Reasoning About Probability via Continuous Functions

Tommaso Flaminio[1] , **Sandro Preto**[2] , **Sara Ugolini**[1]
[1] Artificial Intelligence Research Institute (IIIA - CSIC), Barcelona, Spain.
[2] Institute of Mathematics and Statistics, University of São Paulo, Brazil.
tommaso@iiia.csic.es, spreto@ime.usp.br, sara@iiia.csic.es

Abstract

For a functional representation of a logic we mean a representation of its formulas by means of (possibly real-valued) functions. Functional representations have shown to be a key tool for the study of (algebraizable) non-classical logics, since they allow one to approach the study of typical proof theoretical properties via the functional semantics. Interestingly, the functional representation for Łukasiewicz logic has been very recently shown to have an impact outside the purely logical realm; indeed, it can be applied to study properties of artificial neural networks. In this contribution we will provide a functional representation for a probability modal logic, $\mathrm{FP}(\text{Ł})$, that builds on Łukasiewicz calculus by adding to it a unary operator P reading "it is probable that". While this logic is not algebraizable in the usual sense, we still can provide a functional representation for its probability formulas. Our contribution will present two ways of providing a functional representation of the formulas of the modal logic $\mathrm{FP}(\text{Ł})$: a local one, that relies on de Finetti's coherence argument; and a global one that, instead, relies on probability distributions on a finite domain.

1 Introduction

Formulas of classical logic correspond to Boolean functions, i.e., characteristic functions of a certain subset X of a domain Ω; with the same flavor, formulas of a non-classical logic might be described by non-Boolean functions. Studying the class of functions representing the formulas of a certain logic is a line of research that goes under the name *functional representation*[1]. That of functional representation is a particularly active area among scholars working in fuzzy logics since, being the real unit interval $[0, 1]$ the usual domain of models for fuzzy logics, a formula φ from those systems naturally corresponds to a function $f_\varphi : [0, 1]^k \to [0, 1]$, where k is the number of variables occurring in φ. The book chapter (Aguzzoli, Bova, and Gerla 2011) goes through interesting cases and useful techniques on this subject.

Possibly the most relevant and most well studied fuzzy logic is Łukasiewicz's (Cignoli, D'Ottaviano, and Mundici 1999), and its functional representation was shown in (McNaughton 1951). McNaughton's theorem shows that formulas of Łukasiewicz logic are representable by $[0, 1]$-valued continuous and piecewise linear functions, namely, *McNaughton functions* (more details will be given in the next sections). Vice versa, for every such function f one can effectively construct a Łukasiewicz formula φ that represents it, i.e., such that $f = f_\varphi$, see (Mundici 1994).

Besides its theoretical interest, the representation of formulas via (possibly continuous) functions, and vice versa, has crossed the borders of mathematical logic and scholars from the AI community have found interesting ways to apply these representation results to the problem of integrating symbolic logic with learning algorithms. For instance the authors of (Giannini et al. 2018) apply functional representation to yield convex functional constraints; in (van Krieken, Acar, and van Harmelen 2022) the differentiability of fuzzy logical functions is studied to encode prior background knowledge and, by doing so, to help the training of a neural network. In addition it is worth recalling the recent attempts to provide a ground for the integration between logic and learning proposed in (Badreddine et al. 2022) and (Badia, Fagin, and Noguera 2023), to quote a few.

For the aim of the present paper it is worth highlighting that the functional representation of Łukasiewicz logic has been recently applied in (Preto and Finger 2022) in order to show that properties of binary classification neural networks are characterizable via logical properties of Łukasiewicz infinite-valued calculus. Other approaches in the direction of modeling neural networks via formulas of infinite-valued logical calculi can be found in (Di Nola, Lenzi, and Vitale 2016) and (Aguzzoli et al. 2021).

Notwithstanding the suitability of Łukasiewicz logic to interpret truth-degrees of formulas by real numbers, it is questionable whether the same propositional calculus is suitable to handle uncertainty degrees, such as probability values. Indeed, besides sharing some analogies, Łukasiewicz logic and probability logic are radically different in nature. For instance, while the former is truth-functional, the latter is not; e.g., the probability of a conjunctive formula $\varphi \wedge \psi$ cannot be computed only knowing the probability of φ and that of ψ. This distinction must be taken into account; indeed it marks a non negligible limitation in the applicability of Łukasiewicz logic, and hence of McNaughton functions,

[1] We warn the reader that the name "functional representation" is indeed used for several non-equivalent research topics in logic and computer science. What we mean for "functional representation" will be explained in a short while.

to learning problems that are supposed to handle uncertainty, such as predicting algorithms and neural networks for forecasting.

A suitable logic to handle probability values that builds on Łukasiewicz logic has been defined by Hájek, Esteva, and Godo by expanding the language of Łukasiewicz calculus by a unary modality P standing for "probably". This logic, denoted by $\mathrm{FP}(Ł)$, has formulas of the form $P(\varphi)$ that read as "probably φ" (or "φ is probable"). $\mathrm{FP}(Ł)$ has been shown to be sound and complete with respect to probability models (Hájek, Esteva, and Godo 1995) (see also (Hájek 1998)). Interestingly, in (Baldi, Cintula, and Noguera 2020) an almost immaterial variation of $\mathrm{FP}(Ł)$ has been shown to be syntactically interdefinable with, and hence equivalent to, Fagin, Halpern and Megiddo probability logic AX_{MEAS}, (Fagin, Halpern, and Megiddo 1990).

The aim of the present paper is to generalizes McNaughton theorem from Łukasiewicz to $\mathrm{FP}(Ł)$, so as to be able to handle applications where formulas represent probabilistic formulas. In more details, we presents two functional representations for the modal formulas of $\mathrm{FP}(Ł)$. These are obtained by suitably restricting the domain of McNaughton functions from the unit cube $[0,1]^k$ to an appropriate convex subset (or polytope). The choice of the polytope determining the domain of our functions will distinguish between what we call the *local* and the *global* representation: while the domain of functions in the local representation of formulas depends on the probabilistic events that occur in them, the global representation uses a general domain that only depends on the number of propositional variables of the starting classical language. For the local representation we apply de Finetti's approach to subjective probability (de Finetti 1935), whereas for the global representation, uncertainty degrees are encoded by probability functions on a finite algebra of events.

This paper is organized in the following sections: in Section 2 we will recall the basic notions on Boolean algebras and probability functions; in the same section we will also give the needed introduction to de Finetti's foundation of subjective probability theory. In Section 3 we will study coherence polytopes, that is to say, the geometrical objects that characterize de Finetti's theory. More precisely, we will show a description for these polytopes in terms of their extremal points. Łukasiewicz logic and its probabilistic modal expansion $\mathrm{FP}(Ł)$ are recalled in Section 4. The local functional representation for the modal formulas of $\mathrm{FP}(Ł)$ will be shown in Section 5, while Section 6 will deal with the global functional representation. We end this paper with Section 7 that presents some final comments and discusses our future work.

2 Probability Functions and de Finetti's Coherence in a Nutshell

We assume the reader to be familiar with the basic notions and results of classical propositional logic that are required to read the present paper. To facilitate the reading of this section we recall that, up to isomorphism, every finite Boolean algebra $\mathbf{A} = (A, \wedge, \vee, \neg, \bot, \top)$ is the algebra of subsets of a finite set X, the set of atoms of $\mathbf{A}$, with the operations of intersection, union and complementation, and the constants interpreted as $\emptyset$ and X with the obvious meaning. With this intuition in mind, we now move to recall the possibly less familiar de Finetti's theory of subjective probability theory.

Probability maps will be henceforth understood as functions from Boolean algebras to the real unit interval. More precisely a *probability function* on a Boolean algebra $\mathbf{A} = (A, \wedge, \vee, \neg, \bot, \top)$ is a map $\mu : \mathbf{A} \to [0,1]$ that is *normalized* ($\mu(\top) = 1$) and *finitely additive* ($\mu(a \vee b) = \mu(a) + \mu(b)$ for all $a, b \in A$ such that $a \wedge b = \bot$).

Recall that an element α of a Boolean algebra $\mathbf{A}$ is said to be an *atom* if $\alpha > \bot$, and if $\alpha \geq x \geq \bot$ then either $\alpha = x$ or $x = \bot$. Every finite Boolean algebra $\mathbf{A}$ is *atomic*, that is to say, there exists a subset $\mathrm{at}(\mathbf{A})$ of A such that $\alpha_i \in \mathrm{at}(\mathbf{A})$ is an atom of $\mathbf{A}$, $\bigvee_{\alpha \in \mathrm{at}(\mathbf{A})} \alpha = \top$ and for all $\alpha_i \neq \alpha_j \in \mathrm{at}(\mathbf{A})$, $\alpha_i \wedge \alpha_j = \bot$. Elements of a finite Boolean algebra are in one-one correspondence to (clearly finite) sets of atoms. Therefore, if an algebra $\mathbf{A}$ has n atoms $\{\alpha_1, \ldots, \alpha_n\}$, we can represent each of its elements by a n-element string in $\{0,1\}^n$. For instance an element $a \in A$ is represented by a $\{0,1\}$-string of length n such that each coordinate i of the string is 1 if $\alpha_i \leq a$ or $i = 0$ otherwise.

For every finitely additive probability function μ on a finite Boolean algebra $\mathbf{A}$, the restriction of μ to $\mathrm{at}(\mathbf{A})$ is a *probability distribution*, i.e. $\sum_{\alpha \in \mathrm{at}(\mathbf{A})} \mu(\alpha) = 1$ and for every $a \in A$, $\mu(a) = \sum_{\alpha \in \mathrm{at}(\mathbf{A}) : \alpha \leq \mathrm{a}} \mu(\alpha)$. For a finite algebra $\mathbf{A}$ the set of all probability distributions on $\mathrm{at}(\mathbf{A})$

$$\Delta_{\mathbf{A}} = \{d : \mathrm{at}(\mathbf{A}) \to [0,1] \mid \sum_{\alpha} \mathrm{d}(\alpha) = 1\}$$

is the $|\mathrm{at}(\mathbf{A})| - 1$ simplex. Every finitely additive probability function on $\mathbf{A}$ arises from one (and only one) distribution d from $\Delta_{\mathbf{A}}$.

Proposition 2.1. *For every finite Boolean algebra $\mathbf{A}$ there exists a bijective correspondence between $\Delta_{\mathbf{A}}$ and the set of all finitely additive probability functions on $\mathbf{A}$.*

Particularly important in the remainder of this paper are probabilities on finitely generated free Boolean algebras. Up to isomorphism, the n-generated free Boolean algebra is the Lindenbaum-Tarski algebra of formulas of classical propositional logic in a language with n variables. Conforming to a standard notation, we will denote the n-generated free Boolean algebra by $\mathrm{Free}(n)$. Also, we will indicate with lower case Greek letters $\varphi, \psi, \ldots$, with possible subscripts, both the formulas of classical logic and the elements of a free Boolean algebra like $\mathrm{Free}(n)$.

More or less in parallel with the Kolmogorovian axiomatic definition of probability functions just recalled[2], de Finetti proposed an alternative foundation to (*subjective*) probability theory on an ideal betting game between two players, a bookmaker and a gambler. These characters wager money on the occurrence of certain events $\psi_1, \ldots, \psi_k$. For each event ψ_i, gambler's payoffs are 1 if ψ_i occurs, and

[2]To be precise Kolmogorov axioms are for countably additive probabilities, while here we deal with only finitely additive ones.

0 otherwise. In this setting the *probability* of an event ψ_i becomes the *fair selling price* fixed by the bookmaker for it.

Bookmaker's prices for the events $\psi_1, \ldots, \psi_k$ will be referred to as *betting odds* and an assignment $\beta : \{\psi_1, \ldots, \psi_k\} \to [0, 1]$ of betting odds $\beta(\psi_i) = \beta_i$ will be called a *book*.

De Finetti had no particular inclination towards identifying events in a precise logical ground (Flaminio, Godo, and Hosni 2014). However, in order for his main result to be stated in precise terms, events will be understood as elements of a finitely generated free Boolean algebra $\mathrm{Free}(n)$, hence coded by Boolean formulas and denoted by lower case Greek letters. De Finetti's result then reads as follows. Let us fix finitely many events $\psi_1, \ldots, \psi_k$ and a book β on them. A gambler must choose *stakes* $\sigma_1, \ldots, \sigma_k \in \mathbb{R}$, one for each event, and pay to the bookmaker the amount $\sigma_i \cdot \beta_i$ for each ψ_i. When a (classical propositional) valuation w determines the truth value of each ψ_i, the gambler gains σ_i if $w(\psi_i) = 1$ and 0 otherwise. The book β is said to be *coherent* if there is no choice of stakes $\sigma_1, \ldots, \sigma_k \in \mathbb{R}$ such that for every valuation w

$$\sum_{i=1}^{k} \sigma_i \cdot \beta_i - \sum_{i=1}^{k} \sigma_i \cdot w(\psi_i) = \sum_{i=1}^{k} \sigma_i(\beta_i - w(\psi_i)) < 0.$$

The left hand side of the above expression captures the bookmaker's payoff, or *balance*, *relative to the book β under the valuation w*.

Note that a stake σ_i may be negative. Following tradition, money transfers are so oriented that "positive" means "gambler-to-bookmaker". Therefore, if $\sigma_i < 0$, the bookmaker is forced to swap her role with the gambler: she has to pay $-\sigma_i \cdot \beta(\psi_i)$ to the gambler in hopes of winning $-\sigma_i$ in case ψ_i occurs.

De Finetti's Dutch-Book theorem characterizes coherent books as follows.

Theorem 2.2. *Let $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ be a finite set of events and let $\beta : \mathcal{E} \to [0, 1]$ be a book on them. Then β is coherent iff there exists a finitely additive probability function $\mu : \mathrm{Free}(n) \to [0, 1]$ such that, for all i, $\mu(\psi_i) = \beta(\psi_i)$.*

3 Coherence Polytopes

De Finetti's theorem, that we recalled in the previous section, has a clear geometric description. Indeed, if $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ denotes the set of events on which the bookmaker is publishing her book, and if $w_1, \ldots, w_t$ stand for the (necessarily finitely many) logical valuations of events ψ_i to $\{0, 1\}$, define the polytope $\mathscr{C}_{\mathcal{E}}$ to be the convex hull of the t points in $\{0, 1\}^k$ obtained by evaluating the events from $\mathcal{E}$ by $w_1, \ldots, w_t$:

$$\mathscr{C}_{\mathcal{E}} = \mathrm{co}(\{\langle w_j(\psi_1), \ldots, w_j(\psi_k)\rangle \in \{0, 1\}^k \mid j = 1, \ldots, t\}).$$

These sets of the form $\mathscr{C}_{\mathcal{E}}$ will play a key role in this paper and they are called *coherence sets* due to the following result that provides a geometric version of de Finetti's theorem.

Theorem 3.1. *Let $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ be a finite set of events and let $\beta : \mathcal{E} \to [0, 1]$ be a book on them. Then β is coherent iff $\langle \beta(\psi_1), \ldots, \beta(\psi_k)\rangle \in \mathscr{C}_{\mathcal{E}}$.*

Finally, let us recall that for every finite set $\mathcal{E}$ of events, $\mathscr{C}_{\mathcal{E}}$ is a polytope (i.e., the convex closure of finitely many points, (Ewald 1996)) of the cube $[0, 1]^k$ (see for instance (de Finetti 1935) and (Flaminio and Ugolini 2023)).

Definition 3.2. A polytope $\mathscr{X} \subseteq [0, 1]^k$ is said to be a *coherence polytope* if there exists a set of events $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ such that $\mathscr{X} = \mathscr{C}_{\mathcal{E}}$.

By Krein-Milman theorem (Krein and Milman 1940), every convex set in a finite dimensional space is the convex hull of its extremal points. The next lemma, that at the best of the authors' knowledge is new, provides a first non trivial fully geometric description of coherence sets.

Lemma 3.3. *Let $\mathscr{X} \subseteq [0, 1]^k$ be a polytope whose extremal points are contained in $\{0, 1\}^k$. Then $\mathscr{X}$ is a coherence set. That is, there exists a set of events $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ such that $\mathscr{X} = \mathscr{C}_{\mathcal{E}}$.*

Proof. The idea is to reverse the construction that defines a coherence set like $\mathscr{C}_{\mathcal{E}}$ from $\mathcal{E}$. Let us then display the extremal points $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_t$ of $\mathscr{X}$ as

$$\begin{array}{lcl} \mathbf{e}_1 & = & \langle e_{1,1}, e_{1,2}, \ldots, e_{1,k}\rangle \\ \mathbf{e}_2 & = & \langle e_{2,1}, e_{2,2}, \ldots, e_{2,k}\rangle \\ \cdots & & \\ \mathbf{e}_t & = & \langle e_{t,1}, e_{t,2}, \ldots, e_{t,k}\rangle. \end{array}$$

For every $j = 1, \ldots k$, let

$$\mathbf{a}_j = \langle e_{1,j}, e_{2,j}, \ldots, e_{t,j}\rangle \in \{0, 1\}^t.$$

Each $\mathbf{a}_j$ is hence an element of the Boolean algebra $\mathbf{A}$ with $|\mathrm{at}(\mathbf{A})| = \mathrm{t}$. If $t = 2^n$ for some n then $\mathbf{A}$ is, up to isomorphism, the algebra of formulas with n variables $\mathrm{Free}(n)$. Thus, for each j, there exists a formula ψ_j such that $\mathbf{a}_j = \psi_j$. By construction it is hence straightforward to see that $\mathscr{X} = \mathscr{C}_{\mathcal{E}}$ for $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$.

Otherwise, let n be the minimum positive integer such that $t < 2^n$. Hence, there is an injective homomorphism $\iota : \mathbf{A} \to \mathrm{Free}(n)$ (see for instance (Cramer 1970, Theorem 4)). Therefore let, for all $j = 1, \ldots, k$, $\psi_j = \iota(\mathbf{a}_j)$. Again, it is almost immediate to see that, calling $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$, one has $\mathscr{X} = \mathscr{C}_{\mathcal{E}}$. Thus the claim is settled. $\square$

The previous lemma hints at a connection between formulas and coherence sets. However, while every finite set $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ of events determines a unique coherence polytope $\mathscr{C}_{\mathcal{E}}$, for every polytope $\mathscr{X}$ with extremal points in $\{0, 1\}^k$ there are in general several sets of events that have $\mathscr{X}$ as a coherence set.

4 Łukasiewicz Logic and Its Probability Expansion FP(Ł)

Łukasiewicz logic is the formal system based on a propositional language $\mathcal{L}$ containing a denumerable set of variables Var, the binary operator $\oplus$ usually understood as Łukasiewicz *strong disjunction* connective, the unary operator $\neg$ standing for the *negation* connective and a constant $\bot$ for *falsum*. Formulas are inductively defined as usual and, to distinguish them from those of classical logic recalled in Section 2, are here denoted by capital Greek letters

$\Phi, \Psi, \ldots$. A *valuation* is hence a function $v : Var \to [0,1]$ extending to all formulas in the following way: $v(\Phi \oplus \Psi) = \min(1, v(\Phi) + v(\Psi))$, $v(\neg\Phi) = 1 - v(\Phi)$ and $v(\bot) = 0$.

From the basic connectives just introduced, other operators are defined as follows:

$\Phi \odot \Psi := \neg(\neg\Phi \oplus \neg\Psi)$; $\Phi \to \Psi := \neg\Phi \oplus \Psi$;
$\Phi \vee \Psi := (\Phi \to \Psi) \to \Psi$; $\Phi \wedge \Psi := \neg(\neg\Phi \vee \neg\Psi)$;
$\Phi \equiv \Psi := (\Phi \to \Psi) \wedge (\Psi \to \Phi)$.

Their semantics in $[0,1]$ is given by:

$v(\Phi \odot \Psi) = \max(0, v(\Phi) + v(\Psi) - 1)$; $\Phi \to \Psi = \min(1, 1 - v(\Phi) + v(\Psi))$; $\Phi \vee \Psi = \max(v(\Phi), v(\Psi))$; $\Phi \wedge \Psi = \min(v(\Phi), v(\Psi))$; $\Phi \equiv \Psi = 1 - |\max(v(\Phi) - v(\Psi))|$.

Łukasiewicz logic is sound and complete with respect to the so called *standard MV-algebra*, that is to say, the algebra $[0,1]_{MV} = ([0,1], \oplus, \neg, 0)$ where for all $x, y \in [0,1]$, $x \oplus y = \min\{1, x+y\}$ and $\neg x = 1 - x$.

For the next definition we refer to (Cignoli, D'Ottaviano, and Mundici 1999) and (McNaughton 1951).

Definition 4.1. A *McNaughton* function $f : [0,1]^k \to [0,1]$ is a function that is continuous with respect to the usual topology of $[0,1]$, and such that there are linear polynomials $p_1, \ldots, p_m$ over $[0,1]^k$ with integer coefficients such that, for every $x \in [0,1]^k$, there exists $i \in \{1, \ldots, m\}$ and $f(x) = p_i(x)$.

As recalled in the first section, McNaughton functions provide a functional representation for Łukasiewicz formulas. In the next result, and elsewhere in this paper, for every Łukasiewicz formula Φ on variables $q_1, \ldots, q_k$, and for every valuation of the q_i's to the standard MV-algebra $[0,1]_{MV}$, we denote by $\Phi^{[0,1]}(v(q_1), \ldots, v(q_k))$ the value that the interpretation of Φ in $[0,1]_{MV}$ gives to the vector $\langle v(q_1), \ldots, v(q_k)\rangle \in [0,1]^k$.

Theorem 4.2 ((Mundici 1994))**.** *For every McNaughton function $f : [0,1]^k \to [0,1]$, there exists a Łukasiewicz formula Φ on k propositional variables $q_1, \ldots, q_k$ such that for every valuation v,*

$$\begin{aligned} v(\Phi) &= f(v(q_1), \ldots, v(q_k)) \\ &= \Phi^{[0,1]}(v(q_1), \ldots, v(q_k)) \end{aligned}$$

It is clear that valuations of Łukasiewicz language on k variables to $[0,1]$ and points $\mathbf{x} = (x_1, \ldots, x_k) \in [0,1]^k$ are in bijective correspondence. We will henceforth denote, for every $\mathbf{x} \in [0,1]^k$, by $v_{\mathbf{x}}$ its corresponding valuation. It is convenient to restate the above theorem as follows.

Corollary 4.3. *For every McNaughton function $f : [0,1]^k \to [0,1]$, there exists a Łukasiewicz formula Φ on k propositional variables $q_1, \ldots, q_k$ such that for every $\mathbf{x} \in [0,1]^k$,*

$$f(x_1, \ldots, x_k) = v_{\mathbf{x}}(\Phi).$$

We are now in the position to define the logic FP(Ł) whose language is that of Łukasiewicz logic plus a unary modality P. We stress that formulas are of a special kind:

Non-modal formulas: these are built from a set V of propositional variables $\{q_1, q_2, \ldots\}$ using the classical binary connectives $\wedge$ and $\neg$ and the constant $\bot$. Other connectives like $\vee$, $\to$ and $\equiv$ are defined from $\wedge$ and $\neg$ in the usual way and as specified above. Non-modal formulas (we will also refer to them as Boolean propositions) will be denoted by lower case Greek letters φ, ψ, etc.

Modal formulas: these are built from *basic modal formulas* of the form $P\varphi$, where φ is a non-modal formula and using the connectives and constants of Łukasiewicz logic Ł. Adopting the same notation used for Łukasiewicz formulas, we denote them by upper case Greek letters Φ, Ψ, etc with no danger of confusion. (Notice that we do not allow nested modalities of the form $P(P(\psi) \oplus P(\varphi))$, nor mixed formulas of the kind $\psi \to P\varphi$).

Definition 4.4. The axioms of the logic FP(Ł) are the following:

(i) Axioms of classical propositional logic for non-modal formulas
(ii) Axioms of Łukasiewicz logic for modal formulas
(iii) Probabilistic modal axioms:
(FP0) $P\varphi$, for φ being a theorem of classical propositional logic
(FP1) $P(\varphi \to \psi) \to (P\varphi \to P\psi)$
(FP2) $P(\neg\varphi) \equiv \neg P\varphi$
(FP3) $P(\varphi \vee \psi) \equiv (P\varphi \to P(\varphi \wedge \psi)) \to P\psi$

The only deduction rule of FP(Ł) is that of Ł (i.e. *modus ponens*). The notion of proof is defined as usual in a Hilbert-style calculus; for a modal formula Φ we will write $\vdash_{FP} \Phi$ to denote that Φ is a *theorem* of FP(Ł).

Let us remark that every (compound) modal formula of FP(Ł) is hence nothing else than a Łukasiewicz formula Φ having for variables basic modal formulas, say $P(\psi_1), \ldots, P(\psi_k)$. We will write $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ to highlight the basic modal formulas occurring in Φ.

The next definition will be useful in what follows.

Definition 4.5. Two formulas Φ_1 and Φ_2 of FP(Ł) are said to be *logically equivalent* iff $\vdash_{FP} \Phi_1 \equiv \Phi_2$.

Formulas of FP(Ł) are evaluated by probability functions on free Boolean algebras. Indeed, if $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ is a modal formula, denote by $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ the set of non-modal Boolean formulas (i.e., the set of *events*) occurring in it, and let $\mathrm{Free}(n)$ be the free Boolean algebra over the variables appearing in $\mathcal{E}$. Then, if $\mu : \mathrm{Free}(n) \to [0,1]$ is a probability function, we evaluate Φ via μ as follows:

$$\mu(\Phi[P(\psi_1), \ldots, P(\psi_k)]) = \Phi^{[0,1]}(\mu(\psi_1), \ldots, \mu(\psi_k)),$$

where $\Phi^{[0,1]}$ stands for the interpretation of the formula Φ in the standard MV-algebra over $[0,1]$.

Let Φ be as above. By de Finetti's theorem (Theorem 2.2), if $\mathbf{x} \in \mathscr{C}_{\mathcal{E}}$ then every probability μ on $\mathrm{Free}(n)$ that extends $\mathbf{x}$ agrees on $\psi_1, \ldots, \psi_k$. Therefore, we can simply denote by $\mu_{\mathbf{x}}$ a generic probability function that extends the coherent book $\mathbf{x}$ and write

$$\mu_{\mathbf{x}}(\Phi[P(\psi_1), \ldots, P(\psi_k)]).$$

Notice that in this case

$$\mu_{\mathbf{x}}(\Phi[P(\psi_1), \ldots, P(\psi_k)]) = \Phi^{[0,1]}(\mu_{\mathbf{x}}(\psi_1), \ldots, \mu_{\mathbf{x}}(\psi_k)).$$

The next result is as a special case of (Flaminio and Ugolini 2023, Corollary 3.5).

Theorem 4.6. *The logic* $\mathrm{FP}(Ł)$ *is sound and complete with respect to the class of probability functions* $\mu_{\mathbf{x}}$. *In particular, if* $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ *is not a theorem of* $\mathrm{FP}(Ł)$*, then letting* $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ *there exists* $\mathbf{x} \in \mathscr{C}_{\mathcal{E}}$ *such that* $\mu_{\mathbf{x}}(\Phi[P(\psi_1), \ldots, P(\psi_k)]) < 1$.

An immediate consequence of the above theorem is that, if $\Phi_1[P(\psi_1), \ldots, P(\psi_k)]$ and $\Phi_2[P(\psi_1), \ldots, P(\psi_k)]$ are two modal formulas on the same set $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ of events, then Φ_1 and Φ_2 are logically equivalent formulas of FP(Ł) iff $\mu_{\mathbf{x}}(\Phi_1[P(\psi_1), \ldots, P(\psi_k)]) = \mu_{\mathbf{x}}(\Phi_2[P(\psi_1), \ldots, P(\psi_k)])$ for every $\mathbf{x} \in \mathscr{C}_{\mathcal{E}}$.

5 A Local Representation of Probability Formulas via Coherence Sets

We start by considering a formula $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ of FP(Ł). The outer formula Φ, regarded as a propositional Łukasiewicz formula on variables $q_1, \ldots, q_k$, corresponds to a McNaughton function $f_\Phi : [0,1]^k \to [0,1]$. We will now provide a functional representation for $\Phi[P(\psi_1), \ldots, P(\psi_k)]$. The idea is to represent the correct ways of evaluating the basic probability formulas $P(\psi_1), \ldots, P(\psi_k)$, in the same way every possible $\mathbf{x} \in [0,1]^k$ determines a unique Łukasiewicz valuation for the formula $\Phi(q_1, \ldots, q_k)$ (recall Corollary 4.3). This amounts to evaluating the basic modal formulas $P(\psi_i)$'s by means of probability functions. As in Section 3, let us call $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ the set of classical events appearing in Φ.

Since by Theorem 4.6 the basic formulas $P(\psi_i)$'s are evaluated in points of the convex set $\mathscr{C}_{\mathcal{E}} \subseteq [0,1]^k$, it should be clear that the formula $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ can be represented by the McNaughton function f_Φ once restricted to $\mathscr{C}_{\mathcal{E}}$.

Proposition 5.1. *For every formula* $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ *the McNaughton function* f_Φ *is such that for every* $\mathbf{x} = (x_1, \ldots, x_k) \in \mathscr{C}_{\mathcal{E}}$

$$f_\Phi(x_1, \ldots, x_k) = \mu_{\mathbf{x}}(\Phi[P(\psi_1), \ldots, P(\psi_k)]).$$

Our aim is now to show the converse, and indeed interesting, direction of the above result. In other words, we want to prove that every McNaughton function restricted to a coherence set represents a formula of FP(Ł). For this direction we will make use of the results shown in Section 3. To facilitate the reading, we will henceforth adopt the notation defined below.

Definition 5.2. For every $k \in \mathbb{N}$ and every subset $\mathscr{X}$ of $[0,1]^k$, we will say that a function $f : \mathscr{X} \to [0,1]$ is an $\mathscr{X}$-*McNaughton function* if it is the restriction of a McNaughton function $g : [0,1]^k \to [0,1]$ to $\mathscr{X}$, i.e., $f = g\restriction_{\mathscr{X}}$.

Clearly $[0,1]^k$-McNaughton functions are McNaughton functions in their usual form. We observe in passing that if $\mathscr{X}$ is closed in the product topology of $[0,1]^k$, then $\mathscr{X}$-McNaughton functions are, up to isomorphism, elements of the semisimple MV-algebra obtained as the quotient $\mathrm{Free}(k)/\mathscr{X}$ (see e.g., (Cignoli, D'Ottaviano, and Mundici 1999, Theorem 3.6.7)). In the next result we will show how to represent formulas of FP(Ł) by means of $\mathscr{X}$-McNaughton functions, for appropriate choices of $\mathscr{X}$.

Theorem 5.3. *For every McNaughton function* $f : [0,1]^k \to [0,1]$ *and for every coherence polytope* $\mathscr{X} \subseteq [0,1]^k$, *there exists a probabilistic formula* $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ *such that for all* $\mathbf{x} = (x_1, \ldots, x_k) \in \mathscr{X}$,

$$f_\Phi(x_1, \ldots, x_k) = \mu_{\mathbf{x}}(\Phi[P(\psi_1), \ldots, P(\psi_k)]).$$

Proof. Via Theorem 4.2, let $\Phi[x_1, \ldots, x_k]$ be the Łukasiewicz formula such that $f = f_\Phi$. Now, by Lemma 3.3, there is a set of events $\mathcal{E} = \{\psi_1, \ldots, \psi_k\}$ such that $\mathscr{X} = \mathscr{C}_{\mathcal{E}}$. Then $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ is a FP(Ł) formula satisfying the claim. Indeed, for every $\mathbf{x} = (x_1, \ldots, x_k) \in \mathscr{X}$, $\mu_{\mathbf{x}}(\Phi[P(\psi_1), \ldots, P(\psi_k)]) = \Phi^{[0,1]}(\mu_{\mathbf{x}}(\psi_1), \ldots, \mu_{\mathbf{x}}(\psi_k))$. Since $\mu_{\mathbf{x}}(\psi_i) = x_i$, we finally get that $\Phi^{[0,1]}(\mu_{\mathbf{x}}(\psi_1), \ldots, \mu_{\mathbf{x}}(\psi_k)) = f_\Phi(x_1, \ldots, x_k)$. $\square$

A first consideration on the functional representation shown above concerns logically equivalent formulas. Indeed, logically equivalent formulas might be represented by formally different functions. Let us focus on the following example.

Example 5.4. Let us consider the two probability formulas $\Phi_1[P(q)] = P(q) \wedge \neg P(q)$ and $\Phi_2[P(q), P(\neg q)] = P(q) \wedge P(\neg q)$. By axiom (FP2), $\neg P(q) \equiv P(\neg q)$ and since FP(Ł) satisfies the rule of substitution of equivalents, it immediately follows that Φ_1 and Φ_2 are logically equivalent.

However, notice that Φ_1 and Φ_2 are defined on different sets of events and the functions that represent Φ_1 and Φ_2 are formally different. Indeed, the McNaughton function f_{Φ_1} is such that

$$f_{\Phi_1} : x \in [0,1] \mapsto x \wedge \neg x \in [0,1],$$

while f_{Φ_2} is defined as follows:

$$f_{\Phi_2} : (x, y) \in [0,1]^2 \to x \wedge y \in [0,1].$$

The next step is to determine the coherence polytopes to restrict each f_{Φ_i}. As for $\mathscr{C}_{\{q\}}$, since q is a propositional variable, every book on q is coherent. Thus, $\mathscr{C}_{\{q\}} = [0,1]$ (see Figure 1).

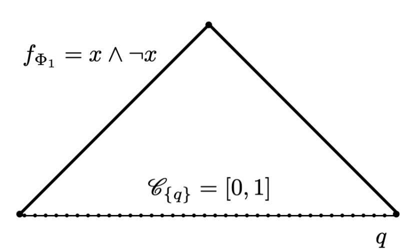

Figure 1: The coherence polytope $\mathscr{C}_{\{q\}} = [0,1]$ (dotted line) and the McNaughton function $f_{\Phi_1} = x \wedge \neg x$.

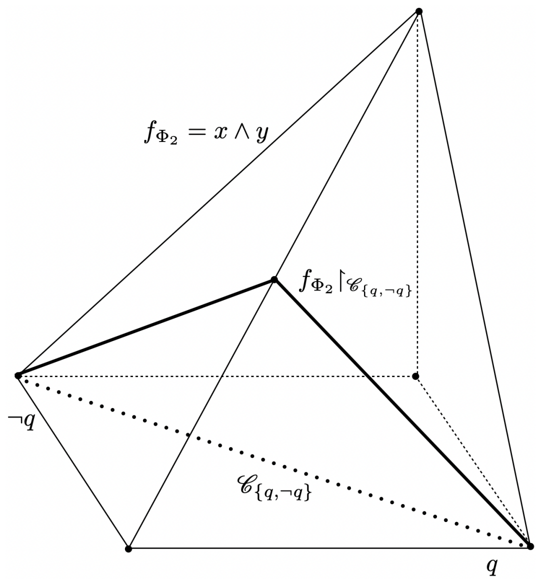

Figure 2: The coherence polytope $\mathscr{C}_{\{q,\neg q\}} = \{(\beta, 1-\beta) \mid \beta \in [0,1]\}$ (dotted line), the McNaughton function $f_{\Phi_2} = x \wedge y$ and its restriction to $\mathscr{C}_{\{q,\neg q\}}$ (boldfaced).

On the other hand, in the case of Φ_2, the events involved are q and $\neg q$, e.g., the typical events of a coin tossing game. Thus $\mathscr{C}_{\{q,\neg q\}} = \{(\beta, 1-\beta) \mid \beta \in [0,1]\}$ (see Figure 2).

Direct inspection of the Figures 1 and 2 shows that the functions $f_{\Phi_1}\restriction_{\mathscr{C}_{\{q\}}} = f_{\Phi_1}\restriction_{[0,1]} = f_{\Phi_1}$ (representing the probability formula Φ_1) and $f_{\Phi_2}\restriction_{\mathscr{C}_{\{q,\neg q\}}}$ (representing the probability formula Φ_2) are different; in particular, they have different domains. However, intuitively, the functions once restricted to the proper coherence polytopes have the same shape.

We can actually transform Φ_1 in an equivalent formula that shares the same basic modal formulas of Φ_2 as follows:

$$\Phi_1^*[P(q), P(\neg q)] = (P(q) \wedge \neg P(q)) \wedge (P(\neg q) \to P(\neg q)).$$

Notice that the coherence polytope of Φ_1^* is the same as Φ_2, i.e. $\mathscr{C}_{\{q,\neg q\}}$. Also notice that $P(\neg q) \to P(\neg q)$ takes value 1 for all possible assignments on q (this is because $x \to x$ is a Łukasiewicz tautology). This means that $f_{\Phi_1^*}$ is the cylindrification of f_{Φ_1} through the $\neg q$ axis (see Figure 3).

Notice that

$$f_{\Phi_2}\restriction_{\mathscr{C}_{\{q,\neg q\}}} = f_{\Phi_1^*}\restriction_{\mathscr{C}_{\{q,\neg q\}}}$$

witnessing the equivalence between Φ_1 and Φ_2 via the functional representation.

The following result presents what we have briefly shown in the above example from a more general perspective. Its proof is a straightforward application of the basic definitions and it is hence omitted. In what follows we will adopt the following notation: let $\Phi_1[P(\psi_1), \ldots, P(\psi_k)]$ and $\Phi_2[P(\gamma_1), \ldots, P(\gamma_m)]$ be probability formulas on events from $\mathcal{E}_1 = \{\psi_1, \ldots, \psi_k\}$ and $\mathcal{E}_2 = \{\gamma_1, \ldots, \gamma_m\}$ respectively. Let us call $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2$ and , for $i = 1, 2$,

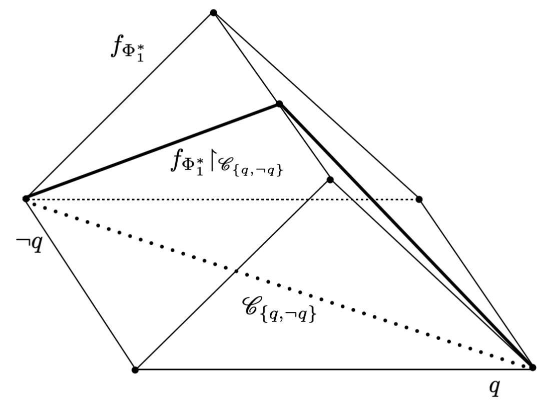

Figure 3: The coherence polytope $\mathscr{C}_{\{q,\neg q\}} = \{(\beta, 1-\beta) \mid \beta \in [0,1]\}$ (dotted line), the McNaughton function $f_{\Phi_1^*}$ obtained by cylindrification of f_{Φ_1} and its restriction to $\mathscr{C}_{\{q,\neg q\}}$ (boldfaced).

$$\Phi_i^* = \Phi_i \wedge \left(\bigwedge_{\delta \in \mathcal{E} \setminus \mathcal{E}_i} (P(\delta) \to P(\delta)) \right).$$

Thus, the coherence polytopes of Φ_1^* and Φ_2^* coincide with $\mathscr{C}_{\mathcal{E}}$. Hence, the following result is a direct consequence of what we recalled after Theorem 4.6 and Theorem 5.3.

Proposition 5.5. *Let Φ_1 and Φ_2 be two probabilistic formulas whose sets of events are $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively. Φ_1 and Φ_2 are logically equivalent in* $\mathrm{FP}(Ł)$ *iff*

$$f_{\Phi_1^*}\restriction_{\mathscr{C}_{\mathcal{E}_1 \cup \mathcal{E}_2}} = f_{\Phi_2^*}\restriction_{\mathscr{C}_{\mathcal{E}_1 \cup \mathcal{E}_2}}.$$

Let us end this section with some considerations on an issue we hinted at in the introduction; that is, whether propositional Łukasiewicz logic could be regarded itself as a probability logic. By Theorem 5.3 McNaughton functions, besides being the functional semantics for propositional sentences, also provide a functional interpretation of particular probability modal formulas. As a matter of fact, if $\mathcal{E}$ is a finite set of Łukasiewicz propositional variables $\{q_1, \ldots, q_k\}$, the coherence set $\mathscr{C}_{\mathcal{E}}$ coincides with the whole cube $[0,1]^k$. Therefore, if $f : [0,1]^k \to [0,1]$ is a McNaughton function and if $\Phi[q_1, \ldots, q_k]$ is a Łukasiewicz formula such that $f = f_\Phi$, the probability formula $\hat{\Phi} = \Phi[P(q_1), \ldots, P(q_k)]$ is also such that $f = f_{\hat{\Phi}}$. These observations are made precise in the next result whose proof is straightforward and therefore omitted.

Proposition 5.6. *Let $\{q_1, \ldots, q_k\}$ be a set of propositional variables, $\Phi(q_1, \ldots, q_k)$ be a Łukasiewicz formula, and $\hat{\Phi} = \Phi[P(q_1), \ldots, P(q_k)]$ be a probability formula on events $\mathcal{E} = \{q_1, \ldots, q_k\}$. Then, $f_\Phi = f_{\hat{\Phi}}$. As a consequence Φ is a theorem of Łukasiewicz logic iff for every map $\tau : \{P(q_1), \ldots, P(q_k)\} \to [0,1]$, $\tau(\hat{\Phi}) = 1$.*

Therefore, if from one side (full) McNaughton functions, and hence propositional formulas of Łukasiewicz logic, are capable of handling probability formulas whose events are propositional variables, things change when we consider modal formulas on less trivial *compound* events. To clarify this last claim by an example, consider the two formulas $\Phi_1 = P(q_1) \wedge P(q_2)$ and $\Phi_2 = P(q_1 \wedge q_2)$. While the former is representable by the McNaughton function $f_{\Phi_1}(q_1, q_2) = q_1 \wedge q_2$, the latter is associated to the 1-variable identity function $f_{\Phi_2} = id : [0,1] \to [0,1]$ on $\mathscr{C}_{\{q_1 \wedge q_2\}} = [0,1]$. In other words, the connection between the two formalisms is lost, as expected, when dealing with compound events on which, in contrast with Łukasiewicz logic, probability logic is not truth-functional. To sum up, we can therefore say that although propositional Łukasiewicz logic can express and reason about probabilistic statements on the basic events that can be formalized by propositional variables, it cannot faithfully handle formulas that interpret the probability of complex events.

6 A Global Representation of Probability Formulas via Probability Distributions

In the previous section we have shown a functional representation for the formulas of $\mathrm{FP}(\text{Ł})$ that is *local* in the sense that it depends on the classical formulas occurring as events in the modal formula we are dealing with. Thus, formulas on different sets of events are represented by McNaughton functions restricted on different coherence sets.

We now show that the functional representation can also be described in *global* terms; i.e., we will see how to reduce oneself to fix, once and for all, a finite dimensional polytope Δ so that all modal formulas of $\mathrm{FP}(\text{Ł})$ can be represented as McNaughton functions restricted on Δ.

The only assumption we make is to start with a finite set of propositional variables $V = \{q_1, \ldots, q_n\}$. Since any formula is written on a finite set of variables, this assumption comes without any loss of generality in potential applications. For this global representation we will then employ probability distributions on the set of atoms of the free Boolean algebra $\mathrm{Free}(n)$ generated by the finite set of variables V. More precisely, let $\alpha_1, \ldots, \alpha_{2^n}$ be the atoms of $\mathrm{Free}(n)$, and let us denote, as in Section 2, by $\Delta_{\mathrm{Free}(n)}$ the $2^n - 1$-simplex of probability distributions on $\alpha_1, \ldots, \alpha_{2^n}$. From Proposition 2.1, $\Delta_{\mathrm{Free}(n)}$ encodes all possible finitely additive probability functions on $\mathrm{Free}(n)$.

The atoms of $\mathrm{Free}(n)$ correspond, syntactically, to *minterms* of the classical language over n variables $q_1, \ldots, q_n$. These are formulas of the form

$$m_j = \bigwedge_{i=1}^{n} q_i^*$$

where q_i^* is a literal standing for either q_i or $\neg q_i$. Therefore, we will henceforth refer to minterms, rather than atoms, when dealing with formulas. In the same way every element of $\mathrm{Free}(n)$ can be written as a disjunction of atoms (as explained in Section 2), every formula ψ on n variables is logically equivalent to $\bigvee_{m_j \vdash \psi} m_j$, i.e., the disjunction of all the min terms that entail ψ.

Now, we need to introduce a normal form for formulas of $\mathrm{FP}(\text{Ł})$ that relies on minterms rather than on generic Boolean events.

Definition 6.1. A formula of $\mathrm{FP}(\text{Ł})$ is said to be in *atomic normal form* if all its basic modal subformulas are in the form $P(m_j)$ for some minterm m_j.

As we are going to show, every probability formula of $\mathrm{FP}(\text{Ł})$ is equivalent to a formula in atomic normal form. Thus, from the syntactical perspective this restriction comes with no loss of generality. To start with, let us pick a generic formula $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ and let us write each ψ_i as a disjunction of minterms as explained above:

$$\psi_i = \bigvee_{m \vdash \psi_i} m.$$

It is easy to check (but see (Flaminio and Ugolini 2023, Proposition 2.9)) that every basic probability formula $P(\psi_i)$ is logically equivalent to $P\left(\bigvee_{m \vdash \psi_i} m\right)$. Similarly, $\Phi[P(\psi_1), \ldots, P(\psi_k)]$ is logically equivalent to

$$\Phi\left[P\left(\bigvee_{m \vdash \psi_1} m\right), \ldots, P\left(\bigvee_{m \vdash \psi_k} m\right)\right].$$

Notice that in $\mathrm{FP}(\text{Ł})$, for each i, $P\left(\bigvee_{m \vdash \psi_i} m\right)$ is equivalent to $\bigoplus_{m \vdash \psi_i} P(m)$. Thus, finally, we denote by $\Phi^*[P(m_1), \ldots, P(m_{2^n})]$ the formula obtained by uniformly replacing, in Φ, $P(\psi_i)$ for $\bigoplus_{m \vdash \psi_i} P(m)$. The following holds.

Proposition 6.2. *For every probability formula Φ of $\mathrm{FP}(\text{Ł})$ there exists a probability formula Φ^* in atomic normal form such that $\vdash_{FP} \Phi \equiv \Phi^*$.*

In the same lines of the previous section, before proving the global representation theorem for probability formulas in atomic normal form, let us briefly show to which functions such formulas correspond. In the next result we assume that the formula $\Phi[P(m_1), \ldots, P(m_{2^n})]$ is in atomic normal form without loss of generality.

Proposition 6.3. *For every formula $\Phi[P(m_1), \ldots, P(m_{2^n})]$ in atomic normal form, the McNaughton function f_Φ is such that, for every $d_{\mathbf{x}} = (x_1, \ldots, x_{2^n}) \in \Delta_{\mathrm{Free}(n)}$,*

$$f_\Phi(x_1, \ldots, x_{2^n}) = d_{\mathbf{x}}(\Phi[P(m_1), \ldots, P(m_{2^n})]).$$

We now prove the converse, and hence the non trivial direction of the characterization theorem.

Theorem 6.4. *For every n and for every McNaughton function $f : [0,1]^{2^n} \to [0,1]$ there exists a probability formula in atomic normal form $\Phi[P(m_1), \ldots, P(m_{2^n})]$ such that, for every $d_{\mathbf{x}} = (x_1, \ldots, x_{2^n}) \in \Delta_{\mathrm{Free}(n)}$,*

$$f_\Phi(x_1, \ldots, x_{2^n}) = d_{\mathbf{x}}(\Phi[P(m_1), \ldots, P(m_{2^n})]).$$

Proof. The claim can be shown by applying the same proof of the above Theorem 5.3. Indeed, for all n, $\Delta_{\mathrm{Free}(n)}$ is the coherence polytope relative to the set of events $\mathscr{E} = m_1, \ldots, m_{2^n}$. Thus, let $\Phi[P(m_1), \ldots, P(m_{2^n})]$ be the probabilistic formulas given by Theorem 5.3. By Proposition 2.1 for every $d_{\mathbf{x}} \in \Delta_{\mathrm{Free}(n)}$, let $\mu_{\mathbf{x}}$ denote the unique finitely additive probability measure induced by $d_{\mathbf{x}}$. Therefore,

$$\begin{aligned} d_{\mathbf{x}}(\Phi[P(m_1),...,P(m_{2^n})]) &= \mu_{\mathbf{x}}(\Phi[P(m_1),...,P(m_{2^n})]) \\ &= f_\Phi(x_1,...,x_{2^n}). \end{aligned}$$

Thus, the claim is settled. □

7 Conclusion and Future Work

In this paper we have presented some results on the functional representation for the formulas of a modal logic meant to reason about probability. This logic, denoted by $\mathrm{FP}(\text{Ł})$, was introduced in (Hájek, Esteva, and Godo 1995) and it is based on Łukasiewicz propositional calculus. However, in contrast with the latter, $\mathrm{FP}(\text{Ł})$ is non truth-functional and is indeed a proper formalism to handle uncertainty values as it is sound and complete with respect to finitely additive probability functions.

In more details we have presented two functional representations for $\mathrm{FP}(\text{Ł})$: a local and a global one. These representations are obtained by restricting McNaughton functions, i.e., real-valued continuous and piecewise linear functions, to suitable domains: in the local representation McNaughton functions are restricted to what we called coherence sets; the global representation is instead obtained considering the more general domain of probability distributions on finite Boolean algebras.

Besides presenting these two main results, a secondary, yet no less important aim of this paper is to suggest a formal environment for the integration of symbolic logic with learning algorithms. Following the recent (Preto and Finger 2022) where the authors apply the functional representation of Łukasiewicz logic to describe properties of binary classification neural networks, we believe that the representations provided in this paper could help improve those algorithms that handle proper uncertainty, rather than vagueness or imprecision.

On these lines, our future work will consider two perspectives. A first one, more theoretical, concerns extensions of the functional representation for probability formulas on events that are more general than classical ones. In particular we will deal with Łukasiewicz events as done in (Flaminio and Godo 2007), and also conditional events that will be handled as in (Flaminio, Godo, and Hosni 2020). The second line of research that we intend to develop aims at finding the most appropriate domain of applicability of this probabilistic framework. As we pointed out in the introduction of the present paper, in fact, we believe that Łukasiewicz logic cannot be the right formalism to handle degrees of uncertainty. For this reason, following the lines of (Preto and Finger 2022), we plan to employ the proposed techniques in an actual (predictive) neural network and to compare the obtained result with those that are present in the literature.

Acknowledgments

The authors thank the anonymous referees for their comments, suggestions and criticism. All the authors acknowledge partial support by the MOSAIC project (H2020-MSCA-RISE-2020 Project 101007627). Flaminio also acknowledges support by the Spanish project PID2019-111544GB - C21/AEI/10.13039/501100011033. Preto was partly supported by the São Paulo Research Foundation (FAPESP) project #2021/03117-2. This work was carried out in part at the Center for Artificial Intelligence (C4AI-USP), with support by FAPESP, project #2019/07665-4, and by the IBM Corporation. Ugolini acknowledges support from the Marie Skłodowska-Curie grant agreement No 890616 (H2020-MSCA-IF-2019), and the Ramón y Cajal programme RyC2021-032670-I.

References

Aguzzoli, S.; Nola, A. D.; Gerla, B.; and Russo, C. 2021. Mv-tropical polynomials and neural networks. In et.al, A. C., ed., *Proceedings of WILF 2021, the 13th International Workshop on Fuzzy Logic and Applications (WILF 2021), Vietri sul Mare, Italy, December 20-22, 2021*, volume 3074 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Aguzzoli, S.; Bova, S.; and Gerla, B. 2011. *Free Algebras and Functional Representation for Fuzzy Logics*. University College. Chapter IX.

Badia, G.; Fagin, R.; and Noguera, C. 2023. New foundations of reasoning via real-valued first-order logics. Submitted. Preprint available at https://arxiv.org/abs/2207.00086.

Badreddine, S.; d'Avila Garces, A.; Serafini, L.; and Spranger, M. 2022. Logic tensor network. *Artificial Intelligence* 103649.

Baldi, P.; Cintula, P.; and Noguera, C. 2020. Classical and Fuzzy Two-Layered Modal Logics for Uncertainty: Translations and Proof-Theory. *Internationa Journal pof Computational Intelligence Systems* 13(1):988–1001.

Cignoli, R.; D'Ottaviano, I. M.; and Mundici, D. 1999. *Algebraic Foundations of Many-Valued Reasoning*. Kluwer Academic Publishers.

Cramer, T. 1970. Countable boolean algebras as subalgebras and homomorphs. *Pacific Journal of Mathematics* 35(3):321–326.

de Finetti, B. 1935. The logic of probability. *Philos. Stud. (1995)* 77:181–190.

Di Nola, A.; Lenzi, G.; and Vitale, G. 2016. Łukasiewicz Equivalent Neural Networks. In Bassis, S.; Esposito, A.; Morabito, F. C.; and Pasero, E., eds., *Advances in Neural Networks*, 161–168. Cham: Springer International Publishing.

Ewald, G. 1996. *Combinatorial Convexity and Algebraic Geometry*. Graduate Texts in Mathematics. Springer International Publishing.

Fagin, R.; Halpern, J.; and Megiddo, N. 1990. A Logic for Reasoning about Probabilities. *Information and Computation* 87(1-2):78–128.

Flaminio, T., and Godo, L. 2007. A logic for reasoning about the probability of fuzzy events. *Fuzzy Sets and Systems* 158:625–638.

Flaminio, T., and Ugolini, S. 2023. Enconding de Finetti's coherence in Łukasiewicz logic and MV-algebras. Submitted. Preprint available at https://arxiv.org/pdf/2303.06963.pdf.

Flaminio, T.; Godo, L.; and Hosni, H. 2014. On the logical structure of de Finetti's notion of event. *Journal of Applied Logic* 12(3):279–301.

Flaminio, T.; Godo, L.; and Hosni, H. 2020. Boolean algebras of conditionals, probability and logic. *Artificial Intelligence* 286:103347.

Giannini, F.; Diligenti, M.; Gori, M.; and Maggini, M. 2018. On a convex logic fragment for learning and reasoning. In *IEEE Transactions on Fuzzy Systems*.

Hájek, P.; Esteva, F.; and Godo, L. 1995. Fuzzy logic and probability. In Besnard, P., and (Eds.), S. H., eds., *Proc. of Uncertainty in Artificial Intelligence UAI'95*, 237–244.

Hájek, P. 1998. *Metamathematics of Fuzzy Logic*. Kluwer Academic Publishers.

Krein, M., and Milman, D. 1940. On extreme points of regular convex sets. *Studia Mathematica* 9:133–138.

McNaughton, R. 1951. A theorem about infinite valued sentential calculi. *The Journal of Symbolic Logic* 16:336–348.

Mundici, D. 1994. A constructive proof of McNaughton's theorem in infinite-valued logic. *The Journal of Symbolic Logic* 59(2):596–602.

Preto, S., and Finger, M. 2022. Proving properties of binary classification neural networks via Łukasiewicz logic. *Logic Journal of the IGPL*.

van Krieken, E.; Acar, E.; and van Harmelen, F. 2022. Analyzing Differentiable Fuzzy Logic Operators. *Artificial Intelligence* 302:103602.