Prolog Extensions to Many–Valued Logics

F. Klawonn
Department of Computer Science
Technical University of Braunschweig
38106 Braunschweig, Germany

Abstract

The aim of this paper is to show that a restriction of a logical language to clauses like Horn clauses, as they are used in Prolog, applied to [0,1]–valued logics leads to calculi with a sound and complete proof theory. In opposition to other models where generally the set of axioms as well as the deduction schemata are enriched we restrict ourselves to a simple modification of the deduction rules of classical logic without adding new axioms.

In our model the truth values from the unit interval can be interpreted in a probabilistic sense, so that a value between 0 and 1 is not just intuitively interpreted as a 'degree of truth'.

Keywords: Prolog; [0,1]–valued logic; probabilistic logic; possibilistic logic

1 Introduction

N. Rescher [23] pointed that there are at least three different approaches to the field of many–valued logic, namely

- the metalogical viewpoint, which is mainly concerned with proof theoretic and algebraic aspects of logical systems as for example described in [22],

- the semantical standpoint, from which N. Rescher's book is written, where the set of truth values is enriched with values like *undetermined* or more abstract values like '0.5',
- and the practical view, which concentrates on applications of many–valued systems for example in physics as indicated in [14].

In this paper we emphasize the semantical viewpoint and focus our attention to applications in the domain of approximate reasoning. We deliberately restrict our investigations to [0,1]–valued logics, so that we are enabled to provide probabilistic interpretations for our concepts.

In the field of approximate reasoning it is very common to attach a (truth) value to a (logical) formula, expressing for instance a degree of truth, possibility, necessity, plausibility, or belief. This weighting or valuation of formulae enforces an extension of the logical language in order to be able to express the truth value attached to a formula. Although the definition of the notion of a model or an interpretation (i.e. the semantical part) of such a language is straight forward, the rules for logical deduction have to be modified and new rules have to be added for the sake of completeness. Examples for such extensions can be found in [21, 16, 17, 18, 19]. The completeness results of these papers are obtained for the price of a complex deduction mechanism, that guarantees completeness, but does not provide efficient methods for finding proofs. Therefore, these approaches are very valuable from a theoretical point of view, but are subject to limitations for practical applications.

Another problem for some applications is a missing interpretation for the truth values between 0 and 1. It is often not enough to understand truth values in an intuitive sense as degrees of truth, possibility, necessity, plausibility, or belief. These notions without making more precise, what the meaning of a certain degree of 0.8 is, and when we should attach this degree to a formula instead of the degree 0.7, can cause undesired results in applications or may even lead to the rejection of such approaches according to the inherent arbitrariness in the choice of the numbers (degrees).

A simple way to overcome these problems is to interpret the unit interval as an ordinal scale as proposed by Dubois and Prade in [5] for their possibilistic logic. Disadvantages of such an approach are that values specified by different persons cannot be compared and that the richer structure of the unit interval is reduced to a linear ordering, although generally more than a simple ranking is associated with numbers between zero and one.

In this paper we present the following approach. As in many other models for approximate reasoning on a logical basis, we consider the unit interval as the set of 'truth values'. In Section 2 a purely formal approach is described without discussing an interpretation of the 'truth values'. In order to avoid a complicated proof theory we restrict our considerations to a subset of a first order logical language. We only admit formula that are similar to Horn clauses, so that we obtain a language suitable for 'fuzzy' Prolog including some completeness results.

Section 3 is devoted to possible interpretations of the truth values. We provide probabilistic models that can be used for an underlying semantics of the truth values. It turns out that the probabilistic interpretation can also be applied to possibilistic logic, so that the we obtain an equivalence between fuzzy Prolog based on the Gödel implication, possibilistic Prolog, and a probabilistic model.

2 Extending Prolog to [0,1]–valued Logics

We consider a first order logical language L containing the logical connectives $\rightarrow$, a set $\bigoplus = \{\oplus_1, \ldots, \oplus_n\}$ of binary connectives, and the universal quantifier $\forall$. The set of truth values is the unit interval [0,1]. The valuation function associated to the logical connective $\rightarrow$ is either

$$/\varphi \rightarrow \psi/ = \min\{1 - /\varphi/ + /\psi/, 1\} \qquad \text{(Łukasiewicz implication)}$$

or

$$/\varphi \rightarrow \psi/ = \begin{cases} 1 & \text{if } /\varphi/ \le /\psi/ \\ /\psi/ & \text{otherwise.} \end{cases} \qquad \text{(Gödel implication)}$$

If $\rightarrow$ is intended to be the Łukasiewicz implication we write $L_Ł$ for L, in case of the Gödel implication we write L_G. For the connectives in $\bigoplus$ we only assume that the corresponding valuation functions are continuous and non–decreasing in both arguments. Examples for such operators are continuous t–norms and t–conorms. For the universal quantifier we define $/(\forall x)(\varphi(x))/ = \inf_x\{/\varphi(x)/\}$.

The following definition describes a restricted subset of the logical language L, which generalizes the notion of Horn–clauses for our purposes.

Definition 2.1 *An implication clause is a closed well formed formula of the form*

$$(\forall x_1)\dots(\forall x_k)(\varphi \to \psi) \tag{1}$$

or

$$(\forall x_1)\dots(\forall x_k)(\psi), \tag{2}$$

where ψ *is an atomic formula with no other free variables than* $x_1, \dots, x_k$. φ *is a formula containing only connectives belonging to* $\bigoplus$ *and no quantifiers.*

A rule base in Prolog consists of rules and facts. Such a rule base is interpreted as a set of axioms known to be true. Instead of a crisp set of axioms (facts and rules) $A \subseteq L$ as in classical logic or Prolog, we consider a mapping $a : L \to [0,1]$ assigning to each logical formula $\varphi \in L$ a lower bound $a(\varphi)$ for the truth value of φ. For classical logic a would correspond to the characteristic function of the given axioms. In practical applications $a(\varphi)$ will in general only be specified for some $\varphi \in L$, whereas for all other $\psi \in L$ the default lower bound zero is assumed, i.e. $a(\psi) = 0$. Since we want to restrict our considerations to implication clauses, we allow a lower bound greater than zero only for implication clauses.

Definition 2.2 *A mapping* $a : L \to [0,1]$ *is called regular, if only implication clauses belong to the support of* a*, i.e.* $a(\varphi) > 0$ *implies that* φ *is an implication clause.*

Our intention is to keep the lower bounds specified by a out of the logical language. For this reason, we also use only the classical deduction schemata, i.e. modus ponens and substitution. Of course, we have to compute corresponding lower bounds for formulae involved in the deduction procedure. This means that, instead of adding a new valid formula in each deduction step as in classical logic, we improve the lower bound for a formula in a deduction step. Formally, a deduction step derives from $a : L \to [0,1]$ a mapping $b : L \to [0,1]$, where $b \geq a$. This motivates the following definition for the inference procedure.

Definition 2.3 *Let* $a, b : L_{\mathrm{L}} \to [0,1]$ *be regular.*

(i) b *is directly derivable from* a *if*

(a) there exists an implication clause $(\forall x_1)\dots(\forall x_k)(\varphi \to \psi)$ *in* $L_{\rm L}$ *such that*

(a1) If ψ_0 *is an implication clause with free variables* $x_1,\dots,x_r$ *and* $\psi_0 \neq \psi$*, then* $a((\forall x_1)\dots(\forall x_r)(\psi_0)) = b((\forall x_1)\dots(\forall x_r :)(\psi_0))$.

(a2)

$$\begin{aligned} b((\forall x_1)\dots(\forall x_k)(\psi)) \;=\; & \max\{/(\forall x_1)\dots(\forall x_k)(\varphi)/_a \\ & +a((\forall x_1)\dots(\forall x_k)(\varphi \to \psi)) \\ & -1, a((\forall x_1)\dots(\forall x_k)(\psi))\ \}, \end{aligned} \tag{3}$$

where the value $/(\forall x_1)\dots(\forall x_k)(\varphi)/_a$ *is obtained by considering the Herbrand universe of* $L_{\rm L}$ *and valuating atomic formulae according to* a*,*

holds or

(b) there exists an implication clause $(\forall x_1)\dots(\forall x_k)(\chi)$ *and terms* $t_{i_1},\dots,t_{i_r}$ $(i_1,\dots i_r \in \{1,\dots,k\})$ *without free variables such that for the formula* χ'*, which is obtained by substituting* x_{i_j} $(j = 1,\dots,r)$ *by* t_{i_j} *in* χ *and quantifying over the remaining free variables,*

$$b(\chi') = \max\{a((\forall x_1)\dots(\forall x_k)(\chi)), a(\chi')\}. \tag{4}$$

is satisfied.

(ii) b *is derivable from* a $(a \lhd b)$ *if there is a regular sequence* $a_0,\dots,a_n : L_{\rm L} \to [0,1]$ *where* a_{k+1} *is directly derivable from* a_k *for each* $k \in \{0,\dots,n-1\}$ *and* $b \le a_n$ *holds.*

(iii) The mapping $\mathrm{th}^{(a)} : L_{\rm L} \to [0,1]$ *is given by*

$$\mathrm{th}^{(a)}(\varphi) = \begin{cases} \sup\{b(\varphi) \mid a \lhd b\} & \text{if } \varphi \text{ is an implication clause} \\ & \text{of the form (2)} \\ 0 & \text{otherwise.} \end{cases}$$

We use the same terminology as in definition 2.3 for L_G. But we have to replace (3) by

$$\begin{aligned} b((\forall x_1)\dots(\forall x_k)(\psi)) \;=\; \max\Big\{ & \min\{/(\forall x_1)\dots(\forall x_k)(\varphi)/_a, \\ & a((\forall x_1)\dots(\forall x_k)(\varphi \to \psi))\}, \\ & a((\forall x_1)\dots(\forall x_k)(\psi))\ \Big\}. \end{aligned} \tag{5}$$

The application of modus ponens and the substitution of free variables by other terms is formalized in (i)(a) and (i)(b), respectively, in the above definition. (ii) describes the application of a finite number of deduction steps of the form introduced in (i). Finally, (iii) specifies what can be obtained from a if we allow an arbitrary number of deduction steps.

For the semantical or model–theoretic part of our $[0,1]$–valued logic we define accordingly to the interpretation of a as a specification of lower bounds for the truth values the mapping $\mathrm{Th}^{(a)}$ that describes the consequences for the lower bounds for all formulae induced by a.

Definition 2.4 *Let $a : L \to [0,1]$. Let $/\varphi/_I$ denote the truth value the formula φ obtains under the [0,1]–valued interpretation I. I is called compatible with a if $/\varphi/_I \geq a(\varphi)$ holds for all $\varphi \in L$.*

$\mathrm{Th}^{(a)} : L \to [0,1]$ *denotes the infimum over all [0,1]–valued interpretations of L, that are compatible with a.*

We write $/\varphi/$ instead of $/\varphi/_I$ if it is clear to which interpretation I we refer.

Theorem 2.5 (Soundness of L_{L}) *Let $a : L_{\mathrm{L}} \to [0,1]$ be regular. For all closed formulae $\varphi \in L_{\mathrm{L}}$*

$$\mathrm{th}^{(a)}(\varphi) \leq \mathrm{Th}^{(a)}(\varphi)$$

holds.

Proof. We have to prove that direct derivability preserves compatibility. Let I be an interpretation compatible with a and let b be directly derivable from a.
Case 1. b is obtained from a by applying (3).
From the compatibility of the interpretation I with a and the monotonicity of the valuation functions in $\bigoplus$, we derive

$$\begin{array}{rcll} /(\forall x_1)\ldots(\forall x_k)(\psi)/ & \geq & \max\{ & /(\forall x_1)\ldots(\forall x_k)(\varphi)/ \\ & & & +/(\forall x_1)\ldots(\forall x_k)(\varphi \to \psi)/_a - 1, \\ & & & /(\forall x_1)\ldots(\forall x_k)(\psi)/ \ \} \\ & \geq & \max\{ & /(\forall x_1)\ldots(\forall x_k)(\varphi)/_a \\ & & & +/(\forall x_1)\ldots(\forall x_k)(\varphi \to \psi)/_a - 1, \\ & & & /(\forall x_1)\ldots(\forall x_k)(\psi)/_a \ \} \\ & = & b((\forall x_1)\ldots(\forall x_k)(\psi)). & \end{array}$$

Case 2. b is obtained from a by applying (4).
For the same reasons as in case 1, we obtain

$$\begin{aligned} /\chi'/ &= \max\{/(\forall x_1)\dots(\forall x_k)(\chi)/, /\chi'/\} \\ &\geq \max\{a((\forall x_1)\dots(\forall x_k)(\chi)), a(\chi')\}. \end{aligned}$$

□

Theorem 2.6 (Soundness of L_G) *Let $a : L_G \to [0,1]$ be regular. For all closed formulae $\varphi \in L_{\mathrm{L}}$*

$$\mathrm{th}^{(a)}(\varphi) \leq \mathrm{Th}^{(a)}(\varphi)$$

holds.

Proof. The proof is analogous to the proof of Theorem 2.5, except that we have to consider equation (5) instead of (3) in case 1. □

Theorem 2.7 (Completeness of L_{L}) *Let $a : L_{\mathrm{L}} \to [0,1]$ be regular and let ψ be an implication clause of the form* (2). *Then*

$$\mathrm{th}^{(a)}(\psi) \geq \mathrm{Th}^{(a)}(\psi).$$

holds.

Proof. Let U be the Herbrand universe of L_{L}. We show that the Herbrand interpretation I induced by $\mathrm{th}^{(a)}$ is compatible with a. For implication clauses χ of the form (2) the definition of $\mathrm{th}^{(a)}$ yields

$$/\chi/ \geq a(\chi).$$

Thus, we only have to consider implication clauses like $(\forall x_1)\dots(\forall x_k)(\varphi \to \psi)$ of the form (1) where

$$/(\forall x_1)\dots(\forall x_k)(\varphi \to \psi)/ < a((\forall x_1)\dots(\forall x_k)(\varphi \to \psi)).$$

There exists a tuple $u = (u_1, \dots, u_k) \in U^k$ such that

$$/\varphi(u) \to \psi(u)/ < a((\forall x_1)\dots(\forall x_k)(\varphi \to \psi)) \qquad (6)$$

holds where $\varphi(u)$ and $\psi(u)$ are obtained by substitution of $x_1, \ldots, x_k$ by $u_1, \ldots, u_k$ in φ and ψ, respectively. By applying part (i)(b) of Definition 2.3 to the formula $(\forall x_1) \ldots (\forall x_k)(\varphi \to \psi)$ by substituting $x_1, \ldots, x_k$ by $u_1, \ldots, u_k$ we derive $b : L_{\mathrm{L}} \to [0,1]$ directly from a. Then we have

$$\begin{aligned} 1 &\geq b(\varphi(u) \to \psi(u)) \\ &\geq a((\forall x_1) \ldots (\forall x_k)(\varphi \to \psi)) \\ &> /\varphi(u) \to \psi(u)/ \\ &= \min\{1 - /\varphi(u)/ + /\psi(u)/, 1\} \\ &= 1 - /\varphi(u)/ + /\psi(u)/. \end{aligned} \tag{7}$$

(7) implies

$$/\varphi(u)/ + b(\varphi(u) \to \psi(u)) - 1 > /\psi(u)/.$$

According to the continuity of the valuation functions for the connectives in $\bigoplus$ and the definition of $\mathrm{th}^{(a)}$ and $/./$ there exists $b' : L_{\mathrm{L}} \to [0,1]$ directly derivable from a such that

$$/\varphi(u)/_{b'} + b(\varphi(u) \to \psi(u)) - 1 > /\psi(u)/ \tag{8}$$

holds. Let $\tilde{b} = \max\{b, b'\}$. Obviously, $\tilde{b}$ is also derivable from a. Taking (8) into account we get

$$/\varphi(u)/_{\tilde{b}} + \tilde{b}(\varphi(u) \to \psi(u)) - 1 > /\psi(u)/.$$

Therefore, there exists $\hat{b} : L_{\mathrm{L}} \to [0,1]$ directly derivable from $\tilde{b}$ by Definition 2.3(i)(a) such that

$$\begin{aligned} \hat{b}(\psi(u)) &= /\varphi(u)/_{\tilde{b}} + \tilde{b}(\varphi(u) \to \psi(u)) - 1 \\ &> /\psi(u)/. \end{aligned}$$

This leads to the contradiction

$$/\psi(u)/ = \mathrm{th}^{(a)}(\psi(u)) \geq \hat{b}(\psi(u)) > /\psi(u)/.$$

□

Theorem 2.8 (Completeness of L_G) *Let $a : L \to [0,1]$ be regular and let ψ be an implication clause of the form* (2)*. Then*

$$\mathrm{th}^{(a)}(\psi) \geq \mathrm{Th}^{(a)}(\psi).$$

holds.

Proof. The proof is analogous to that of Theorem 2.7 except for modifications induced by the differing valuation functions for the Łukasiewicz– and the Gödel implication. Note, that in the proof of Theorem 2.7 we only needed the continuity of the valuation functions for the connectives in $\bigoplus$, but not for the Łukasiewicz implication. Therefore, the discontinuity of the Gödel implication does not lead to any problems. □

Theorems 2.5 – 2.8 show that what we can derive by the deduction steps defined in Definition 2.3 coincides with what is deducible from a in the model–theoretic sense of Definition 2.4. This result holds for $L_Ł$ as well as for L_G. Note, that we allow an infinite number of deduction steps according to the supremum in Definition 2.3. Therefore, it is possible that the value $\mathrm{Th}^{(a)}(\varphi)$ can only be approximated (with arbitrary exactness) when we only allow a finite number of deduction steps.

The possibility of an infinite number of deduction steps is also considered in [16, 21]. But a number of additional axiom schemata and inference rules is needed for the completeness results in these papers.

3 A Probabilistic Interpretation for Prolog Extensions

The previous section was devoted to a purely formal approach to $[0,1]$–valued Prolog without giving an interpretation of the truth values. In the following we provide a formal framework in which the truth values originate from probabilities.

The probabilistic setting for our investigations is related to probabilistic logic [15], but does generalize the assumption of probabilistic logic that the probability for a formula plus the probability for its negation sum up to one to the weaker requirement that the sum is at most one. This corresponds to the idea that there are some 'worlds' in which the formula φ is known to be true, some different 'worlds' in which the negation of φ holds, and other 'worlds' in which nothing is known about φ.

Instead of the term '(possible) worlds' we will use the notion of a (consideration) context in the following, since our probabilistic model is motivated by the context model [9, 6], which was introduced as an integrating model for vagueness and uncertainty and later on also adopted for logical approaches [8].

Definition 3.1 *Let L be the set of (closed) well formed formulae (wff's) of a first order predicate language L and let $\mathcal{C} = (C, \mathcal{A}, P)$ be a probability space with σ-algebra $\mathcal{A}$ together with a mapping $\mu : C \to 2^L$ s.t.*

(i) for all $c \in C : \mathrm{TH}(\mu(c)) = \{\varphi \in L \mid \mu(c) \vdash_L \varphi\} = \mu(c)$

(ii) for all $c \in C : \perp \notin \mu(c)$, *(where* $\perp \leftrightarrow \varphi \wedge \neg\varphi$*)*

(iii) for all $\varphi \in L : \{c \in C \mid \varphi \in \mu(c)\} \in \mathcal{A}$.

Then $(\mathcal{C}, \mu)$ is called a context evaluation of L.

$\mathrm{TH}(\mu(c))$ denotes the set of (classical) logical consequences of the set $\mu(c)$.

C can be understood as a set of contexts or possible worlds. $\mu(c)$ represents the set of formulae that are known in context $c \in C$. It is assumed that all possible deductions are carried out in c (condition (i)), that $\mu(c)$ is consistent (condition (ii)), and that we can assign a number $P_\mu(\varphi)$ (i.e. a probability) to each formula $\varphi \in L$ due to the measurability condition (iii) via

$$P_\mu(\varphi) = P(\{c \in C \mid \varphi \in \mu(c)\}).$$

$P_\mu(\varphi)$ is the probability for those contexts in which φ is known to be true.

In the same way as we have defined compatibility of (logical) interpretations with a mapping $a : L \to [0, 1]$, that specifies lower bounds for the truth values (compare Definition 2.4), we introduce the notion of compatibility for context evaluations.

Definition 3.2 *Let* $a : L \to [0, 1]$.

(i) A context evaluation $(\mathcal{C}, \mu)$ *of* L*, where* $\mathcal{C} = (C, \mathcal{A}, P)$ *is a probability space, is compatible with* a *if*

$$\text{for all } \varphi \in L : a(\varphi) \leq P_\mu(\varphi)$$

(ii) The mapping $\text{Th}_a : L \to [0, 1]$ *is given by*

$$\text{Th}_a(\varphi) = \inf\Big\{P_\mu(\varphi) \mid \quad ((C, \mathcal{A}, P), \mu) \text{ is a context evaluation compatible with } a\Big\},$$

where $\inf \emptyset = 1$*.*

Note, that Th_a is not truth-functional.

3.1 A Probabilistic Interpretation for Prolog Based on the Łukasiewicz Implication

The following theorem shows that the logic $L_{\text{Ł}}$ can be understood as a cautious interpretation compared to the concept of context evaluations. In other words, if we use the inference procedure for $L_{\text{Ł}}$, which was introduced in Definition 2.3, we obtain a sound but not complete proof theory for our probabilistic interpretation.

Theorem 3.3 *Let* $a : L_{\text{Ł}} \to [0, 1]$ *be regular. Let* $\bigoplus = \{\wedge, \vee\}$ *where we associate the valuation functions*

$$\begin{aligned} /\varphi \wedge \psi/ &= \max\{/\varphi/ + /\psi/ - 1, 0\}, \quad \text{and} \\ /\varphi \vee \psi/ &= \max\{/\varphi/, /\psi/\} \end{aligned}$$

with $\wedge$ *and* $\vee$*, respectively. If* φ *is an atomic formula with free variables* $x_1, \ldots, x_n$*, then*

$$\text{th}^{(a)}((\forall x_1) \ldots (\forall x_n)(\varphi)) \;\leq\; \text{Th}_a((\forall x_1) \ldots (\forall x_n)(\varphi))$$

holds.

Proof. By induction we prove that for any b such that $a \lhd b$, and for all atomic formulae with free variables $y_1, \ldots, y_k$

$$b((\forall y_1) \ldots (\forall y_k)(\psi)) \;\leq\; \mathrm{Th}_a((\forall y_1) \ldots (\forall y_k)(\psi)) \tag{9}$$

holds.

By definition we have $a \leq \mathrm{Th}_a$, which gives us the basis for the induction. Now we have to consider a' derivable from a in n steps and b directly derivable from a' and a context evaluation $\big((C, \mathcal{A}, P), \mu\big)$ compatible with a. By the hypothesis of the induction we obtain that $\big((C, \mathcal{A}, P), \mu\big)$ is also compatible with a'. For the induction step we have to prove that $\big((C, \mathcal{A}, P), \mu\big)$ is compatible with b. There are two possibilities of deriving b directly from a'.

Case 1: b is obtained from a' by substitution of quantified variables by terms containing no free variables, for instance by replacing y_i by the term t.

$$b((\forall y_1) \ldots (\forall y_{i-1})(\forall y_{i+1}) \ldots (\forall y_k)(\psi(y_1, \ldots, y_{i-1}, t, y_{i+1}, \ldots, y_k)))$$
$$= a'((\forall y_1) \ldots (\forall y_k)(\psi(y_1, \ldots, y_k)))$$

For all other formulae b coincides with a'. Obviously,

$$\begin{aligned}
& a'((\forall y_1) \ldots (\forall y_k)(\psi(y_1, \ldots, y_k))) \\
& \quad \leq \quad P_\mu\big(\{c \in C \mid ((\forall y_1) \ldots (\forall y_k)(\psi(y_1, \ldots, y_k))) \in \mu(c)\}\big) \\
& \quad \leq \quad P_\mu\big(\{ \quad c \in C \mid \\
& \qquad\qquad ((\forall y_1) \ldots (\forall y_{i-1})(\forall y_{i+1}) \ldots (\forall y_k) \\
& \qquad\qquad (\psi(y_1, \ldots, y_{i-1}, t, y_{i+1}, \ldots, y_n))) \in \mu(c)\}\big)
\end{aligned}$$

holds.

Case 2: b is obtained from a' by the application of a deduction rule of the form

$$(\forall y_1) \ldots (\forall y_k)(\chi \rightarrow \psi)),$$

where χ and ψ contain at most $y_1, \ldots, y_n$ as free variables. Furthermore, ψ is an atomic formula, whereas χ can be composed of atomic formulae and the connectives $\wedge$ and $\vee$. In the following, we use $(\forall y)$ as an abbreviation for $(\forall y_1) \ldots (\forall y_k)$.

We carry out an induction on the number of connectives in χ. For the basis of the induction χ is an atomic formula. Since $(\forall y)(\chi)$ together with $(\forall y : (\chi \to \psi))$ implies $(\forall y)(\psi)$, we derive

$$\begin{aligned} P_\mu((\forall y)(\psi)) &\geq P_\mu\Big(\{c \in C \mid ((\forall y)(\chi)) \in \mu(c) \text{ and } ((\forall y)(\chi \to \psi)) \in \mu(c)\}\Big) \\ &\geq P_\mu((\forall y)(\chi)) + P_\mu((\forall y)(\chi \to \psi)) - 1 \\ &\geq a'((\forall y)(\chi)) + a'((\forall y)(\chi \to \psi)) - 1 \qquad (10) \\ &= b((\forall y)(\psi)). \end{aligned}$$

Now let $\chi = \chi_1 \vee \chi_2$ or $\chi = \chi_1 \wedge \chi_2$. This implies

$$P_\mu(\chi) \geq \max\{P_\mu(\chi_1), P_\mu(\chi_2)\} \quad \text{or}$$

$$P_\mu(\chi) \geq P_\mu(\chi_1) + P_\mu(\chi_2) - 1, \qquad (11)$$

respectively. In the same way as in (10) we obtain

$$P_\mu((\forall y)(\psi)) \geq b((\forall y)(\psi)).$$

□

The following example shows that equality in Theorem 3.3 is not satisfied in general, i.e. that the proof theory is indeed incomplete.

Example 3.4 Let L_{L} be the propositional calculus induced by the propositional constants φ_0, χ_0, and ψ_0. $a : L_{\mathrm{L}} \to [0,1]$ is given by

$$a(\varphi) = \begin{cases} 0.5 & \text{if } \varphi = \varphi_0 \\ 1 & \text{if } \varphi = (\varphi_0 \to \chi_0) \text{ or } \varphi = (\varphi_0 \wedge \chi_0 \to \psi_0) \\ 0 & \text{otherwise.} \end{cases}$$

This implies $\mathrm{th}^{(a)}(\psi_0) = 0$, since the interpretation with $/\varphi_0/ = /\chi_0/ = 0.5$ and $\psi_0 = 0$ is compatible with a. For a context evaluation $\big((C, \mathcal{A}, P), \mu\big)$ compatible with a we have

$$\begin{aligned} P_\mu(\psi_0) &\geq P\Big(\{c \in C \mid \varphi_0, (\varphi_0 \to \chi_0), (\varphi_0 \wedge \chi_0 \to \psi_0) \in \mu(c)\}\Big) \\ &\geq P\Big(\{c \in C \mid \varphi_0 \in \mu(c)\}\Big) \\ &\geq a(\varphi_0) \\ &= 0.5, \end{aligned}$$

since

$$P\Big(\{c \in C \mid (\varphi_0 \to \chi_0) \in \mu(c)\}\Big) = P\Big(\{c \in C \mid (\varphi_0 \wedge \chi_0 \to \psi_0) \in \mu(c)\}\Big) = 1.$$

Therefore, $\mathrm{Th}_a(\psi_0) \geq 0.5$ holds.

3.2 A Probabilistic Interpretation for Possibilistic Logic and Gödel Prolog

The incompleteness result of the previous subsection can be amended if we restrict the set of contexts. We now allow only context evaluations that are nested. This means that the contexts can be understood as a linearly ordered set of more and more speculative contexts, where the set of true formulae in a context becomes larger with the speculative level of the context.

Definition 3.5 *A context evaluation $\big((C, \mathcal{A}, P), \mu\big)$ of a first order language L is nested if there exists a subset $C_0 \subseteq C$ s.t.*

(i) $P(C_0) = 1$

(ii) for all $c, d \in C_0$: $\big(\mu(c) \subseteq \mu(d)$ *or* $\mu(d) \subseteq \mu(c)\big)$.

If we consider nested context evaluations, we have to modify the notion of Th_a, i.e. which minimal restrictions are induced by a specification $a : L \to [0,1]$ of lower bounds for the probabilities for formulae.

Definition 3.6 *Let $a : L \to [0,1]$. The mapping $\mathrm{Th}_a^{(\mathsf{poss})} : L \to [0,1]$ is given by*

$$\mathrm{Th}_a^{(\mathsf{poss})}(\varphi) = \inf\Big\{P_\mu(\varphi) \mid \; ((C, \mathcal{A}, P), \mu) \textit{ is a nested context evaluation compatible with } a\Big\}.$$

Before we proof that the proof theory for L_G is sound and complete with respect to nested context evaluations, we make a short excursus to possibilistic logic. The aim of this excursus is to show that nested context evaluations provide an appropriate interpretation for possibilistic logic so that our final result is the equivalence between possibilistic, nested context, and the $[0,1]$–valued Gödel Prolog. In order to simplify and clarify the necessary terminology for possibilistic logic, we give here slightly modified definitions compared to the originals from Dubois, Lang, and Prade [3, 1, 2, 5].

Definition 3.7 *Let L be the language of a first order predicate logic. A possibility measure Π on L is a mapping $\Pi : L \to [0,1]$ with the following properties.*

(i) $\Pi(\bot) = 0$

(ii) $\Pi(\top) = 1$

(iii) For all $\varphi, \psi \in L : \Pi(\varphi \vee \psi) = \max\{\Pi(\varphi), \Pi(\psi)\}$

(iv) $(\exists x : \varphi(x)) \in L \quad \Rightarrow \quad \Pi(\exists x : \varphi(x)) = \sup\{\Pi(\varphi(d)) \mid d \in D\}$ *where D is the Herbrand universe of L.*

(v) For all $\varphi, \psi \in L$:
$\Big($*If* $(\varphi \leftrightarrow \psi)$ *is a tautology then,* $\Pi(\varphi) = \Pi(\psi)$ *holds.*$\Big)$

$\Pi(\varphi)$ is interpreted as the degree to which φ is considered to be possible. The corresponding necessity measure N is given by

$$N : L \to [0,1], \qquad \varphi \mapsto 1 - \Pi(\neg\varphi)$$

where $N(\varphi)$ is understood as the degree to which φ is necessarily true.

N also satisfies conditions (i), (ii), and (v) of the above definition. The axioms (iii) and (iv) have to be replaced by the dual axioms, i.e. $\vee$, $\exists$, max, and sup should be substituted by $\wedge$, $\forall$, min, and inf, respectively. If N is a necessity measure, the corresponding possibility measure is obtained by $\Pi(\varphi) = 1 - N(\neg\varphi)$.

Definition 3.8 *Let $a : L \to [0,1]$ be regular. The mapping*

$$\mathrm{Th}_a^{(\mathrm{poss})} : L \to [0,1]$$

is defined by

$$\mathrm{Th}_a^{(\mathrm{poss})}(\varphi) = \inf\Big\{P_\mu(\varphi) \mid \begin{array}{l} \mathcal{C} = (C, \mathcal{A}, P) \textit{ and} \\ (\mathcal{C}, \mu) \textit{ is a context evaluation} \\ \textit{compatible with } a \end{array}\Big\}.$$

The following three theorems elucidate the connection between possibilistic logic and nested context evaluations.

Theorem 3.9 *Let N be a necessity measure on L. Then there exists a nested context evaluation $\big((C, \mathcal{A}, P), \mu\big)$ such that*

$$N = P_\mu$$

holds.

Proof. Define

$$\mu :]0,1] \to 2^L, \quad \alpha \mapsto \{\varphi \in L \mid N(\varphi) \geq \alpha\}.$$

If $N(\varphi) \geq \alpha$ and $N(\varphi \to \psi) = N(\neg\varphi \vee \psi) \geq \alpha$ hold, then

$$\begin{aligned} \alpha &\leq \min\{N(\varphi), N(\neg\varphi \vee \psi)\} = N(\varphi \wedge (\neg\varphi \vee \psi)) \\ &= N(\varphi \wedge \psi) = \min\{N(\varphi), N(\psi)\} \leq N(\psi). \end{aligned} \tag{12}$$

According to the conditions (i), (ii), (iv), and (v) for necessity measures we derive by exploiting (12) that $\mathrm{TH}(\mu(\alpha)) = \mu(\alpha)$ holds and that $\mu(\alpha)$ is consistent.
Let $C =]0,1]$ and let P the probability measure which corresponds to the uniform distribution on $]0,1]$. Let $\mathcal{C} = \big(]0,1], \mathcal{B}(]0,1]), P\big)$ where $\mathcal{B}(]0,1])$ is the Borel σ–algebra on $]0,1]$. For $\varphi \in L$ we have

$$\{c \in]0,1] \mid \varphi \in \mu(c)\} \;=\;]0, N(\varphi)].$$

Therefore, $(\mathcal{C}, \mu)$ is a context evaluation of L. Since $P\big(]0,1]\big) = 1$ and by the definition of μ, we obtain that $(\mathcal{C}, \mu)$ is nested.
Let $\varphi \in L$.

$$P_\mu(\varphi) = P\Big(\{c \in]0,1] \mid \varphi \in \mu(c)\}\Big) = P\Big(]0, N(\varphi)]\Big) = N(\varphi)$$

□

Theorem 3.10 *Let $\big((C, \mathcal{A}, P), \mu\big)$ be a nested context evaluation of L. Then P_μ is a necessity measure.*

Proof. Let $C_0 \subseteq C$ such that $P(C_0) = 1$ is satisfied and for all $c, d \in C_0$: $\big(\mu(c) \subseteq \mu(d)$ or $\mu(d) \subseteq \mu(c)\big)$ holds. Since $\mu(c)$ is consistent for all contexts $c \in C$ and because $\mu(c)$ contains at least all tautologies of L, we obtain $P_\mu(\bot) = 0$ and $P_\mu(\top) = 1$. Condition (v) for necessity measures is fulfilled according to $\mu(c) = \mathrm{TH}(\mu(c))$ (for all contexts $c \in C$).
For $\chi \in L$ we define $C_\chi = \{c \in C_0 \mid \chi \in \mu(c)\}$. Thus $P_\mu(\chi) = P(C_\chi)$ holds because of $P(C_0) = 1$.
Let $\varphi, \psi \in L$. In case of $C_\varphi \not\subseteq C_\psi$, there exists a context $c \in C_0$ with $\varphi \in \mu(c)$ and $\psi \notin \mu(c)$. For any context $d \in C_\psi$ we have $\mu(d) \not\subseteq \mu(c)$. Since $\big((C, \mathcal{A}, P), \mu\big)$ is nested, $\mu(c) \subseteq \mu(d)$ follows and therefore also $\varphi \in \mu(d)$. Thus, for the case $C_\varphi \not\subseteq C_\psi$ the inclusion $C_\psi \subseteq C_\varphi$ holds. Analogously, we obtain $C_\varphi \subseteq C_\psi$ for the case $C_\psi \not\subseteq C_\varphi$. In any case we have $C_\varphi \subseteq C_\psi$ or $C_\psi \subseteq C_\varphi$.
Without loss of generality let $C_\varphi \subseteq C_\psi$.

$$\begin{aligned} P_\mu(\varphi \wedge \psi) &= P\Big(\{c \in C_0 \mid (\varphi \wedge \psi) \in \mu(c)\}\Big) \\ &= P\Big(\{c \in C_0 \mid \varphi \in \mu(c) \text{ and } \psi \in \mu(c)\}\Big) \\ &= P(C_\varphi \cap C_\psi) \\ &= P(C_\varphi) \\ &= \min\{P(C_\varphi), P(C_\psi)\} \end{aligned}$$

Now we consider $(\forall x)(\varphi(x)) \in L$. Let U be the Herbrand universe of L. U is a countable set. Let $U = \{u_n \mid n \in \mathbb{N}\}$. We define

$$C^{(n)} = \bigcap_{i=0}^{n} C_{\varphi(u_i)}.$$

$C^{(n)}$ is a non–increasing sequence of sets and

$$\bigcap_{i=0}^{\infty} C^{(i)} = C_{(\forall x)(\varphi(x))}$$

holds. This implies

$$P\Big(C_{(\forall x)(\varphi(x))}\Big) \;=\; P\Big(\lim_{n\to\infty} \bigcap_{i=0}^{n} C^{(i)}\Big) = \lim_{n\to\infty} P(C^{(n)})$$

$$
\begin{aligned}
&= \lim_{n\to\infty} P\Big(C_{\varphi(u_0)\wedge\ldots\wedge\varphi(u_n)}\Big) \\
&= \lim_{n\to\infty} \min\{P_\mu(\varphi(u_0)),\ldots,P_\mu(\varphi(u_n))\} \\
&= \inf\{P_\mu(\varphi(u_n)) \mid n \in \mathbb{N}\}.
\end{aligned}
$$

This proves that P_μ is a necessity measure. □

Theorem 3.11 *Let $a : L \to [0,1]$ be regular. If there is at least one nested context evaluation compatible with a, then* $\mathrm{Th}_a^{(\mathsf{poss})}$ *is a necessity measure on L.*

Proof. According to Theorem 3.10 $\mathrm{Th}_a^{(\mathsf{poss})}$ is the infimum of a set of necessity measures, which is obviously also a necessity measure. □

In [4, 5] an interpretation of possibility theory and possibilistic logic on the basis of consonant or nested worlds was already suggested and the parallels to Spohn's generalized possible world model [25] were indicated by a purely qualitative interpretation, i.e. the unit interval was only considered as a linear ordering. Refraining from the rich structure of the unit interval by concentrating on the linear ordering imposes two drawbacks for possibility theory. On the one hand, the connection between possibility and necessity measures is based on subtraction which is not a property inherent in linear orderings. On the other hand, generally we associate a quantitative and not only a qualitative ranking with numbers from the unit interval. The interpretation of possibilistic logic in the light of nested context evaluations provides a meaningful, quantitative interpretation of the numbers.

Now, after we have clarified the connection between possibilistic logic and nested context evaluations, we can proof that nested context evaluations also provide a model or interpretation for 'Gödel Prolog', i.e. for L_G.

Theorem 3.12 *Let $a : L_G \to [0,1]$ be regular. Let $\bigoplus = \{\wedge, \vee\}$ where we associate the valuation functions* min *and* max *to $\wedge$ and $\vee$, respectively. If φ is an atomic formula with free variables $x_1,\ldots,x_n$, then*

$$
\mathrm{th}^{(a)}((\forall x_1)\ldots(\forall x_n)(\varphi)) \;=\; \mathrm{Th}_a^{(\mathsf{poss})}((\forall x_1)\ldots(\forall x_n)(\varphi))
$$

holds.

Proof. We again abbreviate $(\forall y_1)\ldots\forall(y_k)$ by $(\forall y)$.
As the first step of the proof we show by induction over the number of (direct) derivation steps that

$$b((\forall y)(\psi(y))) \leq \mathrm{Th}_a^{(\mathrm{poss})}((\forall y)(\psi(y))) \tag{13}$$

holds for all atomic formulae ψ with free variables $y_1,\ldots,y_k$ and for all b with $a \lhd b$.
The basis of the induction is given by $a((\forall y)(\psi(y))) \leq \mathrm{Th}_a^{(\mathrm{poss})}((\forall y)(\psi(y)))$.
Now let a' be derivable from a in n steps. By the induction hypothesis a' satisfies (13). We have to consider a nested context evaluation $\big((C,\mathcal{A},P),\mu\big)$ compatible with a (and according to $a \lhd a'$ also compatible with a') and $b: L \to [0,1]$ directly derivable from a'.

There are two possibilities to obtain b from a'.
Case 1 can be treated analogously to case 1 in the proof of Theorem 3.3.
For the second case we have to replace the inequality (10) by

$$\begin{aligned} P_\mu((\forall y)(\psi)) &\geq P_\mu\Big((\forall y)(\chi)) \wedge ((\forall y)(\chi \to \psi))\Big) \\ &= \min\Big\{P_\mu((\forall y)(\chi)), P_\mu((\forall y)(\chi \to \psi))\Big\} \\ &\geq \min\Big\{a'((\forall y)(\chi)), a'((\forall y)(\chi \to \psi))\Big\} \end{aligned} \tag{14}$$

in case 2 of the proof of Theorem 3.3. (14) is satisfied, since $\big((C,\mathcal{A},P),\mu\big)$ is a nested context evaluation and therefore, by Theorem 3.10 a necessity measure.

We also have to replace the inequality (11) by

$$P_\mu(\chi) \geq \min\{P_\mu(\chi_1), P_\mu(\chi_2)\}. \tag{15}$$

Since $\big((C,\mathcal{A},P),\mu\big)$ is nested, (15) is satisfied. All together we obtain

$$\mathrm{th}^{(a)}((\forall y)(\psi)) \leq \mathrm{Th}_a^{(\mathrm{poss})}((\forall y)(\psi)). \tag{16}$$

The opposite inequality to (16) is also fulfilled, since the interpretation (on the Herbrand universe) induced by $\mathrm{th}^{(a)}$ is by definition a necessity measure compatible with a, to which we obtain a corresponding nested context evaluation by Theorem 3.9, so that this context evaluation contributes to the infimum for $\mathrm{Th}_a^{(\mathrm{poss})}$. □

4 Conclusions

We have introduced extensions of Prolog to $[0,1]$–valued logics. We carry out these extensions on the basis of simple mappings $a : L \to [0,1]$ that can be understood as 'fuzzy' Prolog programs. In opposition to other approaches, we keep the truth values out of the logical language, so that we are able to define simple inference mechanisms, which lead to the soundness and completeness results presented in Section 2.

From the formal point of view these results are satisfactory, but they do not provide any hints for the interpretation of the truth values from the unit interval. To fill this gap of missing semantics for the truth values, we introduced the notion of context evaluations based on the idea of a set of consideration contexts or possible worlds weighted by a probability measure.

It turned out that the general idea of context evaluations provides an interpretation of 'fuzzy' Prolog based on the Łukasiewicz implication with a sound but incomplete proof theory. The restriction to nested context evaluations yields an interpretation for a 'fuzzy' Prolog based on the Gödel implication with a sound and complete proof theory. Nested context evaluations are also a possible interpretation for possibilistic logic, so that we obtain the equivalence of 'fuzzy' Prolog based on the Gödel implication, possibilistic Prolog, and the interpretation in the light of nested context evaluations.

We did not use the resolution principle [24] as the inference mechanism in our extensions of Prolog. This would only make sense if the connectives in $\bigoplus$ can be interpreted as conjunctions (i.e. if they are t–norms). Indeed, in this case we would obtain the same results if we would use the resolution principle modified with respect to the given lower bounds of the truth value. But even in this case, the resolution principle would not yield a severe improvement, since it is an efficient method for finding one proof, where we have to consider all proofs due to the supremum in Definition 2.3 (iii).

The advantage of our approach to $[0,1]$–valued extensions of Prolog compared to other fuzzy Prolog systems [10, 13, 7, 12, 26] is the clearly defined semantics of the truth values.

References

[1] D. Dubois, J. Lang, H. Prade, Automated Reasoning Using Possibilistic Logic: Semantics, Belief Revision and Variable Certainty Weights. Proc. 5th Workshop on Uncertainty in Artificial Intelligence, Windsor, Ontario (1989), 81–87.

[2] D. Dubois, J. Lang, H. Prade: Fuzzy Sets in Approximate Reasoning, Part 2: Logical Approaches. Fuzzy Sets and Systems 40 (1991), 203–244.

[3] D. Dubois, H. Prade, An Introduction to Possibilistic and Fuzzy Logics. In: P. Smets, E.H. Mamdani, D. Dubois, H. Prade (eds.), Non-Standard Logics for Automated Reasoning. Academic Press, London (1988), 287-326.

[4] D. Dubois, H. Prade, Fuzzy Sets, Probability and Measurement. European Journal of Operational Research 40 (1989), 135–154.

[5] D. Dubois, H. Prade, Epistemic Entrenchment and Possibilistic Logic. Artificial Intelligence 50 (1991), 223–239.

[6] J. Gebhardt, R. Kruse, A Possibilistic Interpretation of Fuzzy Sets by the Context Model. Proc. International Conference on Fuzzy Systems, IEEE, San Diego (1992), 1089–1096.

[7] M. Ishizuka, N. Kasai, Prolog–ELF Incorporating Fuzzy Logic. Proc. 9th IJCAI Conference, Los Angeles (1985), 701–703.

[8] F. Klawonn, J. Gebhardt, R. Kruse: Logical Approaches to Uncertainty and Vagueness in the View of the Context Model. Proc. International Conference on Fuzzy Systems, IEEE, San Diego (1992), 1375–1382.

[9] R. Kruse, J. Gebhardt, F. Klawonn, Reasoning with Mass Distributions and the Context Model. In: R. Kruse, P. Siegel (eds.), Symbolic and Quantitative Approaches to Uncertainty. Springer–Verlag, Berlin (1991), 81–85.

[10] R.C.T. Lee, Fuzzy Logic and the Resolution Principle. J. Association of Computation and Machines 19 (1972), 109–119.

[11] J.W. Lloyd: Foundations of Logic Programming (2nd ed.). Springer, Berlin (1987).

[12] T.P. Martin, J.F. Baldwin, B.W. Pilsworth, The Implementation of FProlog – A Fuzzy Prolog Interpreter. Fuzzy Sets and Systems 23 (1987), 119–129.

[13] M. Mukaidono, Z.L. Shen, L. Ding, Fundamentals of Fuzzy Prolog. Int. J. Approximate Reasoning 3 (1989), 179–193.

[14] D. Mundici, Ulam Games, Łukasiewicz Logic, and AF $C^{\star}$–Algebras. Fundamenta Informaticae 18 (1993), 151–161.

[15] N.J. Nilsson, Probabilistic Logic. Artificial Intelligence 28 (1986), 71–87.

[16] V. Novák, On the Syntactico–Semantical Completeness of First–Order Fuzzy Logic, Part I: Syntax and Semantics. Kybernetika 26 (1990), 47–66.

[17] V. Novák, On the Syntactico–Semantical Completeness of First–Order Fuzzy Logic, Part II: Main Results. Kybernetika 26 (1990), 134–154.

[18] V. Novák, The Unifying Role of Fuzzy Logic in Fuzzy Set Theory. Annales Univ. Sci. Budapest, Sect. Comp. 12 (1991), 187–194.

[19] V. Novák, On the Logical Basis of Approximate Reasoning. In: V. Novák, J. Ramík, M. Mareš, M. Černy, J. Nekola (eds), Fuzzy Approach to Reasoning and Decision Making. Academia, Prague (1992), 17–27.

[20] I.P. Orci, Programming in Possibilistic Logic. Int. J. Expert System 2 (1989), 79–96.

[21] J. Pavelka, On Fuzzy Logic I, II, III. Z. Math. Logik Grundlag. Math. 25 (1979), 45–52, 119–134, 447–464.

[22] H. Rasiowa, R. Sikorski, The Mathematics of Metamathematics (3rd ed.). Polish Scientific Publishers, Warszawa (1970).

[23] N. Rescher, Many–Valued Logic. McGraw–Hill, New York (1969).

[24] J.A. Robinson: A Machine–Oriented Logic Based on the Resolution Principle. Journ. of the Association for Computing Machinery 12 (1965), 23–41.

[25] W. Spohn, Ordinal Conditional Functions: A Dynamic Theory of Epistemic States. In: W.L. Harper, B. Skyrms (eds.), Causation in Decision, Belief, Change, and Statistics, Vol.2. Reidel, Dordrecht (1987), 105–134.

[26] M. Umano, Fuzzy Set Prolog. Proc. 2nd IFSA Congress, Tokyo (1987), 750–753.