A Probabilistic Deontic Logic

Vincent de Wit, Dragan Doder(✉), and John Jules Meyer

Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands
{d.doder,J.J.C.Meyer}@uu.nl

Abstract. In this article, we introduce a logic for reasoning about probability of normative statements. We present its syntax and semantics, describe the corresponding class of models, provide an axiomatization for this logic and prove that the axiomatization is sound and complete. We also prove that our logic is decidable.

Keywords: Monadic deontic logic · Normative reasoning · Probabilistic logic · Completeness · Decidability

1 Introduction

The seminal work of von Wright from 1951 [14] initiated a systematic study on formalization of normative reasoning in terms of deontic logic. The latter is a branch of modal logics that deals with obligation, permission and related normative concepts. A plethora of deontic logics have been developed for various application domains like legal reasoning, argumentation theory and normative multi-agent systems [1,7].

Some recent work also studied learning behavioral norms from data [11,13]. In [11], the authors pointed out that human norms are context-specific and laced with uncertainty, which poses challenges to their representation, learning and communication. They gave an example of a learner that might conclude from observations that talking is prohibited in a library setting, while another learner might conclude the opposite when seeing people talking at the checkout counter. They represented uncertainty about norms using deontic operators, equipped with probabilistic boundaries that capture the subjective degree of certainty.

In this paper, we study uncertain norms form a logical point of view. We use probabilistic logic [3–6,12] to represent uncertainty, and we present the proof-theoretical and model-theoretical approach to a logic which allows reasoning about uncertain normative statements. We take two well studied logics, monadic deontic logic (MDL) [9] and probabilistic logic of Fagin, Halpern and Magido (FHM) [4], as the starting points, and combine them in a rich formalism that generalizes each of them. The resulting language makes it possible to adequately model different degrees of belief in norms; for example, we can express statements like "the probability that one is obliged to be quiet is at least 0.9".

J. Vejnarová and N. Wilson (Eds.): ECSQARU 2021, LNAI 12897, pp. 616–628, 2021.
https://doi.org/10.1007/978-3-030-86772-0_44

The semantics for our logic consists of specific Kripke-like structures, where each model contains a probability space whose sample space is the set of states, and with each state carrying enough information to evaluate a deontic formula. We consider so-called *measurable* models, which allow us to assign a probability value to every deontic statement.

The main result of this article is a sound and complete axiomatization for our logic. Like any other real-valued probabilistic logic, it is not compact, so any finitary axiomatic system would fail to be strongly complete ("every consistent set of formulas has a model") [6]. We prove weak completeness ("every consistent formula has a model") combining and modifying completeness techniques for MDL and FHM. We also show that our logic is decidable, combining the method of filtration and a reduction to a system of inequalities.

The rest of the paper is organised as follows: In Sect. 2 the proposed syntax and semantics of the logic will be presented together with other needed definitions. In Sect. 3 the axiomatization of the logic is given, followed by the soundness and completeness proof in Sect. 4. In Sect. 5 we show that our logic is decidable. Lastly, in Sect. 6 a conclusion is given together with future research topics.

2 Syntax and Semantics

In this section, we present the syntax and semantics of our probabilistic deontic logic. This logic, that we call $\mathcal{PDL}$, contains two types of formulas: standard deontic formulas of MDS, and probabilistic formulas. Let $\mathbb{Q}$ denote the set of rational numbers.

Definition 1 (Formula). *Let $\mathbb{P}$ be a set of atomic propositions. The language $\mathcal{L}$ of probabilistic monadic deontic logic is generated by the following two sentences of BNF (Backus Naur Form):*

$$[\mathcal{L}_{deontic}]\ \phi ::= p \mid \neg\phi \mid (\phi \wedge \phi) \mid O\phi \qquad p \in \mathbb{P}$$
$$[\mathcal{L}_{prob-d}]\ f ::= a_1 w(\phi_1) + \cdots + a_n w(\phi_n) \geq \alpha \mid \neg f \mid (f \wedge f) \qquad a_i, \alpha \in \mathbb{Q}$$

The set of all formulas $\mathcal{L}$ is $\mathcal{L}_{deontic} \cup \mathcal{L}_{prob-d}$. We denote the elements of $\mathcal{L}$ with θ and θ', possibly with subscripts.

The construct $O\phi$ is read as "It is obligatory that ϕ", while $w(\phi)$ stands for "probability of ϕ". An expression of the form $a_1 w(\phi_1) + \cdots + a_n w(\phi_n)$ is called *term*. We denote terms with x and t, possibly with subscripts. The propositional connectives, $\vee$, $\rightarrow$ and $\leftrightarrow$, are introduced as abbreviations, in the usual way. We abbreviate $\theta \wedge \neg\theta$ with $\bot$, and $\neg\bot$ with $\top$. We also use abbreviations to define other types of inequalities; for example, $w(\phi) \geq w(\phi')$ is an abbreviation for $w(\phi) - w(\phi') \geq 0$, $w(\phi) = \alpha$ for $w(\phi) \geq \alpha$ and $-w(\phi) \geq -\alpha$, $w(\phi) < \alpha$ for $\neg w(\phi) \geq \alpha$.

Example 1. *Following our informal example from the introduction about behavioral norms in a library, the fact that a person has become fairly certain that it* *is normal to be quiet might be expressed by the probabilistic statement "the probability that one is obliged to be quiet is at least 0.9". This sentence could be formalized using the introduced language as*

$$w(Oq) \geq 0.9.$$

Note that we do not allow mixing of the formulas from $\mathcal{L}_{deontic}$ and $\mathcal{L}_{prob-d}$. For example, $O(p \vee q) \wedge w(Oq) \geq 0.9$ is not a formula of our language. Before we introduce the semantics of $\mathcal{PDL}$ we will introduce Monadic Deontic Logic models.

Definition 2 (Relational model). *A relational model D is a tuple $D = (W, R, V)$ where:*

- *W is a (non-empty) set of states (also called "possible worlds"); W is called the universe of the model.*
- *$R \subseteq W \times W$ is a binary relation over W. It is understood as a relation of deontic alternativeness: sRt (or, alternatively, $(s,t) \in R$) says that t is an ideal alternative to s, or that t is a "good" successor of s. The first one is "good" in the sense that it complies with all the obligations true in the second one. Furthermore, R is subject to the following constraint:*

$$(\forall s \in W)(\exists t \in W)(sRt) \qquad (seriality)$$

  This means that the model does not have a dead end, a state with no good successor.
- *$V : \mathbb{P} \mapsto 2^W$ is a valuation function assigning to each atom p a set $V(p) \subseteq W$ (intuitively the set of states at which p is true.)*

Next, we define the satisfiability of a formula in a model. This definition is in accordance with standard satisfiability relation of MDL.

Definition 3 (Satisfaction in MDL). *Let $D = (W, R, V)$ be a relational deontic model, and let $w \in W$. We define the satisfiability of a deontic formula $\phi \in \mathcal{L}_{deontic}$ in the state w, denoted by $D, w \models_{MDL} \phi$, recursively as follows:*

- *$D, w \models_{MDL} p$ iff $w_s \in V_s(p)$.*
- *$D, w \models_{MDL} \neg\phi$ iff $D, s \not\models_{MDL} \phi$.*
- *$D, w \models_{MDL} \phi \wedge \psi$ iff $D, s \models_{MDL} \phi$ and $D, s \models_{MDL} \psi$.*
- *$D, w \models_{MDL} O\phi$ iff for all $u \in W_s$, if wRu then $D, u \models_{MDL} \phi$.*

Now we introduce the semantics of $\mathcal{PDL}$.

Definition 4 (Model). *A probabilistic deontic model is a tuple $M = \langle S, \mathscr{X}, \mu, \tau \rangle$, where*

- *S is a non-empty set of* states
- *$\mathscr{X}$ is a σ-algebra of subsets of S*
- *$\mu : \mathscr{X} \to [0, 1]$ is a probability measure, i.e.,*

  - $\mu(X) \geq 0$ *for all* $X \in \mathscr{X}$
  - $\mu(S) = 1$
  - $\mu(\bigcup_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} \mu(X_i)$, *if the* X_i*'s are pairwise disjoint members of* $\mathscr{X}$
- τ *associates with each state* s *in* S *a tuple containing a monadic deontic model and one of its worlds, i.e.,* $\tau(s) = (D_s, w_s)$, *where:*
  - $D_s = (W_s, R_s, V_s)$ *is a relational model of monadic deontic logic as defined in Definition 3.*
  - $w_s \in W_s$ *is a world* w_s *in* W_s *of model* D_s.

Let us illustrate this definition.

Example 1 ***(continued)*** *Assume a finite set of atomic propositions* $\{p, q\}$. *Let us consider the model* $M = \langle S, \mathscr{X}, \mu, \tau \rangle$, *where*

- $S = \{s, s', s'', s'''\}$
- $\mathscr{X}$ *is the set of all subsets of* S
- μ *is characterized by:* $\mu(\{s\}) = 0.5$, $\mu(\{s'\}) = \mu(\{s''\}) = 0.2$, $\mu(\{s'''\}) = 0.1$ *(other values follow from the properties of probability measures)*
- τ *is a mapping which assigns to the state* s, $D_s = (W_s, R_s, V_s)$ *and* w_s *such that*
  - $W_s = \{w_1, w_2, w_3, w_4\}$
  - $R_s = \{(w_1, w_2), (w_1, w_3), (w_2, w_2), (w_2, w_3), (w_3, w_2), (w_3, w_3), (w_4, w_2), (w_4, w_3), (w_4, w_4)\}$
  - $V_s(p) = \{w_1, w_3\}$, $V_s(q) = \{w_2, w_3\}$
  - $w_s = w_1$
    Note that the domain of τ *is always the whole set* S, *but in this example we only explicitly specify* $\tau(s)$ *for illustration purposes.*

This model is depicted in Fig. 1. The circle on the right contains the four states of the model, which are measured by μ. *Each of the states is equipped with a standard pointed model of MDL. In this picture, only one of them is shown, the one that corresponds to* s. *It is represented within the circle on the left. Note that the arrows depict the "good" alternative relation* R. *If we assume that* q *stands for "quiet", like in the previous example, in all good successors of* w_1 *the proposition* q *holds. Note that, according to Definition 3 , this means that in* w_1 *people are obliged to be quiet in the library.*

For a model $M = \langle S, \mathscr{X}, \mu, \tau \rangle$ and a formula $\phi \in \mathcal{L}_{deontic}$, let $\|\phi\|_M$ denote the set of states that satisfy ϕ, i.e., $\|\phi\|_M = \{s \in S \mid D_s, w_s \models_{MDL} \phi\}$. We omit the subscript M from $\|\phi\|_M$ when it is clear from context. The following definition introduces an important class of probabilistic deontic models, so-called *measurable* models.

Definition 5 (Measurable model). *A probabilistic deontic model is* measurable *if*

$$\|\phi\|_M \in \mathscr{X}$$

for every $\phi \in \mathcal{L}_{deontic}$.

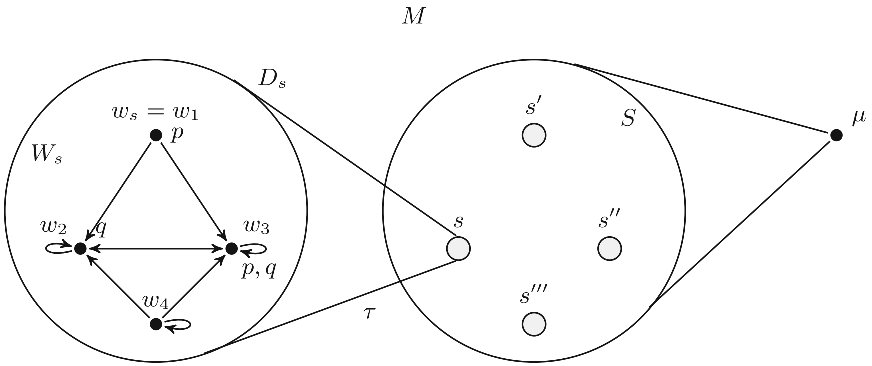

Fig. 1. Model $M = \langle S, \mathscr{X}, \mu, \tau \rangle$.

In this paper, we focus on measurable structures, and we prove completeness and decidability results for this class of structures.

Definition 6 (Satisfaction). *Let $M = \langle S, \mathscr{X}, \mu, \tau \rangle$ be a measurable probabilistic deontic model. We define the satisfiability relation $\models$ recursively as follows:*

- $M \models \phi$ *iff for all* $s \in S$, $D_s, w_s \models_{MDL} \phi$
- $M \models a_1 w(\phi_1) + \cdots + a_k w(\phi_k) \geq \alpha$ *iff* $a_1 \mu(\|\phi_1\|) + \cdots + a_k \mu(\|\phi_k\|) \geq \alpha$.
- $M \models \neg f$ *iff* $M \not\models f$
- $M \models f \wedge g$ *iff* $M \models f$ *and* $M \models g$.

Example 1 ***(continued)*** *Continuing the previous example, it is now also possible to speak about the probability of the obligation to be quiet in a library. First, according to Definition 3 it holds that* $D_s, w_s \models_{MDL} Oq$. *Furthermore, assume that* τ *is defined in the way such that* $D_{s'}, w_{s'} \models_{MDL} Oq$ *and* $D_{s''}, w_{s''} \models_{MDL} Oq$, but $D_{s'''}, w_{s'''} \not\models_{MDL} Oq$. Then $\mu(\|Oq\|) = \mu(\{s, s', s''\}) = 0.5 + 0.2 + 0.2 = 0.9$. *According to Definition 6,* $M \models w(Oq) \geq 0.9$.

Note that, according to Definition 6, a deontic formula is true in a model iff it holds in every state of the model. This is a consequence of our design choice that those formulas represent certain deontic knowledge, while probabilistic formulas express uncertainty about norms. At the end of this section, we define some standard semantical notions.

Definition 7 (Semantical consequence). *Given a set Γ of formulas, a formula θ is a semantical consequence of Γ (notation: $\Gamma \models \theta$) whenever, all the states of the model have, if $M, s \models \theta'$ for all $\theta' \in \Gamma$, then $M, s \models \theta$.*

Definition 8 (Validity). *A formula θ is valid (notations: $\models \theta$) whenever for $M = \langle S, \mathscr{X}, \mu, \tau \rangle$ and every $s \in S$: $M, s \models \theta$ holds.*

3 Axiomatization

The following axiomatization contains 13 axioms and 3 inference rules. It combines the axioms of proof system D of monadic deontic logic [9] with the axioms of probabilistic logic. The axioms for reasoning about linear inequalities are taken form [4].

The Axiomatic System: $AX_{\mathcal{PDL}}$

Tautologies and Modus Ponens

Taut. All instances of propositional tautologies.
MP. From θ and $\theta \rightarrow \theta'$ infer θ'.

Reasoning with O:

O-K. $O(\phi \rightarrow \psi) \rightarrow (O\phi \rightarrow O\psi)$
O-D. $O\phi \rightarrow P\phi$
O-Nec. *From ϕ infer $O\phi$.*

Reasoning About Linear Inequalities:

I1. $x \geq x$ (identity)
I2. $(a_1x_1 + ... + a_kx_k \geq c) \leftrightarrow (a_1x_1 + ... + a_kx_k + 0x_{k+1} \geq c)$ (adding and deleting 0 terms)
I3. $(a_1x_1 + ... + a_kx_k \geq c) \rightarrow (a_{j_1}x_{j_1} + ... + a_{j_k}x_{j_k} \geq c)$, if $j_1, ..., j_k$ is a permutation of $1, ..., k$ (permutation)
I4. $(a_1x_1 + ... + a_kx_k \geq c) \wedge (a'_1x_1 + ... + a'_kx_k \geq c') \rightarrow ((a_1 + a'_1)x_1 + ... + (a_k + a'_k)x_k \geq (c + c'))$ (addition of coefficients)
I5. $(a_1x_1 + ... + a_kx_k \geq c) \leftrightarrow (da_1x_1 + ... + da_kx_k \geq dc)$ if $d > 0$ (multiplication of non-zero coefficients)
I6. $(t \geq c) \vee (t \leq c)$ if t is a term (dichotomy)
I7. $(t \geq c) \rightarrow (t > d)$ if t is a term and $c > d$ (monotonicity).

Reasoning About Probabilities:

W1. $w(\phi) \geq 0$ (nonnegativity).
W2. $w(\phi \vee \psi) = w(\phi) + w(\psi)$, if $\neg(\phi \wedge \psi)$ is an instance of a classical propositional tautology (finite additivity).
W3. $w(\top) = 1$
P-Dis. From $\phi \leftrightarrow \psi$ infer $w(\phi) = w(\psi)$ (probabilistic distributivity).

The axiom Taut allows all $\mathcal{L}_{deontic}$-instances and $\mathcal{L}_{prob-d}$-instances of propositional tautologies. For example, $w(Oq) \geq 0.9 \vee \neg w(Oq) \geq 0.9$ is an instance of Taut, but $w(Oq) \geq 0.9 \vee \neg w(Oq) \geq 1$ is not. Note that Modus Ponens (MP) can be applied to both types of formulas, but only if θ and θ' are both from $\mathcal{L}_{deontic}$ or both from $\mathcal{L}_{prob-d}$. O-Nec is a deontic variant of necessitation rule. P-Dis is an inference rule which states that two equivalent deontic formulas must have the same probability values.

Definition 9 (Syntactical consequence). *A derivation of θ is a finite sequence $\theta_1, \dots, \theta_m$ of formulas such that $\theta_m = \theta$, and every θ_i is either an instance of an axiom, or it is obtained by the application of an inference rule to formulas in the sequence that appear before θ_i. If there is a derivation of θ, we say that θ is a* theorem *and write $\vdash \theta$. We also say that θ is derivable from a set of formulas Γ, and write $\Gamma \vdash \theta$, if there is a finite sequence $\theta_1, \dots, \theta_m$ of formulas such that $\theta_m = \theta$, and every θ_i is either a theorem, a member of Γ, or the result of an application of MP. or P-Nec. to formulas in the sequence that appear before θ_i.*

Note that this definition restricts the application of O-Nec. to theorems only. This is a standard restriction for modal necessitations, which enables one to prove Deduction theorem using induction on the length of the inference. Also, note that only deontic formulas can participate in a proof of another deontic formula, thus derivations of deontic formulas in our logic coincide with their derivations in MDL.

Definition 10 (Consistency). *A set Γ is consistent if $\Gamma \nvdash \bot$, and inconsistent otherwise.*

Now we prove some basic consequences of $AX_{\mathcal{PDL}}$. The first one is probabilistic variant of necessitation. It captures the semantical property that a deontic formula represents certain knowledge, and therefore it must have probability value 1. The third part of the lemma shows that a form of additivity proposed as an axiom in [4] is provable in $AX_{\mathcal{PDL}}$.

Lemma 1. *The following rules are derivable from our axiomatization:*

1. *From ϕ infer $w(\phi) = 1$*
2. $\vdash w(\bot) = 0$
3. $\vdash w(\phi \wedge \psi) + w(\phi \wedge \neg\psi) = w(\phi)$.

Proof.

1. Let us assume that a formula ϕ is derived. Then, using propositional reasoning (Taut and MP), one can infer $\phi \leftrightarrow \top$. Consequently, $w(\phi) = w(\top)$ follows from the rule P-Dis. Since we have that $w(\top) = 1$ (by W3), we can employ the axioms for reasoning about inequalities to infer $w(\phi) = 1$.
2. Then to show that $w(\bot) = 0$ using finite additivity (W2) $w(\top \vee \neg\top) = w(\top) + w(\neg\top) = 1$ and so $w(\neg\top) = 1 - w(\top)$. Since $w(\top) = 1$ and $\neg\top \leftrightarrow \bot$ we can derive $w(\bot) = 0$.
3. To derive additivity we begin with the propositional tautology, $\neg((\phi \wedge \psi) \wedge (\phi \wedge \neg\psi))$ then the following equation is given by W2 $w(\phi \wedge \psi) + w(\phi \wedge \neg\psi) = w((\phi \wedge \psi) \vee (\phi \wedge \neg\psi))$. The disjunction $(\phi \wedge \psi) \vee (\phi \wedge \neg\psi)$ can be rewritten to, $\phi \wedge (\psi \vee \neg\psi)$ which is equivalent to ϕ. From $\phi \leftrightarrow (\phi \wedge \psi) \vee (\phi \wedge \neg\psi)$, using P-Dis, we obtain $w(\phi) = w(\phi \wedge \psi) + w(\phi \wedge \neg\psi)$.

4 Soundness and Completeness

In this section, we prove that our logic is sound and complete with respect to the class of measurable models, combining and adapting the approaches from [2,4].

Theorem 1 (Soundness & Completeness). *The axiom system $AX_{\mathcal{PDL}}$ is sound and complete with respect to the class of measurable probabilistic deontic models. i.e., $\vdash \theta$ iff $\models \theta$.*

Proof. The proof of soundness is straightforward. To prove completeness, we need to show that every consistent formula θ is satisfied in a measurable model. Since we have two types of formulas, we distinguish two cases.

If $\theta \in \mathcal{L}_{deontic}$ we write θ as ϕ. Since ϕ is consistent and monadic deontic logic is complete [9], we know that there is a MDL model (W, R, V) and $w \in W$ such that $(W, R, V), w \models \phi$. Then, for any probabilistic deontic model M with only one state s and $\tau(s) = ((W, R, V), w)$ we have $M, s \models \phi$, and therefore $M \models \phi$ (since s is the only state); so the formula is satisfiable.

When $\theta \in \mathcal{L}_{prob-d}$ we write θ as f, and assuming consistency of, f we need to prove that it is satisfiable. First notice that f can be equivalently rewritten as a formula in disjunctive normal form,

$$f \leftrightarrow g_1 \vee \cdots \vee g_n$$

this means that satisfiability of f can be proven by showing that one of the disjuncts g_i of the disjunctive normal form of f is satisfiable. Note that every disjunct is of the form

$$g_i = \bigwedge_{j=1}^{r} (\sum_k a_{j,k} w(\phi_{j,k}) \geq c_j) \wedge \bigwedge_{j=r+1}^{r+s} \neg(\sum_k a_{j,k} w(\phi_{j,k}) \geq c_j)$$

In order to show that g_i is satisfiable we will substitute each weight term $w(\phi_{j,k})$ by a sum of weight terms that take as arguments formulas from the set Δ that will be constructed below. For any formula θ, let us denote the set of subformulas of θ by $Sub(\theta)$. Then, for considered, g_i we introduce the set of all deontic subformulas $Sub_{DL}(g_i) = Sub(g_i) \cap \mathcal{L}_{deontic}$. We create the set Δ as the set of all possible formulas that are conjunctions of formulas from $Sub_{DL}(g_i) \cup \{\neg e \mid e \in Sub_{DL}(g_i)\}$, such that for every e either e or $\neg e$ is taken as a conjunct (but not both). Then we can prove the following two claims about the set Δ:

- The conjunction of any two different formulas δ_k and δ_l from Δ is inconsistent: $\vdash \neg(\delta_k \wedge \delta_l)$. This is the case because for each pair of δ's at least one subformula $e \in Sub(\phi)$ such that $\delta_k \wedge \delta_l \vdash e \wedge \neg e$ and $e \wedge \neg e \vdash \bot$. If there is no such, e then by construction $\delta_k = \delta_l$.
- The disjunction of all δ's in Δ is a tautology: $\vdash \bigvee_{\delta \in \Delta} \delta$. Indeed, it is clear from the way the set Δ is constructed, that the disjunction of all formulas is an instance of a propositional tautology.

As noted earlier, we will substitute each term of each weight formula of g_i by a sum of weight terms. This can be done by using the just introduced set Δ and the set Φ, which we define as the set containing all deontic formulas $\phi_{j,k}$ that occur in the weight terms of g_i. In order to get all the relevant δ's to represent a weight term, we construct for each $\phi \in \Phi$ the set $\Delta_\phi = \{\delta \in \Delta \mid \delta \vdash \phi\}$ which contains all δ's that imply ϕ. Then we can derive the following equivalence:

$$\vdash \phi \leftrightarrow \bigvee_{\delta \in \Delta_\phi} \delta.$$

From the rule P-Dis we obtain

$$\vdash w(\phi) = w(\bigvee_{\delta \in \Delta_\phi} \delta).$$

Since any two elements of Δ are inconsistent, from W2 and axioms about inequalities we obtain $\vdash w(\bigvee_{\delta \in \Delta_\phi} \delta) = \sum_{\delta \in \Delta_\phi} w(\delta)$. Consequently, we have

$$\vdash w(\phi) = \sum_{\delta \in \Delta_\phi} w(\delta).$$

Note that some of the formulas δ's might be inconsistent (for example, a formula from Δ might be a conjunction in which both Op and $F(p \wedge q)$ appear as conjuncts). For an inconsistent formula δ, we have $\vdash \delta \leftrightarrow \bot$ and, consequently $\vdash w(\delta) = 0$, by the inference rule P-Dis. This can provably filter out the inconsistent δ's from each weight formula, using the axioms about linear inequalities. Thus, without any loss of generality, we can assume in the rest of the proof that all the formulas from Δ are consistent[1].

Lets us consider a new formula f', created by substituting each term of each weight formula of g_i:

$$f' = \left(\bigwedge_{j=1}^{r} (\sum_k a_{j,k} \sum_{\delta \in \Delta_{\phi_{j,k}}} w(\delta) \geq c_j) \right) \wedge \left(\bigwedge_{j=r+1}^{r+s} \neg (\sum_k a_{j,k} \sum_{\delta \in \Delta_{\phi_{j,k}}} w(\delta) \geq c_j) \right)$$

Then we will construct f'' by adding to f': a non-negativity constraint and an equality that binds the total probability weight of δ's to 1. In other words, f'' is the conjunction of the following formulas:

[1] We might introduce Δ^c and Δ^c_ϕ as the sets of all consistent formulas from Δ and Δ_ϕ, respectively, but since we will still have $\vdash w(\phi) = \sum_{\delta \in \Delta^c_\phi} w(\delta)$, we prefer not to burden the notation with the superscripts in the rest of the proof, and we assume that we do not have inconsistent formulas in Δ.

$$\sum_{\delta \in \Delta} w(\delta) = 1$$

$$\forall \delta \in \Delta \qquad w(\delta) \geq 0$$

$$\forall l \in \{1, \ldots, r\} \qquad \sum_k a_{l,k} \sum_{\delta \in \Delta_{\phi_{l,k}}} w(\delta) \geq c_l$$

$$\forall l \in \{r+1, \ldots, r+s\} \qquad \sum_k a_{l,k} \sum_{\delta \in \Delta_{\phi_{l,k}}} w(\delta) < c_l$$

Since the weights can be attributed independently while respecting the system of equations, the formula f'' is satisfiable if the following system of equations is solvable. With $I = \{1, \ldots, |\Delta|\}$:

$$\sum_{i=1}^{|\Delta|} x_i = 1$$

$$\forall i \in I \qquad x_i \geq 0$$

$$\forall l \in \{1, \ldots, r\} \qquad \sum_k a_{l,k} \sum_{i=1}^{|\Delta_{\phi_{l,k}}|} x_i \geq c_l$$

$$\forall l \in \{r+1, \ldots, r+s\} \qquad \sum_k a_{l,k} \sum_{i=1}^{|\Delta_{\phi_{l,k}}|} x_i < c_l$$

Each δ can be identified as a state in the universe of the probability structure. Since MDL is complete, each state in the probability structure corresponds with a pointed deontic model's state via the identification function τ. Furthermore, $w()$ abides to the rules of probability measures due to the axiom system. This means that to prove satisfiability, only a probability measure should be found that corresponds with the representation f'. By adding the constraints to the representation, we can find a probability measure by solving the system of linear inequalities f'' using the axioms for reasoning with inequalities I1-I7. We took f in the beginning of the proof to be a consistent formula and f is either satisfiable or unsatisfiable. When the system can be shown to be satisfiable we have proven completeness, satisfiability of f is proven when satisfiability of f'' is shown. This is the case because if f'' is satisfiable then so is f' which means g_i is satisfiable and if g_i is satisfiable then f is satisfiable. Assume f'' is unsatisfiable then $\neg f''$ is provable from the axioms I1-I7. As just explained f's satisfiability is equivalent to that of f''. Then $\neg f$ is provable, which means that f is inconsistent. This is a contradiction, and therefore we have to reject that f'' is unsatisfiable and to conclude that f is satisfiable.

5 Decidability

In this section, we prove that our logic is decidable. First, let us recall the satisfiability problem: given a formula θ, we want to determine if there exists a model M such that $M \models \theta$.

Theorem 2 (Decidability). *Satisfiability problem for $\mathcal{PDL}$ is decidable.*

Proof. (Sketch). Since we have two types of formulas, we will consider two cases. First, let us assume that $\theta \in \mathcal{L}_{deontic}$. We start with the well-known result that the problem of whether a formula from $\mathcal{L}_{deontic}$ is satisfiable in a standard monadic deontic model is decidable. It is sufficient to show that each $\theta \in \mathcal{L}_{deontic}$ is satisfiable in a monadic deontic model iff it is satisfiable under our semantics. First, if $(W', R', V'), w' \models \theta$ for some deontic model (W', R', V') and $w' \in W'$, let us construct the model $M = \langle S, \mathscr{X}, \mu, \tau\rangle$, with $S = \{s\}$, $\mathscr{X} = \{\emptyset, S\}$, $\mu(S) = 1$ and $\tau(s) = ((W', R', V'), w')$. Since $(W', R', V'), w' \models \theta$, then $M, s \models \theta$. From the fact that s is the unique state of M, we conclude that $M \models \theta$. On the other hand, if θ is not satisfiable in standard monadic deontic logic, then for every $M = \langle S, \mathscr{X}, \mu, \tau\rangle$ and $s \in S$ we will have $M, s \not\models \theta$, so $M \not\models \theta$.

Now, let us consider the case $\theta \in \mathcal{L}_{prob-d}$. In the proof, we use the method of filtration [2,8], and reduction to finite systems of inequalities. We only provide a sketch of the proof, since we use similar ideas as in our completeness proof. We will also use notation introduced in the proof of completeness. In the first part of the proof, we show that a formula is satisfiable iff it is satisfiable in a model with a finite number of (1) states and (2) worlds.

(1) First we show that if $\theta \in \mathcal{L}_{prob-d}$ is satisfiable, then it is satisfiable in a model with a finite set of states, whose size is at most $2^{|Sub_{DL}(\theta)|}$ (where $Sub_{DL}(\theta)$ is the set of deontic subformulas of θ, as defined in the proof of Theorem 1). Let $M = \langle S, \mathscr{X}, \mu, \tau\rangle$ be a model such that $M \models \theta$. Let us define by $\sim$ the equivalence relation over $S \times S$ in the following way: $s \sim s'$ iff for every $\phi \in Sub_{DL}(\theta)$, $M, s \models \phi$ iff $s' \models \phi$. Then the corresponding quotient set $S_{/\sim}$ is finite and $|S_{/\sim}| \leq 2^{|Sub_{DL}(\theta)|}$. Note that every C_i belongs to $\mathscr{X}$, since it corresponds to a formula δ_i of Δ (from the proof of Theorem 1), i.e., $C_i = \|\delta_i\|$. Next, for every equivalence class, C_i we choose one element and denote it s_i. Then we consider the model $M' = \langle S', \mathscr{X}', \mu', \tau'\rangle$, where:

- $S' = \{s_i \mid C_i \in S_{/\sim}\}$,
- $\mathscr{X}'$ is the power set of S',
- $\mu'(\{s_i\}) = \mu(C_i)$ such that $s_i \in C_i$ and for any $X \subseteq S'$, $\mu'(X) = \sum_{s_i \in X} \mu'(\{s_i\})$,
- $\tau'(s_i) = \tau(s_i)$.

Then it is straightforward to verify that $M' \models \theta$. Moreover, note that, by definition of M', for every $s_i \in S$ there is $\delta_i \in \Delta$ such that $M', s_i \models \delta_i$, and that for every $s_j \neq s_i$ we have $M', s_j \not\models \delta_i$. We therefore say that δ_i is the *characteristic formula* of s_i.

(2) Even if S' is finite, some sets of worlds attached to a state might be infinite. Now we will modify τ', in order to ensure that every $W(s_i)$ is finite, and of the size which is bounded by a number that depends on the size of θ. In this part of the proof we refer to the filtration method used to prove completeness of MDL [2], which shows that if a deontic formula ϕ is satisfiable, that it is satisfied in a world of a model $D(\psi) = (W, R, V)$ where the size of W is at most exponential wrt. the size of the set of subformulas of ϕ. Then we can replace τ' with a function τ'' which assigns to each s_i one such $D(\delta_i)$ and the corresponding world, where δ_i is the characteristic formula of s_i. We also assume that each $V(s_i)$ is restricted to the propositional letters from $Sub_{DL}(\theta)$. Finally, let $M'' = \langle S', \mathscr{X}', \mu', \tau'' \rangle$ It is easy to check that for every $\phi \in Sub_{DL}(\theta)$ and $s_i \in S'$, $M', s_i \models \phi$ iff $M'', s_i \models \phi$. Therefore, $M'' \models \theta$.

From the steps (1) and (2) it follows that in order to check if a formula $\theta \in \mathcal{L}_{prob-d}$ is satisfiable, it is enough to check if it is satisfied in a model $M = \langle S, \mathscr{X}, \mu, \tau \rangle$ in which S and each W_s (for every $s \in S$) are of finite size, bounded from above by a fixed number depending on the size of $|Sub_{DL}(\theta)|$. Then there are finitely many options for the choice of S and τ (i.e., (D_s, w_s), for every $s \in S$), and our procedure can check in finite time whether there is a probability measure μ for some of them, such that θ holds in the model. We guess S and τ and check whether we can assign probability values to the states from S, using translation to a system of linear inequalities, in the same way as we have done in the proof of Theorem 1. This finishes the proof, since the problem of checking whether a linear system of inequalities has a solution is decidable.

6 Conclusion

In this article, we introduced the probabilistic deontic logic $\mathcal{PDL}$, a logic in which we can reason about the probability of deontic statements. We proposed a language that extends both monadic deontic logic and probability logic from [4]. We axiomatized that language and proved soundness and completeness with respect to corresponding semantics. We also proved that our logic is decidable.

To the best of our knowledge, we are the first to propose a logic for reasoning about probabilistic uncertainty about norms. It is worth mentioning that there is a recent knowledge representation framework about probabilistic uncertainty in deontic reasoning obtained by merging deontic argumentation and probabilistic argumentation frameworks [10].

Our logic $\mathcal{PDL}$ used MDL as the underlying framework, we used this logic simply because it is one of the most studied deontic logics. On the other hand, MDL is also criticized because of some issues, like representation of contrary-to-duty obligations. It is important to point out that the axiomatization technique developed in this work can be also applied if we replace MDL with, for example, dyadic deontic logic, simply by changing the set of deontic axioms and the function τ in the definition of model, which would lead to a more expressive framework for reasoning about uncertain norms. Another avenue for future research is to extend the language by allowing conditional probabilities. In such a logic, it would be possible to express that one uncertain norm becomes more certain if another norm is accepted or learned.

References

1. Boella, G., van der Torre, L.W.N., Verhagen, H.: Introduction to normative multiagent systems. Comput. Math. Organ. Theor. **12**(2–3), 71–79 (2006). https://doi.org/10.1007/s10588-006-9537-7
2. Chellas, B.F.: Modal Logic: An Introduction. Cambridge University Press (1980). https://doi.org/10.1017/CBO9780511621192
3. Fagin, R., Halpern, J.Y.: Reasoning about knowledge and probability. J. ACM **41**(2), 340–367 (1994)
4. Fagin, R., Halpern, J.Y., Megiddo, N.: A logic for reasoning about probabilities. Inf. Computat. **87**(1), 78–128 (1990)
5. Frisch, A., Haddawy, P.: Anytime deduction for probabilistic logic. Artif. Intell. **69**, 93–122 (1994)
6. van der Hoek, W.: Some considerations on the logic pfd. J. Appl. Non Class. Logics **7**(3), 287–307 (1997)
7. Horty, J.F.: Agency and Deontic Logic. Oxford University Press (2001)
8. Hughes, G.E., Cresswell, M.J.: A Companion to Modal Logic. Methuen London, New York (1984)
9. Parent, X., Van Der Torre, L.: Introduction to Deontic Logic and Normative Systems. Texts in Logic and Reasoning. College Publications (2018). https://books.google.nl/books?id=IyUYwQEACAAJ
10. Riveret, R., Oren, N., Sartor, G.: A probabilistic deontic argumentation framework. Int. J. Approximate Reason. **126**, 249–271 (2020)
11. Sarathy, V., Scheutz, M., Malle, B.F.: Learning behavioral norms in uncertain and changing contexts. In: 2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 000301–000306 (2017). https://doi.org/10.1109/CogInfoCom.2017.8268261
12. Savic, N., Doder, D., Ognjanovic, Z.: Logics with lower and upper probability operators. Int. J. Approx. Reason. **88**, 148–168 (2017)
13. Tomic, S., Pecora, F., Saffiotti, A.: Learning normative behaviors through abstraction. In: Giacomo, G.D. (eds.) 24th European Conference on Artificial Intelligence, ECAI 2020, 29 August–8 September 2020, Santiago de Compostela, Spain, Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020). Frontiers in Artificial Intelligence and Applications, vol. 325, pp. 1547–1554. IOS Press (2020). https://doi.org/10.3233/FAIA200263
14. von Wrigth, G.H.: I. Deontic logic. Mind **LX**(237), 1–15 (1951). https://doi.org/10.1093/mind/LX.237.1