University of Bern
Institute of Computer Science

Soundness and completeness of a first order probabilistic logic

with approximate conditional probabilities

Bachelor thesis

Author:
Stephan Matter

Supervisor:
Prof. Dr. Thomas Studer

1 Introduction

This thesis provides an introduction to a first order probabilistic logic with approximate conditional probabilities as it was presented in [4]. Classical first order logic does not have a built-in mechanism to express statements that are probably true. It does only cover the most basic statements, e.g. "all birds fly". This logic takes the approach of introducing a new probability quantifier which lets you express statements that are probably true like "50% of birds fly". It also provides a tool to formulate approximate conditional probabilities for statements like "generally every bird flies".

In this document we present and prove basic results as the soundness and the completeness theorem. Many proofs which are omitted in the original paper are explained here.

Acknowledgement

I would like to thank Jannis Kokkinis for his explanations and corrections and Thomas Studer for his support and patience for this Thesis.

Disclaimer

Many results in this thesis were taken from [9], [3] and [4]. There won't be a reference for every theorem or proof where exactly it is taken from. For a detailed understanding of the subject I recommend reading those sources.

Contact

Stephan Matter
Weberstrasse 23
3007 Bern
stephan.matter@hispeed.ch

Contents

1 Introduction **1**

2 Conditional Probabilities **3**

2.1 Basics of Probability Theory 3

2.2 Application to First Order Logic 4

2.3 Approximate Conditional Probabilities 5

3 Syntax of $L_{\omega\omega}^{\mathbf{P},\mathbb{I}}$ **6**

4 Semantics **8**

4.1 Structures of $L_{\omega\omega}^{\mathbf{P},\mathbb{I}}$ 8

4.2 Assignments and Satisfaction 9

5 Axiomatization **14**

5.1 Axioms 14

5.2 Inference Rules 15

5.3 Soundness Theorem 16

6 Deduction Theorem **22**

7 Consistency **24**

7.1 Properties of Consistent Sets 24

7.2 Lindenbaum Theorem 25

8 Completeness **27**

8.1 Language Extension 27

8.2 Canonical Model 30

8.3 Truth Lemma 32

References **35**

2 Conditional Probabilities

2.1 Basics of Probability Theory

This section provides a few definitions of terms in probability theory. We will omit everything not directly related to the application on first order logic. For a better understanding of the subject I recommend reading the book by H. Bauer on Probability Theory [2]. The contents of this section are freely adapted from [10] and [11].

Definition 2.1 (Probability Measure)**.** Let $\mathfrak{F}$ be a σ-algebra[1] over a set Ω. A measure $\mathbf{P}$ on $\mathfrak{F}$ is called *probability measure* if $\mathbf{P}(\Omega) = 1$.

A probability measure is defined on all elements of $\mathfrak{F}$ and its codomain is the unit interval $[0, 1]$.

Definition 2.2 (Probability Space)**.** If Ω is a set, $\mathfrak{F}$ a σ-Algebra on Ω and $\mathbf{P}$ a probability measure on $\mathfrak{F}$, then the tuple $(\Omega, \mathfrak{F}, \mathbf{P})$ is called *probability space*.

If A is a subset of $\mathfrak{F}$ we call A an *event* and $\mathbf{P}(A)$ the *probability* of A.

Example 2.3 (Deck of cards)**.** Let Ω be an ordinariy deck of playing cards, $\mathfrak{F} = 2^\Omega$ and A ={First card drawn is a club}. The probability of drawing a club out of a full deck is $1/4$ and therefore $\mathbf{P}(A) = 1/4$.
Let B = {Second card drawn is a club}. If the first card drawn was a club, $\mathbf{P}(B)$ is now slightly less than $\mathbf{P}(A)$. This leads us to the concept of conditional probability.

If A and B are elements of $\mathfrak{F}$ the conditional probability gives us the chance of B given that A already occured.

Definition 2.4 (Conditional Probability)**.** A,$B \in \mathfrak{F}$ and $\mathbf{P}(A) \neq 0$. The *conditional probability* of B given A is defined as

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} \tag{1}$$

Remark 2.5. It follows directly that $\mathbf{P}(B|\Omega) = \mathbf{P}(B)$.

For our example (2.3) we can now write $\mathbf{P}(B|A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} \approx 0.24$

[1]The definition of a σ-algebra can be found in [2] or [11]

2.2 Application to First Order Logic

Classical first order logic makes use of a universal quantifier $\forall$ and its dual $\exists$ to express that a statement holds for all assignments or that there exists at least one assignment that satifies the statement.

Example 2.6. (Deck of cards) Assume Club(x) is a unary relation which states that x is a club. We can use first order language to express for example the following statements:

- Every card of the deck is a club: $(\forall x)Club(x)$
- There is at least one club in the deck: $(\exists x)Club(x)$

While the first sentence is clearly wrong for an ordinary deck of cards the second statement provides us with very few information about the deck. We clearly can't formulate example 2.3 in a simple way.

It would therefore be convenient to be able to express probabilities in first order logic. In their paper [4], Ikodinović, Rašković, Marković and Ognjanović introduced a few new quantifiers to model conditional probabilities in first order logic.

Notation 2.7 (Conditional Probability Quantifier)**.** If α and β are formulas[2] and $r \in \mathbb{I}$ (Unit interval of the Hardy field) then $(CP\vec{x} \leq r)(\alpha|\beta)$ states that the probability of tuples $\vec{x}$ satisfying $\alpha \wedge \beta$ divided by the probability of tuples satisfying β is less or equal to r. $(CP\vec{x} \geq r)(\alpha|\beta)$ is defined analogously.[3]

Example 2.8 (Deck of cards)**.** With those new quantifiers we can now adequately express the statements of example 2.3:

- The probability of the first card being a club is 25%:
$$(CPx = 0.25)(Club(x)|\top).$$
- The probability of the second card being a club, given that the first one was a club, is less than 25%:
$$(CPxy < 0.25)(Club(y)|Club(x)).$$

[2]Definition on page 6.

[3]A formal definition for satisfaction of the CP quantifier follows on page 9.

2.3 Approximate Conditional Probabilities

A lot of real world applications require statements that are generally true, but may have some exceptions like “generally every bird flies” or “normally there are no temperatures above 10°C in winter”.

To be able to formulate such statements in our logic we add a single positive infinitesimal ϵ to the domain of r in the probability quantifier. The construction of infiniteimals from first order logic can be found in [3]. For this thesis it is enough to know that $\epsilon < \frac{1}{n}$ for all $n \in \mathbb{N}$ (we assume that $\mathbb{N}$ does not contain 0).

Definition 2.9 (Hardy Field)**.** The Hardy field $\mathbb{Q}(\epsilon)$ is the smallest ordered field obtained by adding an infinitesimal ϵ to the rational numbers.

We will denote the unit interval of the Hardy field $\mathbb{Q}(\epsilon)$ by $\mathbb{I}$ and the unit interval of rational numbers by $\mathbb{I}_{\mathbb{Q}}$

Definition 2.10 (infinitely close elements)**.** We say, that two elements $x, y \in \mathbb{Q}(\epsilon)$ are infinitely close, if $x - y$ is infinitesimal. We denote this by $x \approx y$.

Definition 2.11. (Monad) We define the set of elements infinitely close to an element $x \in \mathbb{Q}$ as

$$monad(x) = \{y|y \approx x\}.$$

Or by our understanding of infinitesimals:

$$monad(q) = \bigcap_{n\in\mathbb{N}} \left[max\{0, q - \frac{1}{n}\}, min\{1, q + \frac{1}{n}\} \right]$$

Notation 2.12 (approximate conditional probability quantifier)**.** If α and β are formulas and $q \in \mathbb{I}_{\mathbb{Q}}$, then $(CP\vec{x} \approx q)(\alpha|\beta)$ states that the probability of tuples $\vec{x}$ satisfying $\alpha \wedge \beta$ divided by the probability of tuples satisfying β is in $monad(q)$.

The statements from the beginning of this section would now look like this:

$$(CPx \approx 1)(Flies(x)|Bird(x)) \quad \text{“generally every bird flies”}$$

$$(CPx \approx 0)(Warm(x)|Winter(x)) \quad \text{“Normally there are no warm winter days”}$$

3 Syntax of $L_{\omega\omega}^{\mathbf{P},\mathbb{I}}$

In the last section we introduced new quantifiers for a first order logic. We will call this new logic $L_{\omega\omega}^{\mathbf{P},\mathbb{I}}$. This section provides the most important syntactic definitions and notations for $L_{\omega\omega}^{\mathbf{P},\mathbb{I}}$. For a detailed introduction to first order syntax consider reading [3]. All of the following definitions and notions are taken from [9] and [4].

Definition 3.1 (Terms)**.** Terms in $L_{\omega\omega}$ (classical first order logic) are inductively defined as follows:

1. a single variable x_i is a term for each $i \in \mathbb{N}$
2. $f(t_1, t_2, \ldots, t_n)$ is a term for every n-ary function f and terms $t_1, t_2, \ldots, t_n$ for $n \in \mathbb{N}_0$

If c is a 0-ary function, we call c a constant.
We will denote the set of terms of a language L by $\mathrm{Tm}(L)$.

Remark 3.2. The sets of terms in $L_{\omega\omega}^{\mathbf{P},\mathbb{I}}$ and $L_{\omega\omega}$ are equal.

Definition 3.3 (Formulas)**.** Formulas in $L_{\omega\omega}^{\mathbf{P},\mathbb{I}}$ are inductively defined as follows:

1. $R(t_1, t_2, \ldots, t_n)$ for an n-ary relation R and terms $t_1, t_2, \ldots, t_n$
2. if α is a formula, then so is $\neg\alpha$
3. if α and β are formulas, then so is $\alpha \wedge \beta$
4. if α is a formula and x is a variable, then $(\forall x)\alpha$ is a formula
5. if α and β are formulas $r \in \mathbb{I}$ and $q \in \mathbb{I}_{\mathbb{Q}}$, then $(CP\vec{x} \geq r)(\alpha|\beta)$, $(CP\vec{x} \leq r)(\alpha|\beta)$ and $(CP\vec{x} \approx q)(\alpha|\beta)$ are formulas.

We will denote the set of formulas of a language L by $\mathrm{Fm}(L)$.

Remark 3.4. Points 1 to 4 in definition 3.3 are the same as in $L_{\omega\omega}$.

Notation 3.5. In this thesis we use the following short forms:

- $\alpha \vee \beta$ for $\neg(\neg\alpha \wedge \neg\beta)$
- $\alpha \to \beta$ for $\neg\alpha \vee \beta$

- $\alpha \leftrightarrow \beta$ for $(\alpha \to \beta) \wedge (\beta \to \alpha)$
- $(\exists x)\alpha$ for $\neg(\forall x)\neg\alpha$
- $\perp$ for $\alpha \wedge \neg\alpha$
- $\top$ for $\neg\perp$
- $(CP\vec{x} > r)(\alpha|\beta)$ for $\neg(CP\vec{x} \leq r)(\alpha|\beta)$
- $(CP\vec{x} < r)(\alpha|\beta)$ for $\neg(CP\vec{x} \geq r)(\alpha|\beta)$
- $(CP\vec{x} = r)(\alpha|\beta)$ for $(CP\vec{x} \geq r)(\alpha|\beta) \wedge (CP\vec{x} \leq r)(\alpha|\beta)$
- $(P\vec{x} \diamond r)\alpha$ for $(CP\vec{x} \diamond r)(\alpha|\top)$ where $\diamond \in \{\approx, =, \leq, \geq, <, >\}$

Definition 3.6 (Bound Variables, Sentences, Theories)**.** Let $\vec{x}$ be a tuple of variables. $x_i \in \vec{x}$ is called *bound* if it is in the scope of a $(\forall x_i)$ or a $(CP\vec{x} \diamond r)$ quantifier. x_i is called free otherwise.

A formula without any free variables is called a *sentence.*

A set of sentences is called a *theory.*

Notation 3.7. We will denote the set of Sentences of a Language L by $\mathrm{Sen}(L)$.

The sets of free and bound variables of a formula α are denoted by $\mathrm{Fr}(\alpha)$ and $\mathrm{Bd}(\alpha)$ respectively.

If we write $\alpha(\vec{x})$ we mean that $\vec{x}$ contains all free variables of α.

Example 3.8. Let us consider a few examples of formulas and sentences: x, y, z are distinct variables, R is a unary and U is a binary relation.

- $(CPx \geq 0.5)(U(x,y)|R(x))$ (Formula; y occurs free)
- $(CPx \approx 1)((\exists y)U(x,y)|\top)$ or by simplified notation $(Px \approx 1)(\exists y)U(x,y)$ (Sentences)
- $(\forall x)(CPy \geq 0.3)(U(x,y)|R(y))$ (Sentence)
- $(CPxy \leq 0.25)(R(y)|(CPz \approx 0.5)(R(z)|U(x,y)))$ (Sentence)

Definition 3.9 (Substitution in terms)**.** For $s, t \in \mathrm{Tm}(L)$ and x, y distinct variables, the substitution of x by t in s, noted $s(x := t)$, is inductively defined as follows:

1. $x(x := t) = t$
2. $y(x := t) = y$
3. for each n-ary function f and terms $s_1, \dots, s_n$:
   $f(s_1, \dots, s_n)(x := t) = f(s_1(x := t), \dots, s_n(x := t))$

Definition 3.10 (Substitution in formulas)**.** For $\alpha \in \mathrm{Fm}(L)$, $t \in \mathrm{Tm}(L)$ and x a variable, the substitution of free occurences of x by t in α, noted $\alpha(x := t)$, is inductively defined as follows:

1. for each n-ary relation r and terms $s_1, \dots, s_n$:
   $r(s_1, \dots, s_n)(x := t) = r(s_1(x := t), \dots, s_n(x := t))$
2. $(\neg\alpha)(x := t) = \neg(\alpha(x := t))$
3. $(\alpha_1 \wedge \alpha_2)(x := t) = \alpha_1(x := t) \wedge \alpha_2(x := t)$
4. $((\forall y)\alpha)(x := t) = \begin{cases} (\forall y)(\alpha(x := t)) & \text{if } x \neq y \\ (\forall y)\alpha & \text{if } x = y \end{cases}$

To make sure, that a substituted term doesn't contain any quantified variables of α we allow the renaming of bound variables:

Definition 3.11 (Variants)**.** For $\alpha \in \mathrm{Fm}(L)$ we call α' a *variant* of α if it is the result of renaming some (or all) bound variables in α

Remark 3.12. From now on, if we write $\alpha(x := t)$ we always mean the result of substituting x by t in a suitable variant of α which contains no variables in common with t.

4 Semantics

4.1 Structures of $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$

Let us recall section 2: To be able to use probabilities we have to define an underlying probability-space $(\Omega, \mathfrak{F}, \mathbf{P})$. For $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$ those are contained in the $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$-structures:

Definition 4.1 (Structures for $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$)**.** An $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$-structure is a tuple $\bar{\mathfrak{A}} = \langle \mathfrak{A}, \mathfrak{F}_n, \mathbf{P}_n \rangle_{n \in \mathbb{N}}$ such that

1. $\mathfrak{A} = \langle A, r^A, f^A \rangle$ is an $L_{\omega\omega}$-structure for L

2. $(A^n, \mathfrak{F}_n, \mathbf{P}_n)$ is a finitely additive probability space for all $n \geq 1$ with the following additional requirements [4]:

   (a) for each n-ary function f, the graph of $f^{\mathfrak{A}}$ is in $\mathfrak{F}_{n+1}$.

   (b) for each n-ary relation R, $R^{\mathfrak{A}}$ is in $\mathfrak{F}_n$.

   (c) for all $i, j \leq n$, $\{x_1, \ldots, x_n | x_i = x_j\} \in \mathfrak{F}_n$.

   (d) if $X \in \mathfrak{F}_n$, then $A \times X \in \mathfrak{F}_{n+1}$.

   (e) if $X \in \mathfrak{F}_{n+m}$ and $\vec{b} \in A^m$, then $\{\vec{a} \in A^n | (\vec{a}, \vec{b}) \in X\} \in \mathfrak{F}_n$.

   (f) if $X \in \mathfrak{F}_{n+m}$, then $\{\vec{a} \in A^n | \mathbf{P}_m\{\vec{b} \in A^m | (\vec{a}, \vec{b}) \in X\} \diamond r\} \in \mathfrak{F}_n$, where $\diamond \in \{\leq, \geq, \approx\}$.

   (g) if $X \in \mathfrak{F}_n$, then $X^\pi = \{(a_{\pi(1)}, \ldots, a_{\pi(n)} \in A^n | (a_1, \ldots, a_n) \in X\} \in \mathfrak{F}_n$ and $\mathbf{P}_n(X^\pi) = \mathbf{P}_n(X)$ for every permutation π of $\{1, 2, \ldots, n\}$.

   (h) if $X \in \mathfrak{F}_n$, then $\mathbf{P}_{n+1}(A \times X) = \mathbf{P}_n(X)$

We will denote the class of structures for a language L by $\mathrm{Str}(L)$

4.2 Assignments and Satisfaction

Definition 4.2 (Assignement)**.** An assignment in a structure $\bar{\mathfrak{A}}$ is a function ν which maps variables to the universe A.

Notation 4.3. If it is required to specify the assignment we will use the notion $\nu(\vec{x} \mapsto \vec{a})$ to point out that $\vec{a} = (a_1, \ldots, a_n)$ is assigned to the variables $\vec{x} = (x_1, \ldots, x_n)$.

Definition 4.4 (Satisfaction)**.** For $\bar{\mathfrak{A}} \in \mathrm{Str}(L)$, assignment ν and $\alpha \in \mathrm{Fm}(L)$ we define "$\bar{\mathfrak{A}}$ satisfies α for ν", written $\bar{\mathfrak{A}}, \nu \models \alpha$ inductively as follows

1. for each n-ary relation r and terms $t_1, \ldots, t_n$
   $\bar{\mathfrak{A}}, \nu \models r(t_1, \ldots, t_n)$ iff $\langle t_1^{\bar{\mathfrak{A}},\nu}, \ldots, t_n^{\bar{\mathfrak{A}},\nu} \rangle \in r^A$

2. $\bar{\mathfrak{A}}, \nu \models \neg\alpha$ iff $\bar{\mathfrak{A}}, \nu \not\models \alpha$

3. $\bar{\mathfrak{A}}, \nu \models \alpha \wedge \beta$ iff $\bar{\mathfrak{A}}, \nu \models \alpha$ and $\bar{\mathfrak{A}}, \nu \models \beta$

4. $\bar{\mathfrak{A}}, \nu \models (\forall x)\alpha$ iff $\bar{\mathfrak{A}}, \nu(x \mapsto a) \models \alpha$ for all $a \in A$

[4] A justification of those requirements follows on page 10

5. $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \leq r)(\alpha|\beta)$ iff

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} = 0 \text{ and } r = 1$$

or

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} > 0 \text{ and}$$
$$\frac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \alpha \wedge \beta\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\}} \leq r$$

6. $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq r)(\alpha|\beta)$ iff

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} = 0$$

or

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} > 0 \text{ and}$$
$$\frac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \alpha \wedge \beta\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\}} \geq r$$

7. $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \approx q)(\alpha|\beta)$ iff

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} = 0 \text{ and } q = 1$$

or

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} > 0 \text{ and}$$
$$\frac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \alpha \wedge \beta\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\}} \in monad(q)$$

Remark 4.5. Definition 4.4 implies that if $\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} = 0$, the conditional probability of α given β is set to 1.

We now have the exact definition of the CP quantifier and are able to discuss some of the additional properties to the σ-algepras $\mathfrak{F}_n$ and the probability measures $\mathbf{P}_n$ introduced in definition 4.1:

(a) for each n-ary function f, the graph of $f^{\mathfrak{A}}$ is in $\mathfrak{F}_{n+1}$.

This assures measurability of functions

(b) for each n-ary relation R, $R^{\mathfrak{A}}$ is in $\mathfrak{F}_n$.

This assures measurability of all tuples $\langle a_1, \ldots, a_n \rangle \in R^{\mathfrak{A}} \subseteq A^n$

(c) for all $i, j \leq n$, $\{x_1, \ldots, x_n | x_i = x_j\} \in \mathfrak{F}_n$.

Sets with multiple occurences of the same variable have to be measurable.

(d) if $X \in \mathfrak{F}_n$, then $A \times X \in \mathfrak{F}_{n+1}$.

Let us assume that $\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \alpha \wedge \beta\} = X \in \mathfrak{F}_n$.

Consider the formula $\varphi = (CP\vec{x} \leq r)((\forall y)\alpha|\beta), y \notin \mathrm{Fr}(\beta)$.

Then by definition $\bar{\mathfrak{A}}, \nu \models \varphi$ iff

$$\begin{aligned}&\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} = 0 \text{ and } r = 1 \text{ or}\\ &\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} > 0 \text{ and}\\ &\quad \frac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models (\forall y)\alpha \wedge \beta\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\}} \leq r\end{aligned}$$

It follows directly from definition 4.4 that $\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models (\forall y)\alpha \wedge \beta\}$ is equivalent to $\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}; y \mapsto b) \models \alpha \wedge \beta$, for all $b \in A\}$. So $\{(\vec{a}, b) \in A^{n+1} | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}; y \mapsto b) \models \alpha \wedge \beta\} = A \times X$

Therefore $A \times X$ has to be measurable. In other words $A \times X \in \mathfrak{F}_{n+1}$.

(e) if $X \in \mathfrak{F}_{n+m}$ and $\vec{b} \in A^m$, then $\{\vec{a} \in A^n | (\vec{a}, \vec{b}) \in X\} \in \mathfrak{F}_n$.

(f) if $X \in \mathfrak{F}_{n+m}$ then $\{\vec{a} \in A^n | \mathbf{P}_m\{\vec{b} \in A^m | (\vec{a}, \vec{b}) \in X\} \diamond r\} \in \mathfrak{F}_n$, where $\diamond \in \{\leq, \geq, \approx\}$.
The two requirements above assure measurability even in the case of nested quantifiers (especially nested probability quantifiers). The reasoning is similar to (d).

(g) if $X \in \mathfrak{F}_n$, then $X^\pi = \{(a_{\pi(1)}, \ldots, a_{\pi(n)} \in A^n | (a_1, \ldots, a_n) \in X\} \in \mathfrak{F}_n$ and $\mathbf{P}_n(X^\pi) = \mathbf{P}_n(X)$ for every permutation π of $\{1, 2, \ldots, n\}$.
The order of the variables in a formula should not influence the measure.

Definition 4.6 (Validity). $\alpha \in \mathrm{Fm}(L)$ is *valid in* $\bar{\mathfrak{A}} \in \mathrm{Str}(L)$ if $\bar{\mathfrak{A}}, \nu \models \alpha$ for all assignments ν.

α is called *valid* (in general) if it is valid for all $\bar{\mathfrak{A}} \in \mathrm{Str}(L)$.

α is a *semantical consequence* of $\Gamma \subseteq \mathrm{Fm}(L)$, written $\Gamma \models \alpha$, if $\bar{\mathfrak{A}}, \nu \models \Gamma$ implies $\bar{\mathfrak{A}}, \nu \models \alpha$ for every $\bar{\mathfrak{A}} \in \mathrm{Str}(L)$ and every assignement ν.

To conclude this section we will prove a useful lemma for further application:

Lemma 4.7. *For $\bar{\mathfrak{A}} \in Str(L^{\mathbf{P},\mathbb{I}}_{\omega\omega})$, and $\alpha \in Fm(L^{\mathbf{P},\mathbb{I}}_{\omega\omega})$:*

1. *If* ν *and* υ *are* $\bar{\mathfrak{A}}$*-assignments which agree on each free variable of* α*, then*

$$\bar{\mathfrak{A}}, \nu \models \alpha \quad \textit{iff} \quad \bar{\mathfrak{A}}, \upsilon \models \alpha$$

2. *If* α' *is a variant of* α *then for every* $\bar{\mathfrak{A}}$*-assignment* ν

$$\bar{\mathfrak{A}}, \nu \models \alpha \quad \textit{iff} \quad \bar{\mathfrak{A}}, \nu \models \alpha'$$

3. *If* $t \in Tm(L^{\mathbf{P},\mathbb{I}}_{\omega\omega})$ *and* x *is a variable, then for every* $\bar{\mathfrak{A}}$*-assignment* ν

$$\bar{\mathfrak{A}}, \nu \models \alpha(x := t) \quad \textit{iff} \quad \bar{\mathfrak{A}}, \nu(x \mapsto t^{\bar{\mathfrak{A}},\nu}) \models \alpha$$

Proof. We will only prove the first part of the lemma, the other two parts are similarly proovable by induction on the complexity of α.

1. Claim 1: $t^{\bar{\mathfrak{A}},\nu} = t^{\bar{\mathfrak{A}},\upsilon}$ for each $t \in \mathrm{Tm}(L)$

   Proof: by induction over the complexity of t:

   base case $t \equiv x \Rightarrow t^{\bar{\mathfrak{A}},\nu} = \nu(x) = \upsilon(x) = t^{\bar{\mathfrak{A}},\upsilon}$

   induction step $t \equiv f(t_1, \ldots, t_n)$ for a function f and terms t_i

$$\begin{aligned} t^{\bar{\mathfrak{A}},\nu} &= f(t_1, \ldots, t_n)^{\bar{\mathfrak{A}},\nu} = f^{\bar{\mathfrak{A}}}(t_1^{\bar{\mathfrak{A}},\nu}, \ldots, t_n^{\bar{\mathfrak{A}},\nu}) \\ &\overset{I.H.}{=} f^{\bar{\mathfrak{A}}}(t_1^{\bar{\mathfrak{A}},\upsilon}, \ldots, t_n^{\bar{\mathfrak{A}},\upsilon}) = f(t_1, \ldots, t_n)^{\bar{\mathfrak{A}},\upsilon} = t^{\bar{\mathfrak{A}},\upsilon} \end{aligned}$$

■

   We will prove the first part of the Lemma by induction over the complexity of α

   base case $\alpha = R(t_1, \ldots, t_n)$

$$\begin{aligned} \bar{\mathfrak{A}}, \nu \models \alpha \quad &\text{iff} \quad (t_1^{\bar{\mathfrak{A}},\nu}, \ldots, t_n^{\bar{\mathfrak{A}},\nu}) \in R^{\bar{\mathfrak{A}}} \\ &\overset{\text{Claim 1}}{\text{iff}} \quad (t_1^{\bar{\mathfrak{A}},\upsilon}, \ldots, t_n^{\bar{\mathfrak{A}},\upsilon}) \in R^{\bar{\mathfrak{A}}} \\ &\text{iff} \quad \bar{\mathfrak{A}}, \upsilon \models R(t_1, \ldots, t_n) \end{aligned}$$

   induction step

   (i) $\alpha = \neg\beta$

$$\bar{\mathfrak{A}}, \nu \models \neg\beta \quad \text{iff} \quad \bar{\mathfrak{A}}, \nu \not\models \beta \quad \overset{\text{i.h.}}{\text{iff}} \quad \bar{\mathfrak{A}}, \upsilon \not\models \beta \quad \text{iff} \quad \bar{\mathfrak{A}}, \upsilon \models \neg\beta$$

(ii) $\alpha = \beta \wedge \gamma$

$$\begin{aligned} \bar{\mathfrak{A}}, \nu \models \beta \wedge \gamma \quad & \text{iff} \quad \bar{\mathfrak{A}}, \nu \models \beta \text{ and } \bar{\mathfrak{A}}, \nu \models \gamma \\ & \overset{\text{i.h.}}{\text{iff}} \quad \bar{\mathfrak{A}}, \upsilon \models \beta \text{ and } \bar{\mathfrak{A}}, \upsilon \models \gamma \\ & \text{iff} \quad \bar{\mathfrak{A}}, \upsilon \models \beta \wedge \gamma \end{aligned}$$

(iii) $\alpha = (\forall x)\beta$

$$\begin{aligned} \bar{\mathfrak{A}}, \nu \models (\forall x)\beta \quad & \text{iff} \quad \bar{\mathfrak{A}}, \nu(x \mapsto a) \models \beta \text{ for all } a \in A \\ & \overset{\text{i.h.}}{\text{iff}} \quad \bar{\mathfrak{A}}, \upsilon(x \mapsto a) \models \beta \text{ for all } a \in A \\ & \text{iff} \quad \bar{\mathfrak{A}}, \upsilon \models (\forall x)\beta \end{aligned}$$

(iv) $\alpha = (CP\vec{x} \geq r)(\beta|\gamma)$

$$\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq r)(\beta|\gamma) \qquad \text{iff}$$

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \gamma\} = 0 \tag{2}$$

or

$$\left(\begin{gathered} \mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \gamma\} > 0 \\ \text{and} \\ \frac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta \wedge \gamma\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \gamma\}} \geq r \end{gathered} \right) \tag{3}$$

By the inductive hypothesis, (2) is equivalent to:

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \upsilon(\vec{x} \mapsto \vec{a}) \models \gamma\} = 0 \tag{4}$$

And also by the inductive hypothesis, (3) is equivalent to:

$$\left(\begin{gathered} \mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \upsilon(\vec{x} \mapsto \vec{a}) \models \gamma\} > 0 \\ \text{and} \\ \frac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \upsilon(\vec{x} \mapsto \vec{a}) \models \beta \wedge \gamma\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \upsilon(\vec{x} \mapsto \vec{a}) \models \gamma\}} \geq r \end{gathered} \right) \tag{5}$$

Finally the conjunction of (4) and (5) is equivalent to

$$\bar{\mathfrak{A}}, \upsilon \models (CP\vec{x} \leq r)(\beta|\gamma)$$

(v) $\alpha = (CP\vec{x} \leq r)(\beta|\gamma)$ and $\alpha = (CP\vec{x} \approx q)(\beta|\gamma)$ analogously.

□

5 Axiomatization

Maybe the most famous Axioms in the world are the ones Euclid described in his book "the Elements" about geometry. He defines five rules to be true in all possible situations and tries to derive further statements out of these basic rules. When you first learn about geometry, rules like the parallel postulate seem quite obvious. But as soon as you learn about hyperbolic geometry you really see the neccessity for these formal axioms.

Now we want to apply the same strategy - as Euclid did for his geometry - for our first order logic $L_{\omega\omega}^{\mathbf{P};\mathbb{I}}$.

We make a difference between *axioms* and *inference rules*. Axioms are statements that are "obviously valid" in all $L_{\omega\omega}^{\mathbf{P};\mathbb{I}}$-models. An example from below: (CP1) tells us, that all conditional probabilities have to be greater or equal to zero, which is clearly given by the definition of a probability-measure. Inference rules are the only possibility to use the axioms or other formulas to deduce new propositions from them. As an example, if we know α to be valid and we also know that from α follows β we can now assume using (MP) that β is valid too (in a given model $\mathfrak{A}$).

5.1 Axioms

The following statements (FO), (CP1)...(CP7), (P1)...(P6) are axioms:

Let $r, r_1, r_2 \in \mathbb{I}, q \in \mathbb{I}_{\mathbb{Q}}$ and $\alpha, \beta \in \mathrm{Fm}(L)$

(FO) all $L_{\omega\omega}^{\mathbf{P};\mathbb{I}}$-instances of the axioms for $L_{\omega\omega}$

(CP1) $(CP\vec{x} \geq 0)(\alpha|\beta)$

(CP2) $(CP\vec{x} \leq r_1)(\alpha|\beta) \to (CP\vec{x} < r_2)(\alpha|\beta)$ for $r_1 < r_2$

(CP3) $(CP\vec{x} < r)(\alpha|\beta) \to (CP\vec{x} \leq r)(\alpha|\beta)$

(CP4) $(CP\vec{x} \approx q)(\alpha|\beta) \to (CP\vec{x} \geq q - \frac{1}{n})(\alpha|\beta)$ for every $n \in \mathbb{N}$ such that $0 \leq q - \frac{1}{n}$

(CP5) $(CP\vec{x} \approx q)(\alpha|\beta) \to (CP\vec{x} \leq q + \frac{1}{n})(\alpha|\beta)$ for every $n \in \mathbb{N}$ such that $q + \frac{1}{n} \leq 1$

(CP6) $(P\vec{x} = 0)\beta \to (CP\vec{x} = 1)(\alpha|\beta)$

(CP7) $((P\vec{x} = r_1)\beta \wedge (P\vec{x} = r_2)(\alpha \wedge \beta)) \to (CP\vec{x} = \frac{r_2}{r_1})(\alpha|\beta)$

(P1) $(P\vec{x} \geq 1)(\alpha \leftrightarrow \beta) \rightarrow ((P\vec{x} = r)\alpha \rightarrow (P\vec{x} = r)\beta)$

(P2) $(P\vec{x} \leq r)\alpha \leftrightarrow (P\vec{x} \geq 1 - r)\neg\alpha$

(P3) $((P\vec{x} = r_1)\alpha \wedge (P\vec{x} = r_2)\beta \wedge (P\vec{x} = 0)(\alpha \wedge \beta)) \rightarrow (P\vec{x} = \min\{1, r_1 + r_2\})(\alpha \vee \beta)$

(P4) $(Px_1 \cdots x_i \cdots x_n \diamond r)\alpha \leftrightarrow (Px_1 \cdots y \cdots x_n \diamond r)\alpha(x_i := y)$, where y is a variable not occuring in α and $\diamond \in \{\leq, \geq, \approx\}$

(P5) $(Px_1 \cdots x_n \diamond r)\alpha \leftrightarrow (Px_{\pi(1)} \cdots x_{\pi(n)} \diamond r)\alpha$, where π is a permuation of $\{1, \ldots, n\}$ and $\diamond \in \{\leq, \geq, \approx\}$

(P6) $(P\vec{x} \diamond r)\alpha(\vec{x}) \leftrightarrow (P\vec{x}\vec{y} \diamond r)\alpha(\vec{x})$, where the variables $\vec{y}$ are free in α and $\diamond \in \{\leq, \geq, \approx\}$

5.2 Inference Rules

The following statements (MP), (Gen), (Nec), (Ran) and (Approx) are inference rules: $r, \in \mathbb{I}, \alpha, \beta, \gamma \in \mathrm{Fm}(L)$

(MP) $\dfrac{\alpha \quad \alpha \rightarrow \beta}{\beta}$

(Gen) $\dfrac{\alpha}{(\forall x)\alpha}$

(Nec) $\dfrac{(\forall x)\alpha}{(Px = 1)\alpha}$

(Ran) $\dfrac{\alpha \rightarrow (P\vec{x} \neq r)\beta(\vec{x}), \text{ for all } r \in \mathbb{I}}{\alpha \rightarrow \perp}$

(Approx) for every $q \in \mathbb{I}_{\mathbb{Q}} \setminus \{0, 1\}$

$$\frac{\gamma \rightarrow (CP\vec{x} \geq q - \frac{1}{n})(\alpha|\beta), \text{ for all } n \geq \frac{1}{q} \quad (CP\vec{x} \leq q + \frac{1}{n})(\alpha|\beta), \text{ for all } n \geq \frac{1}{1-q}}{\gamma \rightarrow (CP\vec{x} \approx q)(\alpha|\beta)}$$

$$\frac{\gamma \rightarrow (CP\vec{x} \geq 1 - \frac{1}{n})(\alpha|\beta), \text{ for all } n \geq 1}{\gamma \rightarrow (CP\vec{x} \approx 1)(\alpha|\beta)}$$

$$\frac{\gamma \rightarrow (CP\vec{x} \leq \frac{1}{n})(\alpha|\beta), \text{ for all } n \geq 1}{\gamma \rightarrow (CP\vec{x} \approx 0)(\alpha|\beta)}$$

where $n \in \mathbb{N}$

Definition 5.1 (Proof). A *proof of* $\alpha \in \mathrm{Fm}(L)$ *from* $\Sigma \subseteq \mathrm{Fm}(L)$ is a sequence $\gamma_1, \ldots, \gamma_n \subseteq \mathrm{Fm}(L)$, such that $\gamma_n = \alpha$ and

for $i = 1 \ldots n$

either γ_i is an axiom
or $\gamma_i \in \Sigma$
or γ_i is derived from $\gamma_1, \ldots, \gamma_{i-1}$ using (MP), (Gen), (Nec) (Ran) or (Approx)

We write $\Sigma \vdash \alpha$ if α has a proof from Σ.

Remark 5.2.

(a) We say α is deducible (or derivable) from Σ if $\Sigma \vdash \alpha$.

(b) If Σ is the empty set we write $\vdash \alpha$ and call α a theorem.

(c) Do not confuse the term "proof of α" with the same term in meta language. It should always be clear from the context which one is meant.

5.3 Soundness Theorem

Although the validity of those Axioms and inference rules seems obvious, we need to proof that they are "sound in respect to $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$-models". This means that if we deduce a new formula α from a given set of formulas Σ, all models satisfying Σ also satisfy α, or a bit shorter...

Theorem 5.3 (Soundness). *For* $\Sigma \cup \{\gamma\} \subseteq$ *Fm(L):*

$$\Sigma \vdash \gamma \quad \Rightarrow \quad \Sigma \models \gamma$$

Before we jump into the proof of the soundness theorem we introduce a few notations and a useful lemma:

For convenience and readability reasons we write the following terms only once and reference them wherever needed:

$$\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\} \tag{$\mathbf{P}_\beta$}$$

$$\frac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \alpha \wedge \beta\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\}} \tag{$\mathbf{P}_{\frac{\alpha\wedge\beta}{\beta}}$}$$

It is possible that it is needed to replace β or α in the expression above. We will denote this by replacing the corresponding formulas in the subscript: e.g. $\mathbf{P}_{\frac{\gamma_1 \to \gamma_2 \wedge \beta}{\beta}}$ means $\dfrac{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \gamma_1 \to \gamma_2 \wedge \beta\}}{\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \mapsto \vec{a}) \models \beta\}}$

Lemma 5.4.

1. *$\bar{\mathfrak{A}}, \nu \models (CP\vec{x} < r)(\alpha|\beta)$ iff $\mathbf{P}_\beta > 0$ and $\mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} < r$*
2. *$\bar{\mathfrak{A}}, \nu \models (CP\vec{x} > r)(\alpha|\beta)$ iff*
   $(\mathbf{P}_\beta = 0$ and $r < 1)$ or $(\mathbf{P}_\beta > 0$ and $\mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} > r)$
3. *$\bar{\mathfrak{A}}, \nu \models (P\vec{x} \diamond r)\alpha$ iff $\mathbf{P}_\alpha \diamond r$ for $\diamond \in \{\approx, =, \leq, \geq, <, >\}$*

Proof (lemma 5.4).

1.

$$\begin{array}{ll} & \bar{\mathfrak{A}}, \nu \models (CP\vec{x} < r)(\alpha|\beta) \\ \text{iff} & \bar{\mathfrak{A}}, \nu \models \neg(CP\vec{x} \geq r)(\alpha|\beta) \\ \text{iff} & \bar{\mathfrak{A}}, \nu \not\models (CP\vec{x} \geq r)(\alpha|\beta) \\ \text{iff not} & \mathbf{P}_\beta = 0 \text{ or } (\mathbf{P}_\beta > 0 \text{ and } \mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} \geq r) \\ \text{iff} & \mathbf{P}_\beta > 0 \text{ and } \mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} < r \end{array}$$

2.

$$\begin{array}{ll} & \bar{\mathfrak{A}}, \nu \models (CP\vec{x} > r)(\alpha|\beta) \\ \text{iff} & \bar{\mathfrak{A}}, \nu \models \neg(CP\vec{x} \leq r)(\alpha|\beta) \\ \text{iff} & \bar{\mathfrak{A}}, \nu \not\models (CP\vec{x} \leq r)(\alpha|\beta) \\ \text{iff not} & (\mathbf{P}_\beta = 0 \text{ and } r = 1) \text{ or } (\mathbf{P}_\beta > 0 \text{ and } \mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} \leq r) \\ \text{iff} & (\mathbf{P}_\beta = 0 \text{ and } r < 1) \text{ or } (\mathbf{P}_\beta > 0 \text{ and } \mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} > r) \end{array}$$

3.

$$\begin{array}{ll} & \bar{\mathfrak{A}}, \nu \models (P\vec{x} \diamond r)\alpha \\ \text{iff} & \bar{\mathfrak{A}}, \nu \models (CP\vec{x} \diamond r)(\alpha|\top) \end{array}$$

From $\mathbf{P}_\top = 1$ and the definition of CP follows directly that $\mathbf{P}_\alpha \diamond r$

□

Proof (Soundness Theorem). We will proove that every axiom and every derivation rule is sound with respect to the class of $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$-structures.

Claim 2:

1. All axioms for $L_{\omega\omega}$ are valid
2. (MP) and (Gen) are valid

We omit the proof of this claim as it is covered in [9].

Claim 3: The axioms (CP1) to (CP6) are valid

Proof:

(CP1) $(CP\vec{x} \geq 0)(\alpha|\beta)$
This holds for every assignment, as by definition $\mathbf{P}_n(S) \geq 0$ for all $S \in \mathfrak{F}_n$

For all the following axioms we have to show, that if an assignment satisfies the requirements (left side of $\rightarrow$) it also satisfies their logical consequence (right side of $\rightarrow$).

(CP2) $(CP\vec{x} \leq r_1)(\alpha|\beta) \rightarrow (CP\vec{x} < r_2)(\alpha|\beta)$ for $r_1, r_2 \in [0,1]$ with $r_1 < r_2$

Assume that $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \leq r_1)(\alpha|\beta)$.
This holds iff $\mathbf{P}_\beta > 0$ and $\mathbf{P}_{\frac{\alpha\wedge\beta}{\beta}} \leq r_1$

then $\mathbf{P}_\beta > 0$ and $\mathbf{P}_{\frac{\alpha\wedge\beta}{\beta}} < r_2$

And by lemma 5.4 it follows that $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} < r_2)(\alpha|\beta)$

(CP3) $(CP\vec{x} < r)(\alpha|\beta) \rightarrow (CP\vec{x} \leq r)(\alpha|\beta)$

Assume that $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} < r)(\alpha|\beta)$.
By lemma 5.4 this holds iff $\mathbf{P}_\beta > 0$ and $\mathbf{P}_{\frac{\alpha\wedge\beta}{\beta}} < r$.

and by weakening the uneqation it follows that $\mathbf{P}_\beta > 0$ and $\mathbf{P}_{\frac{\alpha\wedge\beta}{\beta}} \leq r$

and therefore $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \leq r)(\alpha|\beta)$

(CP4) $(CP\vec{x} \approx q)(\alpha|\beta) \rightarrow (CP\vec{x} \geq q - \frac{1}{n})(\alpha|\beta)$ for every $n \in \mathbb{N}$ such that

$0 \leq q - \frac{1}{n}$

Assume that $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \approx q)(\alpha|\beta)$.

This holds iff $\mathbf{P}_{\beta} = 0$ and $q = 1$ or ($\mathbf{P}_{\beta} > 0$ and $\mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} \in \text{monad}(q)$)

By the definition of monad(q) the above means that

$\mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} \in \bigcap_{n \in \mathbb{N}} [\max\{0, q - \frac{1}{n}\}, \min\{1, q + \frac{1}{n}\}]$.

and therefore $\mathbf{P}_{\beta} = 0$ and $q = 1$ or $\mathbf{P}_{\beta} > 0$ and $\mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} \geq q - \frac{1}{n}$

which are the requirements for $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq q - \frac{1}{n})(\alpha|\beta)$.

(CP5) $(CP\vec{x} \approx q)(\alpha|\beta) \to (CP\vec{x} \leq q + \frac{1}{n})(\alpha|\beta)$ for every $n \in \mathbb{N}$ such that $q + \frac{1}{n} \leq 1$

Analogue to (CP4).

(CP6) $(P\vec{x} = 0)\beta \to (CP\vec{x} = 1)(\alpha|\beta)$

Assume that $\bar{\mathfrak{A}}, \upsilon \models (P\vec{x} = 0)\beta$.

which is the same as $\bar{\mathfrak{A}}, \upsilon \models (CP\vec{x} = 0)(\beta|\top)$

Therefore $\mathbf{P}_n\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \beta \wedge \top\} = \mathbf{P}_{\beta} = 0$

and it follows from the definition that $\bar{\mathfrak{A}}, \upsilon \models (CP\vec{x} = 1)(\alpha|\beta)$

■

Claim 4: The axioms (P1) to (P6) are valid

Proof:

(P1) $(P\vec{x} \geq 1)(\alpha \leftrightarrow \beta) \to ((P\vec{x} = r)\alpha \to (P\vec{x} = r)\beta)$

Assume that $\bar{\mathfrak{A}}, \nu \models (P\vec{x} \geq 1)(\alpha \leftrightarrow \beta)$

By lemma 5.4 this implies $\mathbf{P}_{\alpha \leftrightarrow \beta} \geq 1$

Which is the same as writing $\mathfrak{A} \models \alpha \leftrightarrow \beta$

ergo every assignment that satisfies α satisfies β and vice-versa.

Therefore if $\vec{c}$ is in $\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \alpha\}$ it is in $\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \beta\}$ too. Thus the two sets and their measure are equal.

By lemma 5.4 it follows directly that $((P\vec{x} = r)\alpha \to (P\vec{x} = r)\beta)$ and $((P\vec{x} = r)\beta \to (P\vec{x} = r)\alpha)$

(P2) $(P\vec{x} \le r)\alpha \leftrightarrow (P\vec{x} \ge 1 - r)\neg\alpha$

$\bar{\mathfrak{A}}, \nu \models (P\vec{x} \le r)\alpha \quad \text{iff} \quad \mathbf{P}_\alpha \le r$
If $\vec{c} \in A^n$ then either $\vec{c} \in \{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \alpha\}$ or
$\vec{c} \in \{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \not\models \alpha\}$ which is equal to $\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \neg\alpha\}$.
Therefore $\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \alpha\} \cap \{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \neg\alpha\} = \emptyset$
and because $\mathbf{P}_n$ is a countably additive probability measure
$\mathbf{P}_\alpha + \mathbf{P}_{\neg\alpha} = 1$ or $\mathbf{P}_{\neg\alpha} = 1 - \mathbf{P}_\alpha$
Hence if $\mathbf{P}_\alpha \le r$ then $\mathbf{P}_{\neg\alpha} \ge 1 - r$ and vice versa.

(P3) $((P\vec{x} = r_1)\alpha \wedge (P\vec{x} = r_2)\beta \wedge (P\vec{x} = 0)(\alpha \wedge \beta)) \to (P\vec{x} = \min\{1, r_1 + r_2\})(\alpha \vee \beta)$

$\bar{\mathfrak{A}}, \nu \models (P\vec{x} = r_1)\alpha \wedge (P\vec{x} = r_2)\beta \wedge (P\vec{x} = 0)(\alpha \wedge \beta))$ iff
$\mathbf{P}_\alpha = r_1$ and $\mathbf{P}_\beta = r_2$ and $\mathbf{P}_{\alpha\wedge\beta} = 0$
$\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \alpha \vee \beta\}$ is the same as
$\{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \alpha\} \cup \{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \beta\}$
Then because of $\mathbf{P}_{\alpha\wedge\beta} = 0$ and the additivity of the measure we get
$\mathbf{P}_{\alpha\vee\beta} = r_1 + r_2$.

(P4) $(Px_1 \cdots x_i \cdots x_n \diamond r)\alpha \leftrightarrow (Px_1 \cdots y \cdots x_n \diamond r)\alpha(x_i := y)$, where y is a variable not occuring in α and $\diamond \in \{\le, \ge, \approx\}$
This follows directly from lemma 4.7.

(P5) $(Px_1 \cdots x_n \diamond r)\alpha \leftrightarrow (Px_{\pi(1)} \cdots x_{\pi(n)} \diamond r)\alpha$, where π is a permuation of $\{1, \ldots, n\}$ and $\diamond \in \{\le, \ge, \approx\}$
This is a direct consequence of requirement (g) to the probability measure $\mathbf{P}_n$.

(P6) $(P\vec{x} \diamond r)\alpha(\vec{x}) \leftrightarrow (P\vec{x}\vec{y} \diamond r)\alpha(\vec{x})$, where $y_i \notin \text{Fr}(\alpha)$ for all i and $\diamond \in \{\le, \ge, \approx\}$
$\bar{\mathfrak{A}}, \nu \models (P\vec{x} \diamond r)\alpha(\vec{x})$ iff $\mathbf{P}_\alpha \diamond r$
Since y_i are not free in α for all i, it follows that
$\{(\vec{a}, \vec{b}) \in A^{n+m} | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}, \vec{y} \to \vec{b}) \models \alpha\} = \{\vec{a} \in A^n | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}) \models \alpha\} \times A^m$ Then by clause 2h in definition 4.1 we get
$\mathbf{P}_{n+m}\{(\vec{a}, \vec{b}) \in A^{n+m} | \bar{\mathfrak{A}}, \nu(\vec{x} \to \vec{a}, \vec{y} \to \vec{b}) \models \alpha\} = \mathbf{P}_\alpha$
And $\mathbf{P}_\alpha \diamond r$

$\blacksquare$

Claim 5: Inference Rules (Nec), (Ran) and (Approx) are valid

Proof:

(Nec) $\dfrac{(\forall x)\alpha}{(Px = 1)\alpha}$

$\bar{\mathfrak{A}}, \nu \models (\forall x)\alpha$ iff $\bar{\mathfrak{A}}, \nu(x \to a) \models \alpha$ for all $a \in A$.
Therefore $\{a \in A | \bar{\mathfrak{A}}, \nu(x \to a) \models \alpha\} = A$
This implies $\mathbf{P}_\alpha = 1$ and by lemma 5.4 $\bar{\mathfrak{A}}, \nu \models (Px = 1)\alpha$

(Ran) $\dfrac{\alpha \to (P\vec{x} \neq r)\beta(\vec{x}), \text{ for all } r \in \mathbb{I}}{\alpha \to \bot}$

Assume $\bar{\mathfrak{A}}, \nu \models \alpha \to (P\vec{x} \neq r)\beta(\vec{x})$ for all $r \in \mathbb{I}$.
Then $\mathbf{P}_\beta \neq r$ for all $r \in \mathbb{I}$, which means that $\mathbf{P}_\beta$ is not a valid value of a probability measure implying $\bot$.

(Approx) for every $q \in (0,1) \cap \mathbb{Q}$

$$\frac{\gamma \to (CP\vec{x} \geq q - \tfrac{1}{n})(\alpha|\beta), \text{ for all } n \geq \tfrac{1}{q} \quad \gamma \to (CP\vec{x} \leq q + \tfrac{1}{n})(\alpha|\beta), \text{ for all } n \geq \tfrac{1}{1-q}}{\gamma \to (CP\vec{x} \approx q)(\alpha|\beta)}$$

$$\frac{\gamma \to (CP\vec{x} \geq 1 - \tfrac{1}{n})(\alpha|\beta), \text{ for all } n \geq 1}{\gamma \to (CP\vec{x} \approx 1)(\alpha|\beta)}$$

$$\frac{\gamma \to (CP\vec{x} \leq \tfrac{1}{n})(\alpha|\beta), \text{ for all } n \geq 1}{\gamma \to (CP\vec{x} \approx 0)(\alpha|\beta)}$$

where $n \in \mathbb{N}$

Assume $\bar{\mathfrak{A}}, \nu$ satisfies the first part above and $\bar{\mathfrak{A}}, \nu \models \gamma$.
Then $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq q - \dfrac{1}{n})(\alpha|\beta)$, for all $n \geq \dfrac{1}{q}$
and $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \leq q + \dfrac{1}{n})(\alpha|\beta)$, for all $n \geq \dfrac{1}{1-q}$.
Therefore $\mathbf{P}_{\frac{\alpha \wedge \beta}{\beta}} \in \text{monad}(q)$.

The other two cases are similar.

■

We have now showed that all axioms and inference rules are sound for structures in $L_{\omega\omega}^{\mathbf{P};\mathbb{I}}$. The conclusion is: for all $\Sigma \cup \alpha \subseteq \text{Fm}(L)$: $\Sigma \vdash \alpha \Rightarrow \Sigma \models \alpha$ □

6 Deduction Theorem

Theorem 6.1 (Deduction Theorem). *For $\varphi \in Sen(L)$ and $\Gamma \cup \{\psi\} \subseteq Fm(L)$*

$$\Gamma \cup \{\varphi\} \vdash \psi \quad \textit{iff} \quad \Gamma \vdash \varphi \to \psi$$

Proof.
<u>$\Leftarrow$:</u> Assume $\Gamma \vdash \varphi \to \psi$, then also $\Gamma \cup \{\varphi\} \vdash \varphi \to \psi$ and by using (MP) it follows that $\Gamma \cup \{\varphi\} \vdash \psi$
<u>$\Rightarrow$:</u> Assume that a proof of ψ from Γ is the sequence $\psi_1, \psi_2 \dots \psi_n = \psi$ where n is the length of the proof. We will show this direction using induction on n.

base case n=1: Either ψ is an axiom or $\psi \in \Gamma \cup \{\varphi\}$. Then we have the following proof for $\varphi \to \psi$:

ψ	assumption
$\psi \to (\varphi \to \psi)$	first-order tautology
$\varphi \to \psi$	(MP)

induction step The cases where $\psi (= \psi_n)$ is obtained using (MP) or (Gen) are covered in [9]. What is left are the ones using (Nec), (Ran) or (Approx).

(i) Suppose $\psi_n = (P\vec{x} = 1)\alpha$ is obtained using (Nec) from $\psi_i = (\forall x)\alpha$ for $i < n$. Then we get a derivation of $\varphi \to \psi_n$ as follows:

$\varphi \to (\forall x)\alpha$	induction hypothesis
$(\forall x)(\varphi \to \alpha)$	φ is a sentence
$(P\vec{x} = 1)(\varphi \to \alpha)$	(Nec)
$\varphi \to (P\vec{x} = 1)\alpha$	φ is a sentence

(ii) Suppose $\psi_n = \gamma \to \bot$ is obtained using (Ran) from $\psi_r = \alpha \to (P\vec{x} \neq r)\beta(\vec{x})$ for all $r \in [0, 1]$. Note, that we have infinitely many premises for this rule. Then we get a derivation of $\phi \to \psi_n$ as follows:

$\varphi \to (\alpha \to (P\vec{x} \neq r)\beta(\vec{x}))$	induction hypothesis
$(\varphi \wedge \alpha) \to (P\vec{x} \neq r)\beta(\vec{x})$	
$(\varphi \wedge \alpha) \to \bot$	(Ran)
$\varphi \to (\alpha \to \bot)$	

(iii) The case where $\psi_n = \gamma \to (CP\vec{x} \approx q)(\alpha|\beta)$ is obtained from (Approx) is covered in [4] and is very similar to the others.

$\square$

7 Consistency

Consistency is a concept which tells us something about "contradictions" in sets of formulas. If such a set is consistent, then there are no contradictions. If it is inconsistent, then we can deduce any formula of our language from this set of formulas. This makes consistency to a very powerful property. As a conclusion we will prove, that every consistent set has a model.

Definition 7.1. (Consitency, Maximal Consistent)

1. A set of formulas Σ is called *consistent* if there is at least one formula that is not deducible from Σ (else it is called *inconsistent*).
2. A set of formulas Γ is called *maximal consistent* if Γ is consistent and for every $\alpha \in \mathrm{Fm}(L)$, either $\alpha \in \Gamma$ or $\neg\alpha \in \Gamma$.

7.1 Properties of Consistent Sets

Lemma 7.2 (Property of consistent sets)**.** *Let Σ be a consistent set of formulas. If $\alpha \in \Sigma$ then $\neg\alpha \notin \Sigma$*

Proof. Assume $\alpha \in \Sigma$ and $\neg\alpha \in \Sigma$, then by using the first order tautology $\alpha \to (\neg\alpha \to \beta)$ (which is valid for any $\beta \in \mathrm{Fm}(L)$) we obtain a deduction of every formula from Σ which is a contradiction to its consistency. □

Lemma 7.3 (Characterisation of inconsistent sets)**.**

(a) Σ is inconsistent iff $\Sigma \vdash \alpha$ for all $\alpha \in Fm(L)$.

(b) Σ is inconsistent iff $\Sigma \vdash \bot$

Proof.

(a) This is just a reformulation of definition 7.1

(b) $\Rightarrow$: Follows directly from (a). $\Leftarrow$: Using the first order tautology $\bot \to \beta$, $\beta \in \mathrm{Fm}(L)$ it follows that $\Sigma \vdash \beta$ for all β.

□

Lemma 7.4 (Properties of maximal consistent sets)**.** *Let Γ be a maximal consistent set of formulas.*

(a) for every consistent set Γ' with $\Gamma \subseteq \Gamma'$ it follows that $\Gamma = \Gamma'$

(b) Γ is deductively closed, i.e. if $\Gamma \vdash \alpha$ for a sentence α then $\alpha \in \Gamma$

(c) $\alpha \wedge \beta \in \Gamma$ iff $\alpha \in \Gamma$ and $\beta \in \Gamma$

(d) $\alpha \vee \beta \in \Gamma$ iff $\alpha \in \Gamma$ or $\beta \in \Gamma$

(e) If $\alpha, \alpha \to \beta \in \Gamma$, then $\beta \in \Gamma$

Proof. (a) Assume $\Gamma \subsetneq \Gamma'$, then there exists $\alpha \in \Gamma'$ with $\alpha \notin \Gamma$, but then by definition $\neg\alpha \in \Gamma$ and therefore $\neg\alpha \in \Gamma'$. So Γ' is inconsistent by Lemma 7.2.

(b) Let $\Gamma \vdash \alpha$ for $\alpha \in \mathrm{Sen}(L)$. Assume $\Gamma \cup \{\alpha\}$ is inconsistent, then $\Gamma \cup \{\alpha\} \vdash \neg\alpha$ and by applying the deduction theorem $\Gamma \vdash \alpha \to \neg\alpha$. Using the first order tautology $(\alpha \to \neg\alpha) \to \neg\alpha$ we get $\Gamma \vdash \neg\alpha$. This is a contradiction to the consistency of Γ and therefore $\Gamma \cup \{\alpha\}$ has to be consistent and by the maximality of Γ it follows $\alpha \in \Gamma$.

(c) Assume $\alpha \wedge \beta \in \Gamma$ then $\Gamma \vdash \alpha \wedge \beta$ and thus $\Gamma \vdash \alpha$ and $\Gamma \vdash \beta$. As Γ is maximal consistent: $\alpha \in \Gamma$ and $\beta \in \Gamma$.

(d) Assume $\alpha \vee \beta \in \Gamma$ then $\Gamma \vdash \alpha \vee \beta$ and thus $\Gamma \vdash \alpha$ or $\Gamma \vdash \beta$. As Γ is maximal consistent: $\alpha \in \Gamma$ or $\beta \in \Gamma$.

(e) Assume $\alpha, \alpha \to \beta \in \Gamma$ then $\Gamma \vdash \alpha$ and $\Gamma \vdash \alpha \to \beta$. Therefore $\Gamma \vdash \beta$ and as Γ is maximal consistent:$\beta \in \Gamma$.

□

7.2 Lindenbaum Theorem

Theorem 7.5 (Lindenbaum)**.** *For every consistent set of formulas Σ there exists a maximal consistent set Σ^* s.t. $\Sigma \subseteq \Sigma^*$.*

Proof.
We use Zorn's lemma to show the theorem:

Zorn's lemma: Any poset in which every chain has an upper bound has a maximal element.

We define $S = \{\Sigma' | \Sigma \subseteq \Sigma', \Sigma' consistent\}$
Let $(\Gamma_1 \subseteq \Gamma_2 \subseteq \dots)$ be a chain in S, then $\Gamma = \bigcup_i \Gamma_i$ is clearly an upper limit for the chain.

Claim 6: If $(\Gamma_1 \subseteq \Gamma_2 \subseteq \dots)$ is a chain in S, then $\Gamma = \bigcup_i \Gamma_i$ is in S

Proof: We prove this by contradiction: Assume Γ is inconsistent (by other words: not in S). Then by Lemma 7.3 it follows that $\Gamma \vdash \perp$. Let $\psi_1, \psi_2 \dots \psi_n = \perp$ be the proof of $\perp$. Then for each $i = 1, 2, \dots, n-1$ there is a set of formulas Γ_j with $\Gamma_j \vdash \psi_i$. As $(\Gamma_1 \subseteq \Gamma_2 \subseteq \dots)$ is a chain, there is a largest set Γ_k which contains all Γ_j needed for this proof and therefore $\Gamma_k \vdash \perp$ which contradicts its consistency. $\Rightarrow \Gamma$ is consistent (Γ is in S). ■

There are now all requirements for Zorn's lemma fulfilled and we can conclude that there is a maximal element Σ^* for each consistent set Σ.

Claim 7: Σ^* is maximal consistent.

Proof: Suppose Σ^* is not maximal consistent. Then there is a ψ for which $\Sigma^* \nvdash \psi$ and $\Sigma^* \nvdash \neg\psi$, from whose follows that $\psi \notin \Sigma^*$ and $\Sigma^* \cup \{\psi\}$ is consistent. Therefore $\Sigma^* \cup \{\psi\}$ is in S which contradicts the maximality of Σ^* ■

□

8 Completeness

Theorem 8.1 (Completeness theorem)**.** *If Σ is a consistent set of formulas, then Σ has a model.*

The proof of this theorem will consist of the following steps:

Step 1 Extend the language L and construct a maximal consistent set Σ^* which is a superset of Σ.

Step 2 Define the canonical model for Σ^*

Step 3 Prove the truth-lemma

Step 1:

8.1 Language Extension

Our target in this first step is to obtain a maximal consistent set Σ^* which has witnesses, i.e. if $\exists x\alpha$ is in Σ^* then there is a constant Symbol c in our extended Language such that $\alpha(x := c) \in \Sigma^*$.

To achieve this, we first need to extend our language L by a denumerable set of constants $C = \{c_n | n \in \mathbb{N}\}$. We call the new language $L^* = L \cup C$.

Let $\{\alpha_n | n \in \mathbb{N}\}$ be an enumeration of all formulas of $(L^*)^{\mathbf{P},\mathbb{I}}_{\omega\omega}$. We construct our desired set Σ^* by defining a sequence $(\Sigma_1 \subset \Sigma_2 \subset \Sigma_3 \subset \dots)$. Following lindenbaums theorem we know, that there is a maximal consistent set Σ^* which is a superset of all Σ_i.

Let $\Sigma_0 = \Sigma$
For all $\alpha_n, n \in \mathbb{N}$:

If α_n is a sentence

then if $\Sigma_n \cup \alpha_n$ is consistent

then if α_n is of the form $\exists x\alpha(x)$ we define
$\Sigma_{n+1} = \Sigma_n \cup \{\alpha_n, \alpha(x := c_n)\}$
else $\Sigma_{n+1} = \Sigma_n \cup \{\alpha_n\}$

else $\Sigma_n \cup \alpha_n$ is inconsistent

then if α_n is of the form $\gamma \to (CP\vec{x} \approx q)(\alpha|\beta)$ we define
$\Sigma_{n+1} = \Sigma_n \cup \{\neg\alpha_n, \gamma \to \neg(CP\vec{x} \geq q - \frac{1}{k})(\alpha|\beta)\}$ or
$\Sigma_{n+1} = \Sigma_n \cup \{\neg\alpha_n, \gamma \to \neg(CP\vec{x} \leq q + \frac{1}{k})(\alpha|\beta)\}$
for a suitable $k \in \mathbb{N}$ such that Σ_{n+1} is consistent.

else $\Sigma_{n+1} = \Sigma_n \cup \{\neg\alpha_n\}$

else α_n contains free variables $\vec{x}$ and we define $\Sigma_{n+1} = \Sigma_n \cup \{(P\vec{x} = r)\alpha_n\}$ for a suitable $r \in \mathbb{I}$ such that Σ_{n+1} is consistent.

It remains the question, if each step of the construction produces a consistent set:

If $\Sigma_{n+1} = \Sigma_n \cup \{\alpha_n\}$ or $\Sigma_{n+1} = \Sigma_n \cup \{\neg\alpha_n\}$ then Σ_{n+1} is consistent by construction. If $\Sigma_{n+1} = \Sigma_n \cup \{\alpha_n, \alpha(x := c_n)\}$ then it is consistent by generalisation on constants (for details see [3]).

For the remaining two cases we prove the following lemma:

Lemma 8.2.

(a) If α_n is of the form $\gamma \to (CP\vec{x} \approx q)(\alpha|\beta)$ then there exists $k \in \mathbb{N}$ s.t.
$\Sigma_{n+1} = \Sigma_n \cup \{\neg\alpha_n, \gamma \to \neg(CP\vec{x} \geq q - \frac{1}{k})(\alpha|\beta)\}$ or
$\Sigma_{n+1} = \Sigma_n \cup \{\neg\alpha_n, \gamma \to \neg(CP\vec{x} \leq q + \frac{1}{k})(\alpha|\beta)\}$ is consistent.

(b) If $\vec{x}$ are the free variables of α_n, then there exists $r \in \mathbb{I}$ s.t.
$\Sigma_{n+1} = \Sigma_n \cup \{(P\vec{x} = r)\alpha_n\}$ is consistent

Proof. (a) Assume $\alpha_n = \gamma \to (CP\vec{x} \approx q)(\alpha|\beta)$ and $\Sigma_n \cup \alpha_n$ is inconsistent. Further assume that the sets Σ_{n+1} defined in (a) are inconsistent for all $k \in \mathbb{N}$. Then for every $k \in \mathbb{N}$

$$\Sigma_n, \neg\alpha, \gamma \to \neg(CP\vec{x} \geq q - \frac{1}{k})(\alpha|\beta) \vdash \bot \qquad \text{(inconsistency)}$$

$$\Sigma_n, \neg\alpha, \gamma \to \neg(CP\vec{x} \leq q + \frac{1}{k})(\alpha|\beta) \vdash \bot \qquad \text{(inconsistency)}$$

$$\Sigma_n, \neg\alpha \vdash \gamma \to \neg(CP\vec{x} \geq q - \frac{1}{k})(\alpha|\beta) \qquad \text{(deduction theorem)}$$

$$\Sigma_n, \neg\alpha \vdash \gamma \to \neg(CP\vec{x} \leq q + \frac{1}{k})(\alpha|\beta) \qquad \text{(deduction theorem)}$$

$$\Sigma_n, \neg\alpha \vdash \gamma \to (CP\vec{x} \approx q)(\alpha|\beta) \qquad \text{(Approx)}$$

So basically $\Sigma_n, \neg\alpha \vdash \alpha$ which contradicts the consistency of Σ_n.

(b) Assume that the set Σ_{n+1} as defined in (b) is inconsistent for all $r \in \mathbb{I}$. Then for every $r \in \mathbb{I}$

$$\begin{aligned} \Sigma_n, (P\vec{x} = r)\alpha_n \vdash \bot & \quad \text{(inconsistency)} \\ \Sigma_n \vdash \top \to \neg(P\vec{x} = r)\alpha_n & \quad \text{(deduction theorem)} \\ \Sigma_n \vdash \top \to \bot & \quad \text{(Ran)} \\ \Sigma_n \vdash \bot & \quad \text{(deduction theorem)} \end{aligned}$$

which contradicts the consistency of Σ_n.

□

We can now conclude, that in every step Σ_n is a consistent set of sentences (it contains no formulas with free variables by construction). We define the maximal element given by the chain $(\Sigma_i | i \in \mathbb{N}_0)$ as Σ^* and discuss directly some of its properties:

Lemma 8.3 (Properties of Σ^*). *The Following properties hold:*

(a) Σ^* *is maximal consistent (and thus has all properties of Lemma 7.4)*

(b) There exists exactly one $r \in \mathbb{I}$ *such that* $(P\vec{x} = r)\alpha(\vec{x}) \in \Sigma^*$

(c) There exists exactly one $r \in \mathbb{I}$ *such that* $(CP\vec{x} = r)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$

(d) If $(CP\vec{x} \geq r)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$*, there is* $r' \in \mathbb{I}$ *such that* $r' \geq r$ *and* $(CP\vec{x} = r)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$

(e) If $(CP\vec{x} \leq r)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$*, there is* $r' \in \mathbb{I}$ *such that* $r' \leq r$ *and* $(CP\vec{x} = r)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$

(f) If $(CP\vec{x} \approx q)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$ *and* $q' \in \mathbb{I}_{\mathbb{Q}} \setminus \{q\}$*, then* $(CP\vec{x} \approx q')(\alpha(\vec{x})|\beta(\vec{x})) \notin \Sigma^*$

(g) If $\exists x\alpha(x) \in \Sigma^*$ *then there is a constant* c *s.t.* $\alpha(x := c) \in \Sigma^*$.

Proof. (a) Σ_n is consistent for each step and thus Σ^* is consistent too (see Lindenbaums theorem). Let $\alpha \in \text{Sen}(L)$ such that $\Sigma \cup \alpha$ is consistent. There exists $n \in \mathbb{N}$ with $\alpha = \alpha_n$ and therefore by construction $\alpha \in \Sigma^*$.

(b) Assume there exist $r_1 < r_2 \in \mathbb{I}$ s.t. $(P\vec{x} = r_1)\alpha(\vec{x}) \in \Sigma^*$ and $(P\vec{x} = r_2)\alpha(\vec{x}) \in \Sigma^*$. Then $\Sigma^* \vdash (P\vec{x} \leq r_1)\alpha$ and $\Sigma^* \vdash (P\vec{x} \geq r_2)\alpha$. As $r_1 < r_2$ it follows that $(P\vec{x} \neq r)\alpha$ for all $r \in \mathbb{I}$. We can now use (Ran) and get $\Sigma^* \vdash \bot$ which is a contradiction to the consistency of Σ^*.

(c) analogue to (b)

(d) Assume $(CP\vec{x} \geq r)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$ and there exists no r' s.t. $r' \geq r$ and $(CP\vec{x} = r)(\alpha(\vec{x})|\beta(\vec{x})) \in \Sigma^*$. Then by definition (notation) $(P\vec{x} \neq r)\alpha$ for all $r \in \mathbb{I}$ and using (Ran) we get $\Sigma^* \vdash \bot$.

(e) analogue to (d)

(f) analogue to (b)

(g) Follows directly from the construction.

□

This concludes Step 1 of the completeness-proof.

Step 2:

8.2 Canonical Model

The canonical model $\bar{\mathfrak{A}}$ for Σ^* is defined through equivalence-classes of constants. Our goal is to be able to use $\bar{\mathfrak{A}}$, restricted on the initial language L, as a model for Σ.

Definition 8.4 (Equivalence classes of constants)**.**

- For every constant $c \in C = \{c_n | n \in \mathbb{N}\}$ we define the equivalence class of c as $[c] := \{c_i | c = c_i\}$
- for a tuple $\vec{c} = (c_1, c_2, \dots)$ of constants we define $[\vec{c}] := ([c_1], [c_2], \dots)$.
- for every formula $\alpha(\vec{x})$ we define $\langle\alpha(\vec{x})\rangle := \{[\vec{c}] \in A^n | \alpha(\vec{x} := \vec{c}) \in \Sigma^*\}$.

We choose our universes A^n as the collection of equivalence classes $[\vec{c}]$, where $\vec{c}$ contains n constants.

As a reminder, a structure of $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$ is a tuple $\bar{\mathfrak{A}} = \langle \mathfrak{A}, \mathfrak{F}_n, \mathbf{P}_n \rangle_{n\in\mathbb{N}}$, where $\mathfrak{A}$ is an $L_{\omega\omega}$ structure and $(A^n, \mathfrak{F}_n, \mathbf{P}_n)$ is a finitely additive probability space for all $n \geq 1$. We now define the missing parts for our canonical model:

Definition 8.5 (Canonical model for Σ^*)**.** The canonical model for Σ^* is defined as follows:

- The definition of $\mathfrak{A}$ is standard (see [9]).
- $\mathfrak{F}_n = \{\langle\alpha(\vec{x})\rangle | \alpha(\vec{x})$ has n free variables$\}$
- All values of probability measures $\mathbf{P}_n$ are given by

$$\mathbf{P}_n(\langle\alpha(\vec{x})\rangle) = r \quad \text{iff} \quad (P\vec{x} = r)\alpha(\vec{x}) \in \Sigma^*$$

By choosing $\alpha = \bigwedge_{i\in\mathbf{I}} c_i$ for an index-set $\mathbf{I}$ with cardinality n we assure, that $\mathbf{P}_n$ is defined for all singleton sets. As short notation we will use $\mathbf{P}_{\alpha(\vec{x})} = r$ instead of $\mathbf{P}_n(\langle\alpha(\vec{x})\rangle) = r$

$\bar{\mathfrak{A}}$ defines a model for $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$. The proof that $\mathfrak{F}_n$ really satisfies all demanded properties will be omitted.

Before we begin with the third step we need to state a lemma about some theorems (formulas deducible from the empty set) of our logic $L^{\mathbf{P},\mathbb{I}}_{\omega\omega}$:

Lemma 8.6. *For $\alpha, \beta \in Fm(L)$*

1. $\vdash (CP\vec{x} \geq r_1)(\alpha|\beta) \to (CP\vec{x} \geq r_2)(\alpha|\beta), r_1 > r_2$
2. $\vdash (CP\vec{x} \leq r_1)(\alpha|\beta) \to (CP\vec{x} \leq r_2)(\alpha|\beta), r_1 < r_2$
3. $\vdash (CP\vec{x} = r_1)(\alpha|\beta) \to \neg(CP\vec{x} = r_2)(\alpha|\beta), r_1 \neq r_2$
4. $\vdash (CP\vec{x} = r_1)(\alpha|\beta) \to \neg(CP\vec{x} \geq r_2)(\alpha|\beta), r_1 < r_2$
5. $\vdash (CP\vec{x} = r_1)(\alpha|\beta) \to \neg(CP\vec{x} \leq r_2)(\alpha|\beta), r_1 > r_2$
6. $\vdash (CP\vec{x} = q)(\alpha|\beta) \to (CP\vec{x} \approx q)(\alpha|\beta), q \in \mathbb{I}_{\mathbb{Q}}$
7. $\vdash (CP\vec{x} \approx q_1)(\alpha|\beta) \to \neg(CP\vec{x} \approx q_2)(\alpha|\beta), q_1, q_2 \in \mathbb{I}_{\mathbb{Q}}, q_1 \neq q_2$
8. $\vdash (P\vec{x} = 0)\beta \to \neg(CP\vec{x} \leq r)(\alpha|\beta), r < 1$
9. $\vdash (P\vec{x} \leq 1)\alpha$

The proof of this lemma will be omitted.

Step 3:

8.3 Truth Lemma

We now want to show that the model we defined before really satisfies Σ^*:

Lemma 8.7 (Truth lemma). *For every formula $\alpha(\vec{y})$ and every assignment ν with $\nu(\vec{y}) = [\vec{c}]$*

$$\bar{\mathfrak{A}}, \nu \models \alpha \quad \textit{iff} \quad \alpha(\vec{y} := \vec{c}) \in \Sigma^*$$

Proof. The cases where $\alpha = \neg\varphi, \alpha = \varphi \wedge \psi, \alpha = (\forall x)\varphi$ are covered in [9]. We now consider the case, where $\alpha = (CP\vec{x} \geq r)(\beta(\vec{x}, \vec{y})|\gamma(\vec{x}, \vec{y}))$ for formulas β, γ and variables $\vec{x}, \vec{y}$:
$\Rightarrow$: Assume $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq r)(\beta(\vec{x}, \vec{y})|\gamma(\vec{x}, \vec{y}))$. By definition of the CP quantifier two cases that can occur:

1. $\mathbf{P}_{\gamma(\vec{x},\vec{c})} = 0$ (where $\vec{c}$ are constants)
2. $\mathbf{P}_{\gamma(\vec{x},\vec{c})} = t > 0$ and $\mathbf{P}_{\beta(\vec{x},\vec{c})\wedge\gamma(\vec{x},\vec{c})} = s$ with $s/t \geq r$

For the first case it follows from definition 8.5 that $(P\vec{x} = 0)\gamma(\vec{x}, \vec{c}) \in \Sigma^*$ and therefore (as Σ^* is deductively closed):

$$\begin{aligned} (CP\vec{x} = 1)(\beta(\vec{x}, \vec{c})|\gamma(\vec{x}, \vec{c})) &\in \Sigma^* && \text{(CP6)} \\ \neg(CP\vec{x} \leq r')(\beta(\vec{x}, \vec{c})|\gamma(\vec{x}, \vec{c})) &\in \Sigma^*, r' < 1 && \text{(lemma 8.6)} \\ (CP\vec{x} > r')(\beta(\vec{x}, \vec{c})|\gamma(\vec{x}, \vec{c})) &\in \Sigma^*, r' < 1 \\ (CP\vec{x} \geq r)(\beta(\vec{x}, \vec{c})|\gamma(\vec{x}, \vec{c})) &\in \Sigma^* \text{ for any } r \in \mathbb{I} \end{aligned}$$

The second case implies by definition 8.5 that $(P\vec{x} = t)\gamma(\vec{x}, \vec{c}) \in \Sigma^*$ and $(P\vec{x} = 0)\beta(\vec{x}, \vec{c}) \wedge \gamma(\vec{x}, \vec{c}) \in \Sigma^*$. Then:

$$\begin{aligned} (CP\vec{x} = s/t)(\beta(\vec{x}, \vec{c})|\gamma(\vec{x}, \vec{c})) &\in \Sigma^* && \text{(CP7)} \\ (CP\vec{x} \geq r)(\beta(\vec{x}, \vec{c})|\gamma(\vec{x}, \vec{c})) &\in \Sigma^* \end{aligned}$$

$\Leftarrow$: Suppose $(CP\vec{x} \geq r)(\beta(\vec{x}, \vec{c})|\gamma(\vec{x}, \vec{c})) \in \Sigma^*$. From lemma 8.3 we know, that there exist unique s, t such that

$$\begin{aligned} &(P\vec{x} = s)(\beta(\vec{x}, \vec{c}) \wedge \gamma(\vec{x}, \vec{c})) \in \Sigma^* \\ \text{and } &(P\vec{x} = t)\gamma(\vec{x}, \vec{c}) \in \Sigma^* \end{aligned}$$

Therefore $\mathbf{P}_{\beta(\vec{x},\vec{c})\wedge\gamma(\vec{x},\vec{c})} = s$ and $\mathbf{P}_{\gamma(\vec{x},\vec{c})} = t \geq 0$ and by definition $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq r)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y}))$

The case $\alpha = (CP\vec{x} \leq r)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y}))$ is similar to the one above.

It remains the one where $\alpha = (CP\vec{x} \approx q)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y}))$:
$\Rightarrow$: Assume $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \approx q)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y}))$. Then by definition:

$$\begin{gathered}\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq q - \frac{1}{n})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \\ \text{and } \bar{\mathfrak{A}}, \nu \models (CP\vec{x} \leq q + \frac{1}{m})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \\ \text{for } 0 \leq q - \frac{1}{n} < q < q + \frac{1}{m} \leq 1\end{gathered}$$

Following the first part of this proof (the case for $(CP\vec{x} \geq r)$ and $(CP\vec{x} \leq r)$) we get

$$(CP\vec{x} \geq q - \frac{1}{n})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \in \Sigma^* \tag{1}$$
$$\text{and } (CP\vec{x} \leq q + \frac{1}{m})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \in \Sigma^* \tag{2}$$
$$\text{for } 0 \leq q - \frac{1}{n} < q < q + \frac{1}{m} \leq 1$$

If $(CP\vec{x} \approx q)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \notin \Sigma^*$ then either (1) or (2) would occur in negated form in Σ^* by construction, but that contradicts its consistency. Therefore $(CP\vec{x} \approx q)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \in \Sigma^*$.

$\Leftarrow$: Assume $(CP\vec{x} \approx q)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \in \Sigma^*$.
From lemma 8.3 we get that there exists exactly one $t \in \mathbb{I}$ s.t. $(P\vec{x} = t)\gamma(\vec{x},\vec{c}) \in \Sigma^*$. The usual two cases can occur:

1. $t = 0$, which forces $q = 1$. Using (CP6) and lemma 8.6 it follows immediately that $\bar{\mathfrak{A}}, \nu \models (CP\vec{x} \approx 1)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y}))$

2. $t > 0$ which together with axioms (CP4) and (CP5) implies

$$\begin{gathered}(CP\vec{x} \geq q - \frac{1}{n})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \in \Sigma^* \\ \text{and } (CP\vec{x} \leq q + \frac{1}{m})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \in \Sigma^* \\ \text{for } 0 \leq q - \frac{1}{n} < q < q + \frac{1}{m} \leq 1\end{gathered}$$

Again following the first part of this proof we can conclude with

$$\begin{aligned} & \bar{\mathfrak{A}}, \nu \models (CP\vec{x} \geq q - \frac{1}{n})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \\ \text{and } & \bar{\mathfrak{A}}, \nu \models (CP\vec{x} \leq q + \frac{1}{m})(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \\ & \text{for } 0 \leq q - \frac{1}{n} < q < q + \frac{1}{m} \leq 1 \\ \Rightarrow \quad & \bar{\mathfrak{A}}, \nu \models (CP\vec{x} \approx q)(\beta(\vec{x},\vec{y})|\gamma(\vec{x},\vec{y})) \end{aligned} \tag{Approx}$$

□

We have proven, that for each consistent set of formulas Σ there exists a maximal consistent set Σ^* which has a mode and are finally able to prove the completeness:

Theorem 8.8 (Completeness theorem)**.** *If Σ is a consistent set of formulas, then Σ has a model.*

Proof. If Σ is consistent, then there exists an extending language $L*$ and a maximally consistent set Σ^* containing Σ. By the truth-lemma Σ^* has a model $\bar{\mathfrak{A}}$. The restriction of $\bar{\mathfrak{A}}$ on the original language L is a model for Σ (for details on how to do the restriction see [3]) □

.

Conclusion

Proving the completeness-theorem concludes this thesis. For further results concerning the logic $L^{\mathbf{P};\mathbb{I}}_{\omega\omega}$, including decidable fragments and some applications, I recommend reading the original paper [4].

Future Works

An interesting topic in the future will be to study the combination of justification logics and approximate conditional probabilities. Justification logics are epistemic logics that feature evidences for an agent's knowledge or belief[1, 7, 8]. Recently probabilistic justification logics were introduced by Kokkinis et al. [5, 6], in which different justifications can lead to different degrees of belief.

References

[1] S. Artemov. Explicit provability and constructive semantics. *Symbolic Logic*, (7):1–36, 2001. 34

[2] Heinz Bauer. *Probability theory and elements of measure theory*. Academic Pr, 1981. 3

[3] Herbert B. Enderton. *A Mathematical Introduction to Logic*. Harcourt Academic Press, 2001. 1, 5, 6, 28, 34

[4] Nebojša Ikodinović, Miodrag Rašković, Zoran Marković, and Zoran Ognjanović. A first-order probabilistic logic with approximate conditional probabilities. *Logic Journal of the IGPL*, 22(4):539–564, 2013. 1, 4, 6, 23, 34

[5] Maksimović P. Ognjanović Z. Studer T. Kokkinis, I. First steps towards probabilistic justification logic. *Logic Journal of IGPL*, (23(4)):662–687, 2015. 34

[6] Ognjanović Z. Studer T. Kokkinis, I. Probabilistic justification logic. *Artemov, S., Nerode, A. (eds.) Proc. LFCS 2016, LNCS, vol. 9537*, pages 174–186, Springer (2016). 34

[7] Studer T. Kuznets, R. Justifications, ontology, and conservativity. *Bolander, T., Braüner, T., Ghilardi, S., Moss, L. (eds.) AiML 9*, pages 437–458, College Publications (2012). 34

[8] Studer T. Kuznets, R. Update as evidence: Belief expansion. *Artemov,S., Nerode, A. (eds.) Proc. LFCS 2013, LNCS*, (vol. 7734):266–279, Springer (2013). 34

[9] George Metcalfe. Mathematical logic and model theory. University of Bern, February 2014. 1, 6, 18, 22, 31, 32

[10] Prof. Ilya Molchanov. Kombinatorik und wahrscheinlichkeitsrechnung. University of Bern, 2015. 3

[11] Johanna F. Ziegel. Wahrscheinlichkeitstheorie. University of Bern, Juni 2015. 3

Statement of Originality

I hereby confirm that I have written the accompanying thesis by myself, without contributions from any sources other than those cited in the text and acknowledgements. This bachelor thesis was not used in the same or in a similar version to achieve an academic grading at the University of Bern or elsewhere.

Bern, 16th of February 2016