Mīmāmsā Deontic Logic (bMDL)

(AC, Genco, Freschi and Lellmann Tableaux 2015)

 \mathbf{bMDL} extends any Hilbert system for $\mathbf{S4}$ with the following axioms:

Mīmāmsā axioms

(1)
$$\Box(\varphi \to \psi) \land \mathcal{O}(\varphi/\theta) \to \mathcal{O}(\psi/\theta)$$

(Rāmānujācārya's Tantrarahasya IV.4.3.3)

(2)
$$\Box(\psi \to \neg \varphi) \to \neg(\mathcal{O}(\varphi/\theta) \land \mathcal{O}(\psi/\theta))$$

(Kumārila's Tantravārttika on PMS 1.3.3)

(3) $\Box((\chi \to \theta) \land (\theta \to \chi)) \land \mathcal{O}(\varphi/\chi) \to \mathcal{O}(\varphi/\theta)$

(ŚBh on PMS 6.1.25)

Basic Mīmāmsā Deontic Logic (bMDL) (AC, Genco, Freschi and Lellmann Tableaux 2015)

 \mathbf{bMDL} extends any Hilbert system for $\mathbf{S4}$ with the following axioms:

Mīmāmsā axioms

(1)
$$\Box(\varphi \to \psi) \land \mathcal{O}(\varphi/\theta) \to \mathcal{O}(\psi/\theta)$$

(2) $\Box(\psi \to \neg \varphi) \to \neg(\mathcal{O}(\varphi/\theta) \land \mathcal{O}(\psi/\theta))$

Given that purposes Y and Z exclude each other, if one should use X for the purpose Y, then it cannot be the case that one should use it at the same time for the purpose Z.

(3)
$$\Box((\chi \to \theta) \land (\theta \to \chi)) \land \mathcal{O}(\varphi/\chi) \to \mathcal{O}(\varphi/\theta)$$

If conditions X and Y are always equivalent, given the duty to perform Z under the condition X, the same duty applies under Y.

Sequent calculus for bMDL

Sequent calculus for bMDL

Standard propositional sequent rules + rules for S4 + Modal rules (B. Lellmann, D. Pattinson 2013):

$$\frac{\Gamma^{\Box} \Rightarrow \varphi}{\Gamma \Rightarrow \Box \varphi, \Delta} \ 4 \qquad \frac{\Gamma, \Box \varphi, \varphi \Rightarrow \Delta}{\Gamma, \Box \varphi \Rightarrow \Delta} \ T$$

$$\frac{\Gamma^{\Box}, \varphi \Rightarrow \psi \quad \Gamma^{\Box}, \theta \Rightarrow \chi \quad \Gamma^{\Box}, \chi \Rightarrow \theta}{\Gamma, \mathcal{O}(\varphi/\theta) \Rightarrow \mathcal{O}(\psi/\chi), \Delta}$$
 Mon

$$\frac{\Gamma^{\Box}, \varphi \Rightarrow}{\Gamma, \mathcal{O}(\varphi/\theta) \Rightarrow \Delta} D_1 \qquad \frac{\Gamma^{\Box}, \varphi, \psi \Rightarrow \Gamma^{\Box}, \theta \Rightarrow \chi \quad \Gamma^{\Box}, \chi \Rightarrow \theta}{\Gamma, \mathcal{O}(\varphi/\theta), \mathcal{O}(\psi/\chi) \Rightarrow \Delta} D_2$$

where Γ^{\Box} contains all formulae of the form $\Box \xi$ contained in Γ .

bMDL at work: an example

"Since the Veda is for the purpose of an action, whatever in it does not aim at an action is meaningless and therefore must be said not to belong to the permanent Veda" (PMS 1.2.1)

- no actual action can have a logical contradiction as an effect
- ▶ a logical contradiction cannot be enjoined by an obligation

 $\neg \mathcal{O}(\perp/\theta)$

is derivable in bMDL

(AC, Gulisano and Lellmann (AI and Law, 2021))

Prohibitions in Mīmāmsā are not definable from obligations

(AC, Gulisano and Lellmann (AI and Law, 2021))

▶ Prohibitions in Mīmāmsā are not definable from obligations
 ~ we have extended the base logic with the prohibition operator *F*(*A*/*B*)

(AC, Gulisano and Lellmann (AI and Law, 2021))

- Prohibitions in Mīmāmsā are not definable from obligations \rightsquigarrow we have extended the base logic with the prohibition operator $\mathcal{F}(A/B)$
- Vedic sacrifices are divided into: fixed sacrifices

 (nitya-karman), occasional sacrifices (naimittika-karman),
 and elective sacrifices (kāmya-karman). For Kumārila the
 latter have no deontic force

(AC, Gulisano and Lellmann (AI and Law, 2021))

- Prohibitions in Mīmāmsā are not definable from obligations \rightsquigarrow we have extended the base logic with the prohibition operator $\mathcal{F}(A/B)$
- Vedic sacrifices are divided into: fixed sacrifices

 (nitya-karman), occasional sacrifices (naimittika-karman),
 and elective sacrifices (kāmya-karman). For Kumārila the
 latter have no deontic force
 ~ the logic for Kumārila, extends that for Prabhākara with
 a new operator £(-/-) for kāmya-karman sacrifices

Chapter 2: The Śyena Controversy

Back to the Syena controversy

(van Berkel, AC, Freschi, Gulisano and Olszewski, submitted)

a millenary-old version of the Gentle Murder Paradox
 ((i) x is obliged not to kill, (ii) if x kills, x is obliged to kill gently, and (iii) gentle killing implies killing)



- 1. "One must perform the Śyena sacrifice, if one wants to harm his enemy"
- 2. "One must not perform violence on any living being"

Knowing that

- **3.** The Śyena harms the enemy
- 4. The enemy is a living being

Back to the Syena controversy

(van Berkel, AC, Freschi, Gulisano and Olszewski, submitted)

- ▶ a millenary-old version of the Gentle Murder Paradox
- three different explanations: Prabhākara, Kumārila (ca. 7th c. CE), and Maņḍana(ca. 8th c. CE)



- 1. "One must perform the Śyena sacrifice, if one wants to harm his enemy"
- 2. "One must not perform violence on any living being"

Knowing that

- **3.** The Śyena harms the enemy
- 4. The enemy is a living being

Back to the Śyena controversy: Prabhākara's view



- 1. "If one wants to harm his enemy, one must perform the Syena sacrifice" $\sim \mathcal{O}(\text{syena/des_harm})$
- 2. "One must not perform violence on any living being"
 → F(harm/T)
 - ▶ The Śyena harms the enemy, hence: syena → harm_e
 - ▶ The enemy is a living being, hence: $harm_e \rightarrow harm$

The Śyena is a Contrary-To-Duty obligation

"the Vedas do not impel one to perform the malevolent sacrifice Śyena, they only say that it is obligatory"

Back to the Śyena controversy: Kumārila's view



- 1. "If one wants to harm his enemy, one must perform the \hat{S} yena sacrifice" $\rightarrow \mathcal{E}(\text{syena/des_harm})$
- 2. "One must not perform violence on any living being"
 → F(harm/T)
 - ▶ The Śyena harms the enemy, hence: syena → harm_e
 - ▶ The enemy is a living being, hence: $harm_e \rightarrow harm$

No dilemma: Elective obligations have no deontic force

Back to the Śyena controversy: Maṇḍana's view



- 1. "If one wants to harm his enemy, one must perform the Syena sacrifice"
- 2. "One must not perform violence on any living being"
 - The Śyena harms the enemy
 - The enemy is a living being

He sees obligations as instruments toward the realisation of certain desires and solves the controversy with a cost-benefit approach

Chapter 3: Defeasible Reasoning in $M\bar{n}m\bar{a}ms\bar{a}$

.. Rules for resolving deontic conflicts

Non-monotonic consequence relation in Mīmāmsā

.. Not all the nyāyas can be simply converted into Hilbert axioms..

Ex.

Guṇapradhāna: "a more specific rule overrules a generic one". $(PMS \ 3.1.26-27)$

Non-monotonic consequence relation in $M\bar{n}m\bar{n}m\bar{s}n$

Ex.

Guṇapradhāna: "a more specific rule overrules a generic one". $(PMS \ 3.1.26-27)$

Without Specificity: if $\Gamma \vdash \alpha$, then $\Gamma \cup \{\beta\} \vdash \alpha$ independently from β

With Specificity: if $\Gamma \vdash \alpha$, and β is a rule more specific than α , then $\Gamma \cup \{\beta\} \not\vdash \alpha$

Example

Let Γ = "One should not remarry", α = "One should never take a second wife" and β = "One should take a second wife, if the first one is not virtuous or fertile"

Our approach to Gunapradhāna/Specificity

(AC, Gulisano and Lellmann DEON 2018, AI and Law, 2021)

Step 0 Explicit prima facie prescriptions in the Vedas O_{pf}(-/-) (śrauta) and propositional "facts" (statements about the world) → The Vedas are the primary source of information about duties and sacrifices

Step 1 Apply the **specificity** principle to generate all possible prescriptions

Step 2 Use the logic to reason about the resulting set of conflict-free obligations
→Mīmāṃsā authors tried to keep their arguments not defeasible "as much as possible" (J. Taber, 2004)

Step 1: sequent calculus for Gunapradhāna / Specificity

 $\Rightarrow \mathcal{O}(A/B)$

• it is entailed by a $\mathcal{O}_{pf}(C/D) \in \mathfrak{L}$

$$\{B \Rightarrow D\} \quad \cup \quad \{C \Rightarrow A\}$$

$$\mathcal{O}_R^{\mathcal{O}_{\mathsf{pf}}(C/D)}$$

• it is entailed by a $\mathcal{O}_{pf}(C/D) \in \mathfrak{L}$ which is not overruled by a conflicting more specific one

 $\Rightarrow \mathcal{O}(A)$

$$\begin{cases} B \Rightarrow D \} & \cup \quad \{C \Rightarrow A \} \\ \left\{ \lor \left\{ \begin{array}{c} \forall \cdot B \Rightarrow F \} \\ \{ \forall \cdot F \Rightarrow D \} \\ \{ \forall \cdot E, A \Rightarrow \} \end{array} \right\} \mid \mathcal{O}_{pf}(E/F) \in \mathfrak{L} \end{cases}$$

 $\mathcal{O}_R^{\mathcal{O}_{\mathsf{pf}}(C/D)}$

- it is entailed by a $\mathcal{O}_{pf}(C/D) \in \mathfrak{L}$ which is not overruled by a conflicting more specific one
- there is no other applicable conflicting obligation

. . .

- it is entailed by a $\mathcal{O}_{pf}(C/D) \in \mathfrak{L}$ which is not overruled by a conflicting more specific one
- there is no other applicable conflicting obligation which is not overruled itself by a more specific one

$$\begin{cases} \boldsymbol{B} \Rightarrow \boldsymbol{D} \} \quad \cup \quad \{\boldsymbol{C} \Rightarrow \boldsymbol{A} \} \\ \begin{cases} \left\{ \boldsymbol{\forall} \ \boldsymbol{B} \Rightarrow \boldsymbol{F} \right\} \\ \left\{ \boldsymbol{\forall} \ \boldsymbol{F} \Rightarrow \boldsymbol{D} \right\} \\ \left\{ \boldsymbol{\forall} \ \boldsymbol{E}, \boldsymbol{A} \Rightarrow \right\} \end{cases} \middle| \mathcal{O}_{pf}(\boldsymbol{E}/\boldsymbol{F}) \in \mathfrak{L} \end{cases} \\ \begin{cases} \left\{ \boldsymbol{\forall} \ \boldsymbol{B} \Rightarrow \boldsymbol{H} \right\} \\ \left\{ \boldsymbol{\forall} \ \boldsymbol{G}, \boldsymbol{A} \Rightarrow \right\} \\ \left\{ \boldsymbol{\forall} \ \boldsymbol{G}, \boldsymbol{A} \Rightarrow \right\} \\ \left\{ \begin{array}{c} \left\{ \boldsymbol{B} \Rightarrow \boldsymbol{J} \right\} \\ \cup \left\{ \boldsymbol{I} \Rightarrow \boldsymbol{H} \right\} & | \mathcal{O}_{pf}(\boldsymbol{I}/\boldsymbol{J}) \in \mathfrak{L} \\ \left\{ \boldsymbol{U} \right\} \\ \left\{ \boldsymbol{I} \Rightarrow \boldsymbol{A} \right\} \end{cases} \right\} \\ \Rightarrow \mathcal{O}(\boldsymbol{A}/\boldsymbol{B}) \end{cases} \qquad \mathcal{O}_{R}^{\mathcal{O}_{pf}(\boldsymbol{C}/\boldsymbol{D})}$$

Wait ... underivability premisses?

The underivability premisses look fishy and smell like circular definitions...

Fortunately:

Theorem.

Derivability for formulae of modal depth n + 1 depends only on derivability of formulae of modal depth at most n.

So everything is well-defined, we escape a fixpoint definition and even get

Theorem.

Derivability from assumptions is decidable in polynomial space.



> When there is a real conflict between obligations, any of the conflicting injunctions may be adopted as option.

> > Discussed by Jaimini $(2^{nd} c. BCE)$



> When there is a real conflict between obligations, any of the conflicting injunctions may be adopted as option.

> > Discussed by Jaimini $(2^{nd} c. BCE)$



> When there is a real conflict between obligations, any of the conflicting injunctions may be adopted as option.

> > Discussed by Jaimini $(2^{nd} c. BCE)$



> When there is a real conflict between obligations, any of the conflicting injunctions may be adopted as option.

> > Discussed by Jaimini $(2^{nd} c. BCE)$



When there is a real conflict between obligations, any of the conflicting injunctions may be adopted as option.

Discussed by Jaimini $(2^{nd} c. BCE)$



Defeasible reasoning in Mīmāmsā: the Bādhas

- ✓ ► Guṇapradhāna/Specificity
 ✓ ► Hierarchy of Sources

Four sources of duty: *śruti* (the Vedas), *smrti* (the 'recollected texts', based on the Vedas), $sad\bar{a}c\bar{a}ra$ (the behaviour of good people, who are learned in the Vedas) and $\bar{a}tmatusti$ (the inner feeling of approval by people who are learned in the Vedas)

- ▶ No empty rule
- Economicity Principle
- ► ...

Defeasible reasoning in Mīmāmsā: the Bādhas

- ▶ Guṇapradhāna/Specificity
- Hierarchy of Sources

Four sources of duty: \acute{sruti} (the Vedas), smrti (the 'recollected texts', based on the Vedas), $sad\bar{a}c\bar{a}ra$ (the behaviour of good people, who are learned in the Vedas) and $\bar{a}tmatusti$ (the inner feeling of approval by people who are learned in the Vedas)

▶ No empty rule

Every injunction has to be used. If the only way for applying a rule α is to block another one β (which is not, in turn, applicable only as long as α is blocked), α should block/suspend β .

• Economicity Principle

▶ .

Defeasible reasoning in Mīmāmsā: the Bādhas

- ▶ Guṇapradhāna/Specificity
- Hierarchy of Sources

Four sources of duty: \acute{sruti} (the Vedas), smrti (the 'recollected texts', based on the Vedas), $sad\bar{a}c\bar{a}ra$ (the behaviour of good people, who are learned in the Vedas) and $\bar{a}tmatusti$ (the inner feeling of approval by people who are learned in the Vedas)

▶ No empty rule

►

Every injunction has to be used. If the only way for applying a rule α is to block another one β (which is not, in turn, applicable only as long as α is blocked), α should block/suspend β .

Economicity Principle

a rule (set of rules) which invalidates as few injunctions as possible is preferable to one which invalidates many injunctions.

On the economicity principle

Towards AI applications





On the economicity principle: the Pac Man game



Pac-Man

- will eat 'food pellet' if it moves inside the cell.
- must avoid a collision with one of the ghosts
- when he comes into contact with a ghost, it dies and loses the game – unless the ghost is in a 'scared' state, which occurs for a set amount of time after Pac-Man consumes special pellet.

On the economicity principle: the Pac Man game



Pac-Man

- will eat 'food pellet' if it moves inside the cell.
- must avoid a collision with one of the ghosts
- when he comes into contact with a ghost, it dies and loses the game – unless the ghost is in a 'scared' state, which occurs for a set amount of time after Pac-Man consumes special pellet.

... (Noothigattu et al, IJCAI 2019) introduced Vegan Pac-Man: "Pac-Man must not eat ghosts" trained with Reinforcement Learning (RL)

Economicity principle at work: experimental results (AC, Neufeld, Bartocci and Governatori, submitted)

 We consider Vegan and Vegetarian Pac-Man : "Pac-Man must not eat blue ghosts"

We add a normative supervisor module to the RL agent



- working as (i) real-time compliance checker (ii) event recorder
- implemented with a theorem prover for defeasible deontic logic
- In case there are no compliant actions, the normative supervisor chooses the next action to do according to the economicity principle
- ▶ 100 Tests \sim very good results

(Some) Research directions



Reasoning tools for Deontic Logic and Applications to Indian Sacred Texts

https://mimamsa.logic.at

(Some) Research directions

- ▶ (*Find* and) Formalize new nyāyas and bādhas, refining the logics and the methods
- ▶ (*Find* and) Analyse new controversies from Mīmāmsā texts and compare them with dilemmas in modern deontic logic
- Provide a logical account of the difference between Mīmāmsā obligations and prohibitions, include desires, ...
- ► Use Logic for comparing the perspectives of the main Mīmāṃsā authors
- ▶ Use formal methods to provide answers (hints) to controversial issues in Mīmāṃsā
- Normative reasoning for AI applications?

