# A first-order logic for reasoning about higher-order upper and lower probabilities

**N. Savić**[1] D. Doder[2] Z. Ognjanović[3]

[1]Institute of Computer Science
University of Bern

[2]Faculty of Mechanical Engineering
University of Belgrade

[3]Mathematical Institute of SASA
Belgrade

Lugano, July 2017

# Outline

1. Previous work on this topic
   - Halpern and Pucella's Paper
   - Our previous work
2. The Logic $\mathcal{L}_{lu}$
   - Syntax and Semantics
   - Axiomatization and Strong Completeness

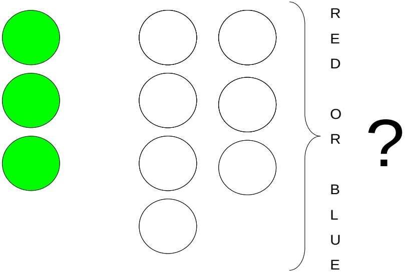
# Outline

1. Previous work on this topic
   - Halpern and Pucella's Paper
   - Our previous work
2. The Logic $\mathcal{L}_{lu}$
   - Syntax and Semantics
   - Axiomatization and Strong Completeness

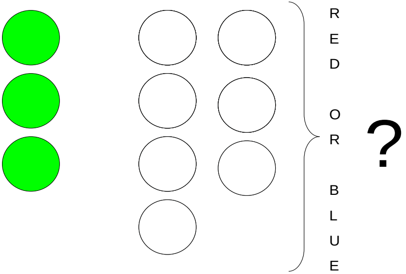
# Halpern and Pucella's Paper

Halpern, J. Y., Pucella, R.: A Logic for Reasoning about Upper Probabilities. Journal of Artificial Intelligence Research, 17: 5781 (2002)

# Example

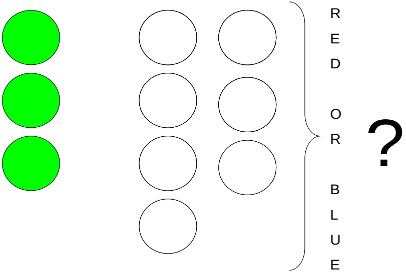


# Example



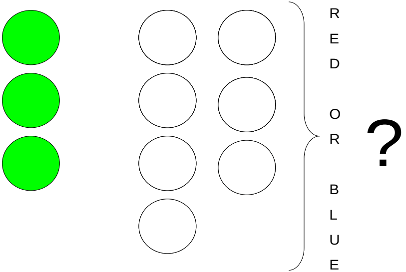

Set of probabilities $P = \{\mu_\alpha \mid \alpha \in [0, 0.7]\}$, where $\mu_\alpha$ gives green-event probability 0.3, blue-event probability $\alpha$, and red-event probability $0.7 - \alpha$.

# Example




Set of probabilities $P = \{\mu_\alpha \mid \alpha \in [0, 0.7]\}$, where $\mu_\alpha$ gives green-event probability 0.3, blue-event probability $\alpha$, and red-event probability $0.7 - \alpha$.
$P_\star(X) = \inf\{\mu(X) \mid \mu \in P\}$

# Example

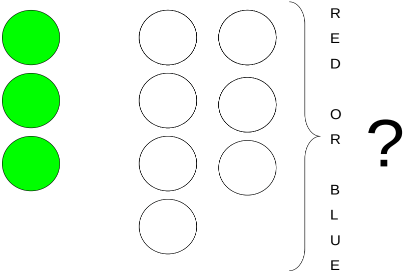



Set of probabilities $P = \{\mu_\alpha \mid \alpha \in [0, 0.7]\}$, where $\mu_\alpha$ gives green-event probability 0.3, blue-event probability $\alpha$, and red-event probability $0.7 - \alpha$.

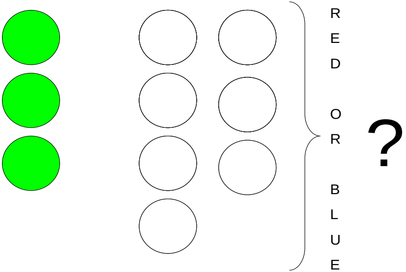
$P_\star(X) = \inf\{\mu(X) \mid \mu \in P\}$

$P^\star(X) = \sup\{\mu(X) \mid \mu \in P\}$

# Example

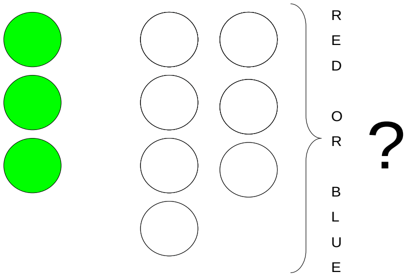


# Example



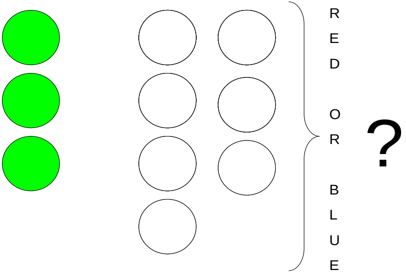

$P_{\star}(R) = 0,$

# Example



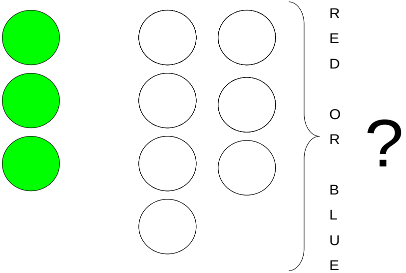

$P_\star(R) = 0, \quad P^\star(R) = 0.7,$

# Example



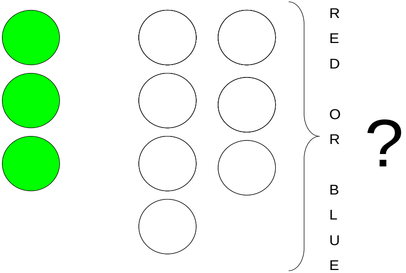

$P_\star(R) = 0, \quad P^\star(R) = 0.7, \quad P_\star(B) = 0,$

# Example




$$P_\star(R) = 0, \quad P^\star(R) = 0.7, \quad P_\star(B) = 0, \quad P^\star(B) = 0.7,$$

# Example




$P_\star(R) = 0, \quad P^\star(R) = 0.7, \quad P_\star(B) = 0, \quad P^\star(B) = 0.7,$
$P_\star(G) = P^\star(G) = 0.3.$

Those two functions are related by the formula $P_\star(X) = 1 - P^\star(X^c)$.

Those two functions are related by the formula $P_\star(X) = 1 - P^\star(X^c)$.
A basic likelihood formulas:

$$\theta_1 l(\varphi_1) + \cdots + \theta_k l(\varphi_k) \geq c,$$

where $c, \theta_i \in \mathbb{R}$, $\varphi_i$ are propositional formulas $i = 1, \ldots, k$.
$l$ is an upper probability operator

## Theorem (Anger and Lembcke 1985)

*Let $W$ be a set, $H$ an algebra of subsets of $W$, and $f$ a function $f : H \longrightarrow [0, 1]$. There exists a set $P$ of probability measures such that $f = P^\star$ iff $f$ satisfies the following three properties:*

(1) $f(\emptyset) = 0$,

(2) $f(W) = 1$,

(3) *for all natural numbers $m, n, k$ and all subsets $A_1, \ldots, A_m$ in $H$, if the multiset $\{\{A_1, \ldots, A_m\}\}$ is an $(n, k)$-cover of $(A, W)$, then $k + nf(A) \leq \sum_{i=1}^{m} f(A_i)$.*

## Theorem (Anger and Lembcke 1985)

*Let $W$ be a set, $H$ an algebra of subsets of $W$, and $f$ a function $f : H \longrightarrow [0, 1]$. There exists a set $P$ of probability measures such that $f = P^{\star}$ iff $f$ satisfies the following three properties:*

(1) $f(\emptyset) = 0$,

(2) $f(W) = 1$,

(3) *for all natural numbers $m, n, k$ and all subsets $A_1, \ldots, A_m$ in $H$, if the multiset $\{\{A_1, \ldots, A_m\}\}$ is an $(n, k)$-cover of $(A, W)$, then $k + nf(A) \leq \sum_{i=1}^{m} f(A_i)$.*

## Definition ($(n, k)$-cover)

A set $A$ is said to be covered $n$ times by a multiset $\{\{A_1, \ldots, A_m\}\}$ of sets if every element of $A$ appears in at least $n$ sets from $A_1, \ldots, A_m$, i.e., for all $x \in A$, there exists $i_1, \ldots, i_n$ in $\{1, \ldots, m\}$ such that for all $j \leq n$, $x \in A_{i_j}$. An $(n, k)$-cover of $(A, W)$ is a multiset $\{\{A_1, \ldots, A_m\}\}$ that covers $W$ $k$ times and covers $A$ $n + k$ times.

# Outline

1. Previous work on this topic
   - Halpern and Pucella's Paper
   - Our previous work
2. The Logic $\mathcal{L}_{lu}$
   - Syntax and Semantics
   - Axiomatization and Strong Completeness

# Our previous work

Savić, N., Doder, D., Ognjanović, Z.: A logic with Upper and Lower Probability Operators. In Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications, 267–276, Pescara, Italy (2015)

# Our previous work

Savić, N., Doder, D., Ognjanović, Z.: A logic with Upper and Lower Probability Operators. In Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications, 267–276, Pescara, Italy (2015)

Instead of using linear combinations...

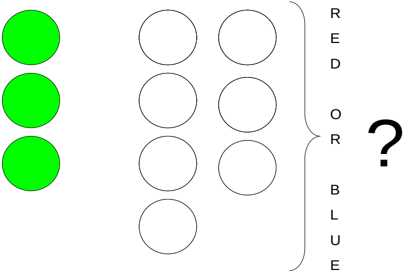
# Our previous work

Savić, N., Doder, D., Ognjanović, Z.: A logic with Upper and Lower Probability Operators. In Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications, 267–276, Pescara, Italy (2015)

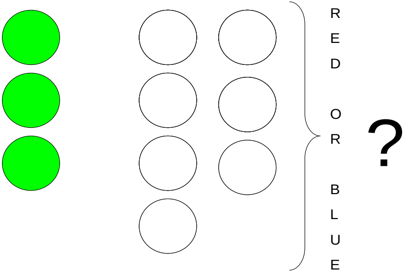
Instead of using linear combinations...

Classical propostional logic + operators $L_{\geq s}$ and $U_{\geq s}$, $s \in \mathbb{Q} \cap [0,1]$.

# Example

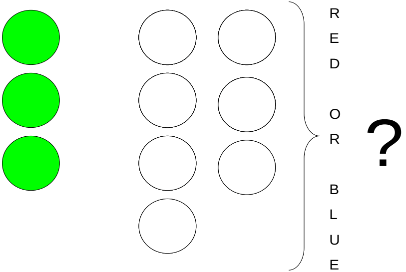


# Example



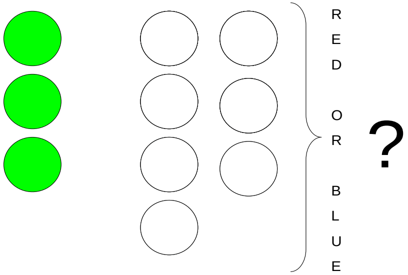

$L_{=0}R, L_{=0}B;$

# Example




$$L_{=0}R, L_{=0}B; \qquad U_{=0.7}R, U_{=0.7}B$$

# Example




$$L_{=0}R, L_{=0}B; \qquad U_{=0.7}R, U_{=0.7}B$$

$$((U_{\leq 0.3}G \wedge L_{\geq 0.3}G) \wedge U_{\leq 0.2}R) \Rightarrow L_{\geq 0.5}B.$$

# Semantics

# Semantics

**Definition ($LUPP$-structure)**

$M = \langle W, H, P, \upsilon \rangle$, where:

- $W$ is a nonempty set of *worlds*.
- $H$ is an algebra of subsets of $W$.
- $P$ is a set of finitely additive probability measures defined on $H$.
- $\upsilon : W \times \mathcal{L} \longrightarrow \{true, false\}$ evaluations of the primitive propositions.

# Semantics

**Definition ($LUPP$-structure)**

$M = \langle W, H, P, v \rangle$, where:

- $W$ is a nonempty set of *worlds*.
- $H$ is an algebra of subsets of $W$.
- $P$ is a set of finitely additive probability measures defined on $H$.
- $v : W \times \mathcal{L} \longrightarrow \{true, false\}$ evaluations of the primitive propositions.

**Definition (Satisfiability relation)**

- $M \models \alpha$ iff $v(w)(\alpha) = true$, for all $w \in W$,
- $M \models U_{\geq s}\alpha$ iff $P^{\star}([\alpha]) \geq s$,
- $M \models L_{\geq s}\alpha$ iff $P_{\star}([\alpha]) \geq s$.

# Axiom schemes

(1) all instances of the classical propositional tautologies

(2) $U_{\leq 1}\alpha \wedge L_{\leq 1}\alpha$

(3) $U_{\leq r}\alpha \rightarrow U_{<s}\alpha$, $s > r$

(4) $U_{<s}\alpha \rightarrow U_{\leq s}\alpha$

(5) $(U_{\leq r_1}\alpha_1 \wedge \cdots \wedge U_{\leq r_m}\alpha_m) \rightarrow U_{\leq r}\alpha$, if $\alpha \rightarrow \bigvee_{J\subseteq\{1,\ldots,m\},|J|=k+n} \bigwedge_{j\in J} \alpha_j$ and $\bigvee_{J\subseteq\{1,\ldots,m\},|J|=k} \bigwedge_{j\in J} \alpha_j$ are propositional tautologies, where $r = \frac{\sum_{i=1}^{m} r_i - k}{n}$, $n \neq 0$

(6) $\neg(U_{\leq r_1}\alpha_1 \wedge \cdots \wedge U_{\leq r_m}\alpha_m)$, if $\bigvee_{J\subseteq\{1,\ldots,m\},|J|=k} \bigwedge_{j\in J} \alpha_j$ is a propositional tautology and $\sum_{i=1}^{m} r_i < k$

(7) $L_{=1}(\alpha \rightarrow \beta) \rightarrow (U_{\geq s}\alpha \rightarrow U_{\geq s}\beta)$

# Inference Rules

(1) From $\rho$ and $\rho \to \sigma$ infer $\sigma$

(2) From $\alpha$ infer $L_{\geq 1}\alpha$

(3) From the set of premises

$$\{\phi \to U_{\geq s - \frac{1}{k}}\alpha \mid k \geq \frac{1}{s}\}$$

infer $\phi \to U_{\geq s}\alpha$

(4) From the set of premises

$$\{\phi \to L_{\geq s - \frac{1}{k}}\alpha \mid k \geq \frac{1}{s}\}$$

infer $\phi \to L_{\geq s}\alpha$.

# Comparison of these two logics

# Comparison of these two logics

Halpern and Pucella's Logic for Reasoning about Upper Probabilities

- Uncountable Language
- Finitary axiomatization
- (Weak) completeness

Our Logic with Upper and Lower Probability Operators

- Countable Language
- Infinitary axiomatization
- Strong completeness

# Outline

# The Idea

Construct the logic that will have the language powerful enough to express:

- $U^a_{\leq \frac{9}{10}} Rain(C)$

# The Idea

Construct the logic that will have the language powerful enough to express:

- $U^{a}_{\leq \frac{9}{10}} Rain(C)$
- $L^{a}_{\geq 0.1} U^{b}_{\leq 0.9} Rain(C)$

# The Idea

Construct the logic that will have the language powerful enough to express:

- $U^a_{\leq \frac{9}{10}} Rain(C)$
- $L^a_{\geq 0.1} U^b_{\leq 0.9} Rain(C)$
- $L^a_{\geq \frac{1}{3}} (\forall x) Rain(x)$

# The Idea

Construct the logic that will have the language powerful enough to express:

- $U^{a}_{\leq \frac{9}{10}} Rain(C)$
- $L^{a}_{\geq 0.1} U^{b}_{\leq 0.9} Rain(C)$
- $L^{a}_{\geq \frac{1}{3}} (\forall x) Rain(x)$
- $(\exists x) U^{a}_{=0} Rain(x)$.

Let $S = \mathbb{Q} \cap [0,1]$, $Var = \{x, y, z, \dots\}$ be a denumerable set of variables and let $\Sigma = \{a, b, \dots\}$ be a finite, non-empty set of agents. The language of the logic $\mathcal{L}_{lu}$ consists of:

- the elements of set $Var$,
- classical propositional connectives $\neg$ and $\wedge$,
- universal quantifier $\forall$,
- for every integer $k \geq 0$, denumerably many function symbols $F_0^k, F_1^k, \dots$ of arity $k$,
- for every integer $k \geq 0$, denumerably many relation symbols $P_0^k, P_1^k, \dots$ of arity $k$,
- the list of upper probability operators $U^a_{\geq s}$, for every $s \in S$,
- the list of lower probability operators $L^a_{\geq s}$, for every $s \in S$,
- comma, parentheses.

## Definition (Formula)

The set $For_{\mathcal{L}_{lu}}$ of formulas is the smallest set containing atomic formulas and that is closed under following formation rules: if $\alpha, \beta$ are formulas, then $L^{a}_{\geq s}\alpha$, $U^{a}_{\geq s}\alpha$, $\neg\alpha$, $\alpha \wedge \beta$, $(\forall x)\alpha$ are formulas as well. The formulas from $For_{\mathcal{L}_{lu}}$ will be denoted by $\alpha, \beta, \ldots$

We use the following abbreviations to introduce other types of inequalities:

- $U^{a}_{<s}\alpha$ is $\neg U^{a}_{\geq s}\alpha$, $U^{a}_{\leq s}\alpha$ is $L^{a}_{\geq 1-s}\neg\alpha$, $U^{a}_{=s}\alpha$ is $U^{a}_{\leq s}\alpha \wedge U^{a}_{\geq s}\alpha$, $U^{a}_{>s}\alpha$ is $\neg U^{a}_{\leq s}\alpha$,
- $L^{a}_{<s}\alpha$ is $\neg L^{a}_{\geq s}\alpha$, $L^{a}_{\leq s}\alpha$ is $U^{a}_{\geq 1-s}\neg\alpha$, $L^{a}_{=s}\alpha$ is $L^{a}_{\leq s}\alpha \wedge L^{a}_{\geq s}\alpha$, $L^{a}_{>s}\alpha$ is $\neg L^{a}_{\leq s}\alpha$.

## Definition ($\mathcal{L}_{lu}$-structure)

An $\mathcal{L}_{lu}$-structure is a tuple $\mathcal{M} = \langle W, D, I, LUP \rangle$, where:

- $W$ is a nonempty set of *worlds*,
- $D$ associates a non-empty domain $D(w)$ with every world $w \in W$,
- $I$ associates an interpretation $I(w)$ with every world $w \in W$ such that:
  - $I(w)(F_i^k) : D(w)^k \to D(w)$, for all $i$ and $k$,
  - $I(w)(P_i^k) \subseteq D(w)^k$, for all $i$ and $k$,
- $LUP$ assigns, to every $w \in W$ and every agent $a \in \Sigma$, a space, such that $LUP(a, w) = \langle W(a, w), H(a, w), P(a, w) \rangle$, where:
  - $\emptyset \neq W(a, w) \subseteq W$,
  - $H(a, w)$ is an algebra of subsets of $W(a, w)$, i.e. a set of subsets of $W(a, w)$ such that:
    - $W(a, w) \in H(a, w)$,
    - if $A, B \in H(a, w)$, then $W(a, w) \setminus A \in H(a, w)$ and $A \cup B \in H(a, w)$,
  - $P(a, w)$ is a set of finitely additive probability measures defined on $H(a, w)$

## Definition

The truth value of a formula $\alpha$ in a world $w \in W$:

- if $\alpha = P_i^m(t_1, \ldots, t_m)$, then $I(w)(\alpha)_\upsilon = true$ if $\langle I(w)(t_1)_\upsilon, \ldots, I(w)(t_m)_\upsilon \rangle \in I(w)(P_i^m)$, otherwise $I(w)(\alpha)_\upsilon = false$,
- if $\alpha = \neg\beta$, then $I(w)(\alpha)_\upsilon = true$ if $I(w)(\beta)_\upsilon = false$, otherwise $I(w)(\alpha)_\upsilon = false$,
- if $\alpha = \beta \wedge \gamma$, then $I(w)(\alpha)_\upsilon = true$ if $I(w)(\beta)_\upsilon = true$ and $I(w)(\gamma)_\upsilon = true$,
- if $\alpha = U^a_{\geq s}\beta$, then $I(w)(\alpha)_\upsilon = true$ if $P^\star(w, a)\{u \in W(w, a) \mid I(u)(\beta)_\upsilon = true\} \geq s$, otherwise $I(w)(\alpha)_\upsilon = false$,
- if $\alpha = L^a_{\geq s}\beta$, then $I(w)(\alpha)_\upsilon = true$ if $P_\star(w, a)\{u \in W(w, a) \mid I(u)(\beta)_\upsilon = true\} \geq s$, otherwise $I(w)(\alpha)_\upsilon = false$,
- if $\alpha = (\forall x)\beta$, then $I(w)(\alpha)_\upsilon = true$ if for every $d \in D(w)$, $I(w)(\beta)_{\upsilon_w[d/x]} = true$, otherwise $I(w)(\alpha)_\upsilon = false$.

# Constraints

We will consider a class of $\mathcal{L}_{lu}$ models that satisfy:

- all the worlds from a model have the same domain, i.e., for all $v, w \in W$, $D(v) = D(w)$,
- for every sentence $\alpha$, for every agent $a \in \Sigma$ and every world $w$ from a model $\mathcal{M}$, the set $\{u \in W(w, a) \mid I(u)(\alpha)_v = \mathit{true}\}$ of all worlds from $W(w, a)$ that satisfy $\alpha$ is measurable,
- the terms are rigid, i.e., for every model their meanings are the same in all the worlds.

# Outline

1. Previous work on this topic
   - Halpern and Pucella's Paper
   - Our previous work

2. The Logic $\mathcal{L}_{lu}$
   - Syntax and Semantics
   - Axiomatization and Strong Completeness

## Axiom schemes

(1) all instances of the classical propositional tautologies

(2) $(\forall x)(\alpha \to \beta) \to (\alpha \to (\forall x)\beta)$, where the variable $x$ does not occur free in $\alpha$

(3) $(\forall x)\alpha(x) \to \alpha(t)$, where $\alpha(t)$ is obtained by substitution of all free occurrences of $x$ in the first-order formula $\alpha(x)$ by the term $t$ which is free for $x$ in $\alpha(x)$

(4) $U^a_{\leq 1}\alpha \wedge L^a_{\leq 1}\alpha$

(5) $U^a_{\leq r}\alpha \to U^a_{< s}\alpha$, $s > r$

(6) $U^a_{< s}\alpha \to U^a_{\leq s}\alpha$

(7) $(U^a_{\leq r_1}\alpha_1 \wedge \cdots \wedge U^a_{\leq r_m}\alpha_m) \to U^a_{\leq r}\alpha$, if $\alpha \to \bigvee_{J\subseteq\{1,\dots,m\},|J|=k+n} \bigwedge_{j\in J}\alpha_j$ and $\bigvee_{J\subseteq\{1,\dots,m\},|J|=k} \bigwedge_{j\in J}\alpha_j$ are tautologies, where $r = \frac{\sum_{i=1}^m r_i - k}{n}$, $n \neq 0$

(8) $\neg(U^a_{\leq r_1}\alpha_1 \wedge \cdots \wedge U^a_{\leq r_m}\alpha_m)$, if $\bigvee_{J\subseteq\{1,\dots,m\},|J|=k} \bigwedge_{j\in J}\alpha_j$ is a tautology and $\sum_{i=1}^m r_i < k$

(9) $L^a_{=1}(\alpha \to \beta) \to (U^a_{\geq s}\alpha \to U^a_{\geq s}\beta)$

## Inference Rules

(1) From $\alpha$ and $\alpha \to \beta$ infer $\beta$

(2) From $\alpha$ infer $(\forall x)\alpha$

(3) From $\alpha$ infer $L^{a}_{\geq 1}\alpha$

(4) From the set of premises

$$\{\alpha \to U^{a}_{\geq s-\frac{1}{k}}\beta \mid k \geq \frac{1}{s}\}$$

infer $\alpha \to U^{a}_{\geq s}\beta$

(5) From the set of premises

$$\{\alpha \to L^{a}_{\geq s-\frac{1}{k}}\beta \mid k \geq \frac{1}{s}\}$$

infer $\alpha \to L^{a}_{\geq s}\beta$.

## Definition (Canonical model)

A canonical model $\mathcal{M}_{Can} = \langle W, D, I, LUP \rangle$ is a tuple such that:

- $W$ is the set of all saturated sets of formulas,
- $D$ is the set of all variable-free terms,
- for every $w \in W$, $I(w)$ is an interpretation such that:
  - for every function symbol $F_i^m$, $I(w)(F_i^m) : D^m \to D$ such that for all variable-free terms $t_1, \ldots, t_m$,
    $I(w)(F_i^m) : \langle t_1, \ldots, t_m \rangle \mapsto F_i^m(t_1, \ldots, t_m)$,
  - for every relation symbol $P_i^m$,
    $I(w)(P_i^m) = \{\langle t_1, \ldots, t_m \rangle \mid P_i^m(t_1, \ldots, t_m) \in w\}$, for all variable-free terms $t_1, \ldots, t_m$,
- for every $a \in \Sigma$ and every $w \in W$,
  $LUP(w, a) = \langle W(w, a), H(w, a), P(w, a) \rangle$ is defined:
  - $W(w, a) = W$,
  - $H(w, a) = \{\{u \mid u \in W(w, a), \alpha \in u\} \mid \alpha \in For_{\mathcal{L}_{lu}}\}$,
  - $P(w, a)$ is any set of probability measures such that
    $P^\star(w, a)(\{u \mid u \in W(w, a), \alpha \in u\}) = \sup\{s \mid U^a_{\geq s}\alpha \in w\}$.

### Definition

Set of formulas $T$ is *saturated* if it is maximally consistent and satisfies: if $\neg(\forall x)\alpha(x) \in T$, then for some term $t$, $\neg\alpha(t) \in T$.

## Definition

Set of formulas $T$ is *saturated* if it is maximally consistent and satisfies:
if $\neg(\forall x)\alpha(x) \in T$, then for some term $t$, $\neg\alpha(t) \in T$.

## Theorem (Lindenbaum's theorem)

*Every consistent set of formulas can be extended to a saturated set.*

**Definition**

Set of formulas $T$ is *saturated* if it is maximally consistent and satisfies: if $\neg(\forall x)\alpha(x) \in T$, then for some term $t$, $\neg\alpha(t) \in T$.

**Theorem (Lindenbaum's theorem)**

*Every consistent set of formulas can be extended to a saturated set.*

**Theorem (Strong completeness)**

*Every consistent set of formulas $T$ is satisfiable.*

# Bibliography


Anger, B., Lembcke, J.: Infnitely subadditive capacities as upper envelopes of measures. Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete, 68: 403–414. (1985)

Halpern, J. Y., Pucella, R.: A Logic for Reasoning about Upper Probabilities. Journal of Artificial Intelligence Research, 17: 57–81 (2002)

Savić, N., Doder, D., Ognjanović, Z.: A logic with Upper and Lower Probability Operators. In Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications, 267–276, Pescara, Italy (2015)