Uncertainty, vagueness and probability of many-valued events

Sara Ugolini

University of Pisa, Department of Computer Science

Santa Fe - May 12, 2017

1. Why fuzzy logic?
2. Uncertainty vs vagueness
3. Probability of many-valued events

Motivation

Classical logic is the main tool for formalizing reasoning, but

- its expressive power is not enough to formalize many facets of commonsense reasoning;
- there is a need to cope with different forms of imperfect information: partial, uncertain, imprecise, vague, etc.

Mathematical Fuzzy Logic

[Hájek, 1998]

- formal systems (syntax, semantics, complete axiomatizations, proof theory, etc...)
- $[0, 1]$: usual choice of truth-value set
- truth-functionality assumption
- logics of comparative truth: $truth(\phi \to \psi) = 1$ iff $truth(\phi) \leq truth(\psi)$
- generalizations of classical logic

Mathematical Fuzzy Logic

Hàjek's idea: to base the semantics on the truth function for conjunction:

A t-norm is a binary operation $*$ on $[0, 1]$ such that:

(i) $*$ is commutative and associative,

(ii) $*$ is non-decreasing in both arguments,

(iii) $1 * x = x$ and $0 * x = 0$ for all $x \in [0, 1]$.

The choice of the t-norm determines the whole calculus, indeed the truth function of implication is the residuum of the t-norm:

$$x \to y = sup\{z : x * z \leq y\}$$

(if the t-norm is continuous, such sup exists and it is unique)

Three main logics:

- Łukasiewicz logic Ł ['20s - '30s]
  - $x *_{Ł} y = max(0, x + y - 1)$
  - $x \rightarrow_{Ł} y = min(1, 1 - x + y)$
  - $\neg_{Ł} x = x \rightarrow_{Ł} 0 = 1 - x$
- Gödel logic G [1930 Heyting, 1933 Gödel, 1959 Dummett]
  - $x *_{\mathrm{G}} y = min(x, y)$
  - $x \rightarrow_{G} y = 1$ if $x \leq y$, or $x \rightarrow_{G} y = y$ otherwise
  - $\neg_{G} x = 1$ if $x = 0$, or $\neg_{G} x = 0$ otherwise
- Product logic Π [Esteva, Godo, Hájek 1996]
  - $x *_{\mathrm{G}} y = x \cdot y$
  - $x \rightarrow_{\Pi} y = 1$ if $x \leq y$, or $x \rightarrow_{\Pi} y = y/x$ otherwise
  - $\neg_{\Pi} x = 1$ if $x = 0$, or $\neg_{\Pi} x = 0$ otherwise

Why Ł, G and Π?

- Hájek's framework is well-established and deeply studied. Between fuzzy logics given by continuous t-norms, Ł, G and Π are fundamental: any other such logic is a combination of them.

- They enjoy interesting and useful properties. For example, the algebra on $[0,1]$ is standard:

  the algebra of formulas with n variables corresponds exactly to the algebra of $[0,1]$- valued functions with domain $[0,1]^n$ and operations defined componentwise by standard ones.

$$\phi \leftrightsquigarrow f_\phi$$

with ϕ formula of n variables, $f_\phi : [0,1]^n \to [0,1]$.

Uncertainty vs vagueness

We said we wanted to deal with imperfect information.

Uncertainty vs vagueness

We said we wanted to deal with imperfect information. This can lead to:

- VAGUENESS
- UNCERTAINTY

Uncertainty vs vagueness

We said we wanted to deal with imperfect information. This can lead to:

- VAGUENESS $\Longrightarrow$ *MANY-VALUED LOGIC*
- UNCERTAINTY $\Longrightarrow$ *PROBABILITY*

Uncertainty vs vagueness

We said we wanted to deal with imperfect information. This can lead to:

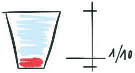
- VAGUENESS $\Longrightarrow$ *MANY-VALUED LOGIC*
- UNCERTAINTY $\Longrightarrow$ *PROBABILITY*

In particular:

- Many-valued logics deal with **vague** concepts and they use intermediate truth values,
- Probability deals with events that are **uncertain** now, but that will become true or false later, and it uses degrees of belief.

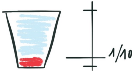
Uncertainty vs vagueness

- Think of a drink that is poisonous with truth-degree 0.1 or a drink with probability 1/10 to be poisonous.



Uncertainty vs vagueness

- Think of a drink that is poisonous with truth-degree 0.1 or a drink with probability 1/10 to be poisonous.



- Fuzzy logics are truth functional:

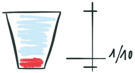
$$truth(A \,\&\, B) = truth(A) \,\&\, truth(B)$$

while probability is not:

$$Prob(A \,\&\, B) \neq Prob(A) \,\&\, Prob(B)$$

Uncertainty vs vagueness

- Think of a drink that is poisonous with truth-degree 0.1 or a drink with probability 1/10 to be poisonous.



- Fuzzy logics are truth functional:

$$truth(A \,\&\, B) = truth(A) \,\&\, truth(B)$$

while probability is not:

$$Prob(A \,\&\, B) \neq Prob(A) \,??\, Prob(B)$$

Probability of vague events

A connection: what does it mean to speak about probability of many-valued events?

Will there be traffic?, is it going to be cold tonight?

Anytime we make a common-life decision we are truly betting on a many-valued event.

Probability of vague events

A connection: what does it mean to speak about probability of many-valued events?

Will there be traffic?, is it going to be cold tonight?

Anytime we make a common-life decision we are truly betting on a many-valued event.

(CLASSICAL) PROBABILITY THEORY $\Longrightarrow$ STATE THEORY

Classical probability functions

Let X be a nonempty set of events. Let $\mathcal{B}$ a collection of subsets of X, closed by intersection, union and complement, containing $\emptyset$ and X (i.e. a *Boolean algebra*).

Classical probability functions

Let X be a nonempty set of events. Let $\mathcal{B}$ a collection of subsets of X, closed by intersection, union and complement, containing $\emptyset$ and X (i.e. a *Boolean algebra*).

A finitely additive probability is a function $P : \mathcal{B} \to [0, 1]$ such that:

(i) If $A, B \in \mathcal{B}$, where $A \cap B = \emptyset$, then

$$P(A \cup B) = P(A) + P(B),$$

(ii) $P(\emptyset) = 0$ and $P(X) = 1$.

Probability of many-valued events: states

Let X a set of many-valued events, and $\mathbf{A}$ be a many-valued structure on X (for example, the equivalent for Łukasiewicz logic of a Boolean algebra).

Probability of many-valued events: states

Let X a set of many-valued events, and $\mathbf{A}$ be a many-valued structure on X (for example, the equivalent for Łukasiewicz logic of a Boolean algebra).

A state is a map $s : A \to [0, 1]$ such that:

(i) For every $a, b \in A$, if $a *_{Ł} b = 0$, then

$$s(a +_{Ł} b) = s(a) + s(b),$$

(ii) $s(1) = 1$.

Probability of many-valued events: states

Let X a set of many-valued events, and $\mathbf{A}$ be a many-valued structure on X (for example, the equivalent for Łukasiewicz logic of a Boolean algebra).

A state is a map $s : A \to [0, 1]$ such that:

(i) For every $a, b \in A$, if $a *_{Ł} b = 0$, then

$$s(a +_{Ł} b) = s(a) + s(b),$$

(ii) $s(1) = 1$.

The condition means additivity with respect to Łukasiewicz sum $+_{Ł}$.

Thus states can be thought of as generalizations of finitely additive probabilities.

States: a developing theory

- States of **Łukasiewicz** logic
  D. Mundici, *Averaging the Truth-value in Łukasiewicz Logic*. Studia Logica 55(1), **1995**.

- States of **Gödel** logic
  S. Aguzzoli, B. Gerla, V. Marra, *Defuzzifying formulas in Gödel logic through finitely additive measures*. Proceedings FUZZ-IEEE, **2008**.

- States of **product** logic
  L. Godo, T. Flaminio, S. Ugolini *States of free product algebras and their integral representation*, to appear

States: why are they relevant?

- States of Ł, G and Π are connected to regular Borel probability measures. This allows to regard them as expected values of bounded random variables.

States: why are they relevant?

- States of Ł, G and Π are connected to regular Borel probability measures. This allows to regard them as expected values of bounded random variables.
- States can be regarded as operators averaging the truth value of Ł, G, Π logics.

States: why are they relevant?

- States of Ł, G and Π are connected to regular Borel probability measures. This allows to regard them as expected values of bounded random variables.
- States can be regarded as operators averaging the truth value of Ł, G, Π logics.
- States characterize the *coherence* criterion of de Finetti's foundation of subjective probability wrt many-valued events. In this sense, states are subjective probability measures.

Integral representation of states

Let $\mathbf{A}$ be the algebra of formulas of n variables of Ł, G or Π respectively.

A map $s : A \to [0,1]$ is a state iff there is a unique (regular Borel) probability measure μ over $[0,1]^n$ such that, for every $f_\phi \in A$,

$$s(f_\phi) = \int_{[0,1]^n} f_\phi \, \mathrm{d}\mu.$$

Łukasiewicz: Kroupa (2005) - Panti (2009)
Gödel: Aguzzoli, Gerla, Marra (2008)
Product: Godo, Flaminio, U. (2017)

Expected value

Let X be a finite set and let $A = [0,1]^X$ be the algebra of Łukasiewicz functions from X in $[0,1]$.

Every $f \in A$ can be regarded as a real-valued and bounded random variable on X.

Expected value

Let X be a finite set and let $A = [0,1]^X$ be the algebra of Łukasiewicz functions from X in $[0,1]$.

Every $f \in A$ can be regarded as a real-valued and bounded random variable on X.

Thus, via the integral representation, states can be seen as expected values of f, indeed:

$$E(f) \;=\; \int_X f_\phi \, \mathrm{d}\mu \;=\; s_\mu(f)$$

Averaging the truth value

The integral representation allows us to associate a real value to each formula of the logic:

$$\phi \;\mapsto\; s(f_\phi) \;=\; \int_{[0,1]^n} f_\phi \,\mathrm{d}\mu_s \;\in \mathbb{R}$$

Averaging the truth value

The integral representation allows us to associate a real value to each formula of the logic:

$$\phi \;\mapsto\; s(f_\phi) \;=\; \int_{[0,1]^n} f_\phi \, \mathrm{d}\mu_s \;\in \mathbb{R}$$

(it is reasonable to consider the integral operator as an averaging process)

Averaging the truth value

The integral representation allows us to associate a real value to each formula of the logic:

$$\phi \;\mapsto\; s(f_\phi) \;=\; \int_{[0,1]^n} f_\phi \, \mathrm{d}\mu_s \quad \in \mathbb{R}$$

(it is reasonable to consider the integral operator as an averaging process)

Moreover, for all Ł, G and Π, we can prove that each possible state belongs to the convex closure of the **valuations of the logic**.

de Finetti's foundation of subjective probability

In the 1930s, de Finetti develops his foundation of subjective probability, in alternative to, for instance, the frequentist approach.

de Finetti's foundation of subjective probability

In the 1930s, de Finetti develops his foundation of subjective probability, in alternative to, for instance, the frequentist approach.

Example:
Suppose that someone wants to build a bridge connecting Reggio Calabria and Messina. Which is the probability that the bridge resists for 200 years?

de Finetti's foundation of subjective probability

In the 1930s, de Finetti develops his foundation of subjective probability, in alternative to, for instance, the frequentist approach.

Example:
Suppose that someone wants to build a bridge connecting Reggio Calabria and Messina. Which is the probability that the bridge resists for 200 years?

Frequentist answer: build a huge number of bridges, wait for 200 years and compute the ratio between the number of bridges which resisted and the total number of bridges.

de Finetti's foundation of subjective probability: coherent betting odds

Events of interest: $e_1 \dots e_k$.

Bookmaker publishes a book β assigning a betting odd $\beta_i \in [0, 1]$ to each e_i.

de Finetti's foundation of subjective probability: coherent betting odds

Events of interest: $e_1 \dots e_k$.

Bookmaker publishes a book β assigning a betting odd $\beta_i \in [0,1]$ to each e_i.

Gambler places stakes $\sigma_1, \dots, \sigma_k \in \mathbb{R}$, for each e_i, and pays to the bookmaker the amount of $\sum_{i=1}^{k} \sigma_i \cdot \beta_i$.

de Finetti's foundation of subjective probability: coherent betting odds

Events of interest: $e_1 \dots e_k$.

Bookmaker publishes a book β assigning a betting odd $\beta_i \in [0,1]$ to each e_i.

Gambler places stakes $\sigma_1, \dots, \sigma_k \in \mathbb{R}$, for each e_i, and pays to the bookmaker the amount of $\sum_{i=1}^{k} \sigma_i \cdot \beta_i$.

Once a future world w is reached, every e_i is either true or false.

de Finetti's foundation of subjective probability: coherent betting odds

Events of interest: $e_1 \dots e_k$.

Bookmaker publishes a book β assigning a betting odd $\beta_i \in [0,1]$ to each e_i.

Gambler places stakes $\sigma_1, \dots, \sigma_k \in \mathbb{R}$, for each e_i, and pays to the bookmaker the amount of $\sum_{i=1}^{k} \sigma_i \cdot \beta_i$.

Once a future world w is reached, every e_i is either true or false.

Bookmaker pays back to the gambler σ_i euros if e_i turns out to be true in w, or nothing if it is false in w.
Total balance: $\sum_{i=1}^{k} \sigma_i(\beta_i - w(e_i))$.

de Finetti's foundation of subjective probability: coherent betting odds

- It is possible for the gambler to pay a negative amount σ_i on e_i. This assumption is called reversibility.

de Finetti's foundation of subjective probability: coherent betting odds

- It is possible for the gambler to pay a negative amount σ_i on e_i. This assumption is called reversibility.
- The book β is called coherent if it does not ensure bookmaker to incur a sure loss,

de Finetti's foundation of subjective probability: coherent betting odds

- It is possible for the gambler to pay a negative amount σ_i on e_i. This assumption is called reversibility.
- The book β is called coherent if it does not ensure bookmaker to incur a sure loss,i.e. for every choice of stakes $\sigma_1, \ldots, \sigma_k$, there exists a world w in which the bookmaker's total balance is not negative. It is called incoherent otherwise.

de Finetti's foundation of subjective probability: coherent betting odds

- It is possible for the gambler to pay a negative amount σ_i on e_i. This assumption is called reversibility.
- The book β is called coherent if it does not ensure bookmaker to incur a sure loss,i.e. for every choice of stakes $\sigma_1, \ldots, \sigma_k$, there exists a world w in which the bookmaker's total balance is not negative. It is called incoherent otherwise.
- The probability of an event is the amount of money a that a coherent and reversible bookmaker would propose for that event.

de Finetti's foundation of subjective probability: coherent betting odds

- It is possible for the gambler to pay a negative amount σ_i on e_i. This assumption is called reversibility.
- The book β is called coherent if it does not ensure bookmaker to incur a sure loss,i.e. for every choice of stakes $\sigma_1, \ldots, \sigma_k$, there exists a world w in which the bookmaker's total balance is not negative. It is called incoherent otherwise.
- The probability of an event is the amount of money a that a coherent and reversible bookmaker would propose for that event.
  Example: How much would you bet on the bridge between Reggio and Messina resisting for 200 years?

de Finetti's foundation of subjective probability: coherent betting odds

A suitable formalization of classical de Finetti's betting game consists in interpreting events, books and possible worlds this way:

1. events are elements of an arbitrary boolean algebra $\mathbf{B}$,
2. a book on a finite subset $\{e_1, \ldots, e_k\} \subseteq B$ is a map $\beta : e_i \mapsto \beta_i \in [0, 1]$,
3. a possible world is a structure preserving map from $\mathbf{B}$ into the two element boolean algebra $\mathbf{2}$, that is, any element of $\mathcal{H}(\mathbf{B}, \mathbf{2})$.

de Finetti's foundation of subjective probability

Classical Coherence Criterion
Let $\mathbf{B}$ be a boolean algebra and let $\{e_1, \ldots, e_k\}$ be a finite subset of B. A book $\beta : e_i \mapsto \beta_i$ is said to be coherent iff for each choice of $\sigma_1, \ldots, \sigma_k \in \mathbb{R}$, there exists $w \in \mathcal{H}(\mathbf{B}, \mathbf{2})$ such that:

$$\sum_{i=1}^{k} \sigma_i(\beta_i - w(e_i)) \geq 0$$

Theorem

Let $\mathbf{B}$ be a boolean algebra, $B' = \{e_1, ..., e_k\}$ be a finite subset of B and let β be a book on B'. Then the following are equivalent:

1. *β is coherent.*
2. *There exists a probability p of $\mathbf{B}$ such that p coincides with β over B'.*

de Finetti's foundation of subjective probability

De Finetti never considered the case of many-valued events, anyway it is not difficult to reframe his coherence criterion in the many-valued realm:

Many-valued Coherence Criterion
Let $\mathbf{A}$ be an MV-algebra and $A' = e_1, \ldots, e_k$ be a finite subset of A. We say that a book $\beta : e_i \mapsto \beta_i$ is coherent iff for each choice of $\sigma_1, \ldots, \sigma_k \in \mathbb{R}$, there exists $w \in \mathcal{H}(\mathbf{A}, [0,1]_{MV})$ such that

$$\sum_{i=1}^{k} \sigma_i(\beta_i - w(e_i)) \geq 0.$$

Theorem

Let $\mathbf{A}$ be an algebra of Łukasiewicz logic, $A' = \{e_1, ..., e_k\}$ be a finite subset of A and let β be a book on A'. Then the following are equivalent:

1. *β is coherent.*
2. *There exists a state s of $\mathbf{A}$ such that s coincides with β over A'.*

Recap

- Distinction between uncertainty and vagueness;

Recap

- Distinction between uncertainty and vagueness;
- Probability of many-valued events: theory of states;

Recap

- Distinction between uncertainty and vagueness;
- Probability of many-valued events: theory of states;
- Integral representation of states of Ł, G and Π: states as expected values of bounded random variables;

Recap

- Distinction between uncertainty and vagueness;
- Probability of many-valued events: theory of states;
- Integral representation of states of Ł, G and Π: states as expected values of bounded random variables;
- States as operators averaging the truth value of Ł, G, Π logics;

Recap

- Distinction between uncertainty and vagueness;
- Probability of many-valued events: theory of states;
- Integral representation of states of Ł, G and Π: states as expected values of bounded random variables;
- States as operators averaging the truth value of Ł, G, Π logics;
- States as subjective probability measures in de Finetti's theory.