U-Sets as a probabilistic set theory

Claudio Sossai
ISIB-CNR, Corso Stati Uniti 4, 35127 Padova, Italy
sossai@isib.cnr.it

Technical Report 05/03 ISIB-CNR, October 2005

Abstract

A topos of presheaves can be seen as an extension of classical set theory, where sets vary over informational states, therefore it is a powerful and expressive mathematical framework. I introduce a suitable topos of presheaves where probabilities and probabilistic reasoning can be represented. In this way we obtain a mathematical definition of probabilistic sets. A valid and complete proof system, w.r.t. the intended semantics of probabilities, is described using the internal language of the topos.

Key words: reasoning with uncertainty, topos theory, sheaf theory

1 Introduction

It is a well known fact that topos and in particular sheaves and presheaves theory describe a set theory whose objects vary over informational states. From the logical point of view, the novelty stays in the local aspect of truth, i.e. the question is not whether a proposition is true (or false) but where, in which informational state or context, a proposition is true (or false). For this reason, the internal language of a topos can be seen as a context-dependent language. In this work I proceed in the investigation of a claim presented in previous works (see [7, 8]): key aspects of uncertainty and hence of uncertainty reasoning can be mathematically described using context-dependent logical languages. For this reason I believe that topos theory is a powerful mathematical tool to describe uncertainty in its various aspects: linguistic uncertainty (possibility theory), stochastic uncertainty (probability theory) and their generalizations: evidence theory and imprecise probabilities (see e.g. [2, 6, 9]). In possibility theory, a possibility distribution can be seen as a context where a linguistic construct takes meaning. For example, a particular possibility distribution that describes the linguistic term *long* can be seen as a context where *long* takes meaning. Using this simple observation, I have proved that the internal language of the topos of presheaves built over the category whose objects are sets of possibility distributions, contains a valid and complete logic for possibilistic reasoning. In this work I prove that all the above observations and mathematical hypotheses hold also for stochastic uncertainty, providing new evidence to the claim that a key aspect of uncertainty is the contextualization of truth and hence of the entire language.

Let us shortly recall the hierarchy that links different semantics of uncertainty.

1. Imprecise probabilities are defined by sets of probability measures [9].
2. Beliefs functions can be seen as special cases of imprecise probabilities. In fact these functions are known to be the lower envelope of the probability measures that dominate them [2, 6].
3. Possibility measures are special cases of belief functions i.e. consonant belief functions. Similarly, probability measures are belief functions where all the mass is given to the atoms.

All these uncertainty models share the property that they are defined using measures over a Boolean algebra. Many logics for uncertainty have been introduced to cope with the above different semantics. For many of them the starting point is that atomic formulae can be represented as pairs (*meta-events*) $\langle L, \alpha \rangle$ meaning: "the degree of L is α". From a logical point of view the following structure is common to many logical system for uncertainty reasoning:

- atomic formulae are meta-events $\langle L, \alpha \rangle$;
- formulae are atomic formulae combined with connectives;
- a model is defined as a measure μ.

The starting point for semantics is:

$$\mu \models \langle L, \alpha \rangle \text{ iff } \mu(L) = \alpha.$$

Every logic has a set of axioms and rules (usually a generalization of modus ponens), that define when a formula ϕ is provable, i.e. $\vdash \phi$. For instance in the case of probabilistic logic additivity of measures is introduced as an axiom (see [3]).

Validity and completeness reads:

$$\vdash \phi \quad \text{iff} \quad (\forall \mu)(\mu \models \phi)$$

Our aim is to find a categorical description of meta-events, abstracting from measures, but with the property that, if interpreted using measures, it gives valid and complete logical systems for the corresponding semantics (without ad hoc axioms).

More precisely, we define an extension of classical set theory (**U**–*Sets*) whose internal logical language has the above properties.

Let us point out the fact that no ad hoc axiom, as in many logics for uncertainty, is necessary to find a valid and complete logic. Set theory has been used to axiomatize probability theory. In set theory it is possible to prove that the cardinality of a set X is lower than the cardinality of the set of subsets of X (i.e. $|X| < |2^X|$), but to prove that $P(\neg L) = 1 - P(L)$ we need the axioms of probability theory.

In the probabilistic version of **U**–*Sets*, $P(\neg L) = 1 - P(L)$ is true without assuming the axioms of probability theory: its truth is a consequence of the internal symmetries of **U**–*Sets* (like for $|X| < |2^X|$ in set theory). Therefore, we can say that **U**–*Sets* is a universe of sets where the symmetries of uncertainty are internally represented (i.e., a *geometry of uncertainty*).

In this work I prove that the same mathematical construction used to define the process of contextualizing linguistic uncertainty is sufficient to describe also stochastic uncertainty. More precisely, I will construct the topos of presheaves that vary over the category of contexts (in this case, sets of probability measures) and I will prove that a valid and complete logic for probabilistic reasoning (see [3]) can be represented in the internal language.

2 U-Sets as probabilistic sets

Let $\mathcal{B}$ be a Boolean algebra and $\mathcal{P}$ be the set of all probability measures μ defined on $\mathcal{B}$.

Def. 1 *A context is an element of* $\mathbf{U} = 2^{\mathcal{P}}$.

Where $2^{\mathcal{P}}$ is the power set of $\mathcal{P}$. It is easy to see that $\mathbf{U}$ determines a category in the following way: the objects are the elements of $\mathbf{U}$, the arrows are the immersions, i.e. there is an arrow $f : B \to A$ iff $B \subseteq A$. Again, I will call this category $\mathbf{U}$.

Def. 2 $\mathbf{U}$*–Sets is the topos of presheaves over* $\mathbf{U}$

The internal language of $\mathbf{U}$–*Sets* is defined as usual (see [1]). To fix the notation I will only recall the definition of the connectives. Formulae are terms of type Ω, if $F \xrightarrow{\alpha} \Omega$, $F \xrightarrow{\beta} \Omega$ are formulae, then the connectives are defined, by induction, for all A, $a \in F(A)$ as: $(\alpha \wedge \beta)_A(a) = (\alpha)_A(a) \cap (\beta)_A(a)$; $(\alpha \vee \beta)_A(a) = (\alpha)_A(a) \cup (\beta)_A(a)$; $(\alpha \to \beta)_A(a) = \{f \in max^A : \Omega(f)((\alpha)_A(a)) \subseteq \Omega(f)((\beta)_A(a))\}$; $(\neg\alpha)_A(a) = \{f \in max^A : \Omega(f)((\alpha)_A(a)) = \emptyset\}$, where max^A is the set of all arrows with codomain A. If $F \times G \xrightarrow{\alpha} \Omega$, i.e. α contains an additional variable x_G of type G, then: $((\exists x_G)\alpha)_A(a) = \{f \in max^A : (\exists b \in G(B))((\alpha)_B(\langle F(f)(a), b\rangle) = max^B)\}$; $((\forall x_G)\alpha)_A(a) = \{f \in max^A : (\forall\ C \xrightarrow{g} B)(\forall b \in G(C))((\alpha)_C(\langle F(f \circ g)(a), b\rangle) = max^C\}$. The *semantical entailment* is defined, for $F \xrightarrow{\alpha} \Omega$, $F \xrightarrow{\beta} \Omega$, as: $\alpha \models \beta$ iff $(\forall A)(\forall a \in F(A))(\alpha_A(a) \subseteq \beta_A(a))$. Moreover, validity of a formula α in $\mathbf{U}$–*Sets* is defined as: $\models \alpha$ iff $(\forall A)(\forall a \in F(A))(\alpha_A(a) = max^A)$. As usual in topos theory:

Def. 3 *Probabilistic sets are objects of* $\mathbf{U}$*–Sets of power type.*

3 Reasoning about probabilities

Let me describe some useful objects of $\mathbf{U}$–*Sets*.

Def. 4 Δ^S *(where* S *is a set) is the constant functor defined as* $\Delta^S(A) = S$ *for every object* A *and* $\Delta^S(f) = 1_S$ *for every arrow* f.

In particular, I will consider $\Delta^{\mathcal{B}}$, where $\mathcal{B}$ is the Boolean algebra and Δ^R, where R is the set of real numbers. Using the internal language of $\mathbf{U}$–*Sets*, it is possible to prove that the elements of the constant presheaf Δ^R satisfy the properties of Dedekind cuts (see [5]), hence I will take Δ^R as the set of real numbers in $\mathbf{U}$–*Sets*. To describe measures in $\mathbf{U}$–*Sets*, let me introduce the following functor:

Def. 5 *Let* $\mathcal{S} : \mathbf{U}^{op} \to$ *Sets be defined as:*

1. $\mathcal{S}(A) = \{k : A \to R\}$, *i.e. the set of all functions k from A to the set R of usual real numbers*

2. *for* $B \xrightarrow{f} A$, $\mathcal{S}(A) \xrightarrow{\mathcal{S}(f)} \mathcal{S}(B)$ *is the function that sends every $g \in \mathcal{S}(A)$ to g restricted to B.*

Note that Δ^R can be embedded in $\mathcal{S}$. In fact, for every usual real number $r \in R$, let $\ulcorner r \urcorner \in \mathcal{S}(A)$ be the function that sends A to r, i.e. $\ulcorner r \urcorner : A \longmapsto r$, then $\iota^{\Delta^R} : \Delta^R \to \mathcal{S}$ is defined, for all A, $r \in \Delta^R(A)$ as : $\iota_A^{\Delta^R}(r) = \ulcorner r \urcorner$. When the context is clear, to simplify the notation, I will write r for the internal representation of the real number r, i.e. $r : 1 \to \mathcal{S}$ defined for all A as $r_A(\star) = \ulcorner r \urcorner$. Now we have a mathematical machinery sufficient to give the categorical definition of measure.

Def. 6 *For every $L \in \Delta^{\mathcal{B}}$ let $p_A(L) \in \mathcal{S}(A)$ be the function that maps every $\mu \in A$ to $\mu(L)$.*

The *variable measure* can be represented as follows:

Def. 7 *For every A, P_A is the function $P_A : \Delta^{\mathcal{B}} \to \mathcal{S}(A)$ that sends $L \in \Delta^{\mathcal{B}}$ to the function $p_A(L) \in \mathcal{S}(A)$.*

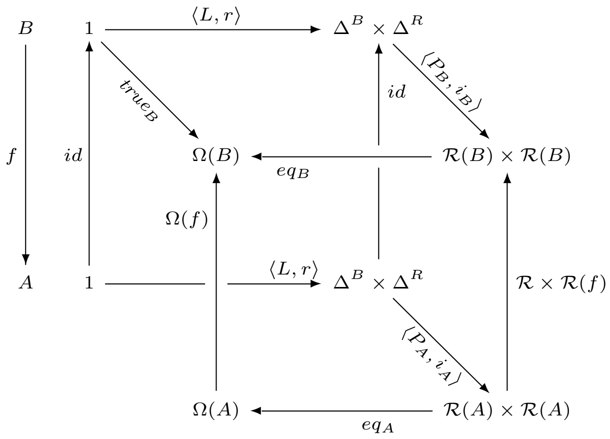
It is easy to see that the family $\{P_A : A \, \text{object of } U\}$ defines a natural transformation $P : \Delta^{\mathcal{B}} \to \mathcal{S}$. Every element L of the Boolean algebra $\mathcal{B}$ is represented by the arrow $L : 1 \to \Delta^{\mathcal{B}}$ defined as $L_A(\star) = L \in \Delta^{\mathcal{B}}(A)$. The following lemma shows that P is the internal representation of probability measures.

Lemma 1 *For all A, $P(L) = r$ is true in A (i.e. $(P(L) = r)_A(< \star, \star >) = max^A$) iff $(\forall \mu \in A)(\mu(L) = r)$*

Proof As usual, equality $S \times S \xrightarrow{=} \Omega$ is defined as the characteristic arrow of the diagonal $S \xrightarrow{\delta} S \times S$. It holds that: $(P(L) = r)_A(< \star, \star >) = \{B \xrightarrow{f} A | < S(f)((P(L))_A(\star)), S(f)(r_A(\star)) > \in \delta_B(S(B))\} = \{B \xrightarrow{f} A | S(f)((P(L))_A(\star)) = S(f)(r_A(\star))\} = \{B \xrightarrow{f} A | (\forall \mu \in B)(\mu(L) = r))\}$. Using the above equalities it easy to see that $(P(L) = r)_A < \star, \star >= max^A$ iff $(\forall \mu \in A)(\mu(L) = r))$. □
The above lemma says that $P(L) = r$ is true in the informational state A iff for all $\mu \in A$ it holds that $\mu(L) = r$. This can be expressed in the language of category theory requiring that the following diagram commutes:

$$\begin{array}{ccc} 1(A) & \xrightarrow{\langle L_A, r_A \rangle} & \Delta^B \times \Delta^R \\ {\scriptstyle true_A}\downarrow & & \downarrow{\scriptstyle \langle P_A, i_A \rangle} \\ \Omega(A) & \xleftarrow[eq_A]{} & \mathcal{R}(A) \times \mathcal{R}(A) \end{array} \qquad (1)$$

This gives a method to compute the truth value of $P(L) = r$ in a generic informational state A. In fact the truth value of $P(L) = r$ in A is given by the set of all arrows $f : B \to A$ s.t. the above diagram commutes in B as in:



This definition of categorical measure remains the same for different semantics of uncertainty (possibilities, probabilities and imprecise probabilities), and is the key idea for a topos description of uncertainty.

Due to the fact that $1 \times 1 \simeq 1$, to simplify the notation, I will assume that all arrows with domain $1 \times .. \times 1$ are defined on 1 implicitly assuming the above morphism. The usual order relation over R can be defined in the internal language as follows:

Def. 8 *For every A, $\langle h, k\rangle \in \mathcal{S}(A) \times \mathcal{S}(A)$, let $\geq_A$: $\mathcal{S}(A) \times \mathcal{S}(A) \to \Omega(A)$ be the function defined as $\geq_A (\langle h, k\rangle) = \{B \xrightarrow{f} A : (\forall \mu \in B)(\mathcal{S}(f)(h)(\mu) \geq \mathcal{S}(f)(k)(\mu))\}$.*

Lemma 2 *The family of functions $\{\geq_A$: A object of $U\}$ defines a natural transformation.*

Proof We must prove that for every arrow $B \xrightarrow{f} A$ it holds that $\geq_B \circ \langle \mathcal{S}(f), \mathcal{S}(f)\rangle = \Omega(f) \circ \geq_A$. Let $\langle h, k\rangle \in \mathcal{S}(A) \times \mathcal{S}(A)$. Note that:

1. $(\Omega(f) \circ \geq_A)(\langle h, k\rangle) = \{C \xrightarrow{g} B : f \circ g \in \geq_A (\langle h, k\rangle)\} = \{C \xrightarrow{g} B : (\forall \mu \in C)(\mathcal{S}(f \circ g)(h)(\mu) \geq \mathcal{S}(f \circ g)(k)(\mu))\}$.

2. $\geq_B (\langle \mathcal{S}(f)(h), \mathcal{S}(f)(k)\rangle) = \{C \xrightarrow{g} B : (\forall \mu \in C)((\mathcal{S}(g) \circ \mathcal{S}(f))(h)(\mu) \geq (\mathcal{S}(g) \circ \mathcal{S}(f))(k)(\mu))\}$

The identity follows from the fact that $\mathcal{S}(f \circ g) = \mathcal{S}(g) \circ \mathcal{S}(f)$ □

Let me note that the usual operations of product and sum defined over R can be extended to $\mathcal{S}$. In fact, let the sum be the natural transformation $+ : \mathcal{S} \times \mathcal{S} \to \mathcal{S}$ defined for $< g, k > \in \mathcal{S}(A) \times \mathcal{S}(A)$ and $\mu \in A$ as $+(< g, k >)(\mu) = g(\mu) + k(\mu) \in R$. To simplify the notation I will often write $g + k$ for $+(< g, k >)$. Product is defined similarly, and the reader can easily check the naturality of the two transformations. Using the above definitions I will prove that the internal language $\mathcal{L}^{\mathbf{U}\text{–}Sets}$ of $\mathbf{U}$–*Sets* contains a sublanguage $\mathcal{L}^P$, that is a conservative extension of $\mathcal{L}^Q$ (see [3]) a valid and complete logic w.r.t. the semantics of probabilities. An atomic formula of the language $\mathcal{L}^Q$ is a term of the form $r_1 l(L_1) + ... + r_n l(L_n) \geq r$, where $L_i \in \mathcal{B}$ and $r, r_i \in R$. A formula is a Boolean combination of atomic formulae. A model is a measure μ and the definition of semantics is:

- $\mu \models r_1 l(L_1) + ... + r_n l(L_n) \geq r$ iff $r_1 \mu(L_1) + ... + r_n \mu(L_n) \geq r$
- $M \models \neg_{\mathcal{B}} f$ iff $M \not\models f$
- $M \models f \wedge_{\mathcal{B}} g$ iff $M \models f$ and $M \models g$

where $\neg_{\mathcal{B}}$ and $\wedge_{\mathcal{B}}$ are the Boolean connectives. When the context is clear I will drop the subscript.

Def. 9 *The translation* $T : \mathcal{L}^Q \rightarrow \mathcal{L}^{\mathbf{U}\text{-}Sets}$ *is defined as:*

- $T(r_1 l(L_1) + ... + r_n l(L_n) \geq r) = r_1 P(L_1) + ... + r_n P(L_n) \geq r$
- $T(\neg_{\mathcal{B}} f) = \neg T(f)$
- $T(f \wedge_{\mathcal{B}} g) = T(f) \wedge T(g)$

Let me check that definition 9 is a good definition, i.e. if $f = g$ then $T(f) = T(g)$. To this aim we need the following lemma.

Lemma 3 *For all measures* μ *it holds that:* $(T(f))_{\{\mu\}}(\star) = max^{\{\mu\}}$ *iff* $\mu \models f$.

Proof By induction on the length of f

- let $f = \sum_{i \leq n} r_i l(L_i) \geq r$, it is easy to see that $(T(f))_{\{\mu\}}(\star) = max^{\{\mu\}}$ iff $\sum_{i \leq n} r_i \mu(L_i) \geq r$ iff $\mu \models f$
- let $f = f_1 \wedge f_2$ then $max^{\{\mu\}} = (T(f_1 \wedge f_2))_{\{\mu\}}(\star)$ iff $(T(f_1))_{\{\mu\}}(\star) = max^{\{\mu\}}$ and$(T(f_2))_{\{\mu\}}(\star) = max^{\{\mu\}}$ iff (by inductive hypothesis) $\mu \models f_1$ and $\mu \models f_2$ iff $\mu \models f$
- Let $f = \neg f_1$.
  - if $max^{\{\mu\}} = (T(\neg f_1))_{\{\mu\}}(\star) = (\neg T(f_1))_{\{\mu\}}(\star)$ then not $max^{\{\mu\}} = (T(f_1))_{\{\mu\}}(\star)$, therefore, by inductive hypothesis, not $\mu \models f_1$ and hence $\mu \models \neg f_1$.
  - if $\mu \models \neg f_1$ then not $\mu \models f_1$ and hence not $max^{\{\mu\}} = (T(f_1))_{\{\mu\}}(\star)$. Note that $max^{\{\mu\}} = \{1_{\{\mu\}} : \{\mu\} \rightarrow \{\mu\}\}$, hence $(T(f_1))_{\{\mu\}}(\star) = \emptyset$. From this we obtain: $(T(\neg f_1))_{\{\mu\}}(\star) = (\neg T(f_1))_{\{\mu\}}(\star) = \{B \xrightarrow{f} \{\mu\} | \Omega(f)((T(f_1))_{\{\mu\}}(\star)) = \emptyset\} = \{B \xrightarrow{f} \{\mu\} | \Omega(f)(\emptyset) = \emptyset\} = max^{\{\mu\}}$.

□

Fact 1 *If* $f = g$ *then* $T(f) = T(g)$

Proof

If $f = g$ then $(\forall \mu)(\mu \models f$ iff $\mu \models g)$ iff (by lemma 3) $(\forall \mu)(T(f)_{\{\mu\}}(\star) = T(g)_{\{\mu\}}(\star))$. From this it holds that, for all A, $(T(f))_A(\star) = (T(g))_A(\star)$ therefore $T(f) = T(g)$. □

Note that as a consequence of fact 1 we have that $\neg\neg T(f) = \neg(T(\neg f)) = T(\neg\neg f) = T(f)$ therefore the formulae of $\mathcal{L}^P$ behave classically. A valid and complete proof system is given by the following set of axioms (AX^P) and modus ponens (see [3]):

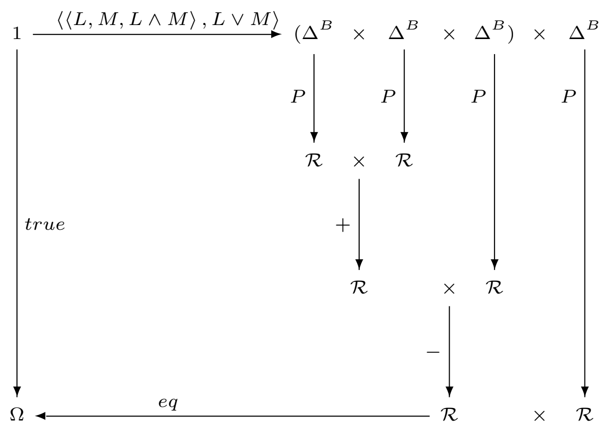
- TAUT: all instances of propositional tautologies in $\mathcal{L}^Q$
- MP: from f and $f \to g$ infer g
- INEQ: all instances of valid formulae about linear inequalities
- L1: $l(L_1) = l(L_2)$ iff $L_1 \equiv L_2$ is a propositional tautology.
- L2: $l(\top_{\mathcal{B}}) = 1$
- L3: $l(L) \geq 0$
- L4: $l(L_1 \vee L_2) = l(L_1) + l(L_2) - l(L_1 \wedge L_2)$.

An inequality formula is a formula of the form $r_1x_1 + ... + r_nx_n \geq r$. It is said valid if it is true under every possible assignment of real numbers to variables. To get an instance of INEQ, we replace each variable x_i that occurs in a valid inequality formula by a primitive likelihood term of the form $l(L_i)$. L2 and L4 are the defining properties of probabilities. Validity and completeness means that for all μ it holds that $\mu \models f$ iff f is provable using the above sets of axioms and rules.

Corollary 1 *For all formula $f \in \mathcal{L}^Q$ it holds that $T(f) = true$ iff $(\forall\mu)(\mu \models f)$*

Proof $T(f) = true$ iff $(\forall A)((T(f))_A(\star) = max^A$ iff $(\forall \mu \in A)((T(f))_{\{\mu\}}(\star) = max^{\{\mu\}})$ iff $(\forall\mu)(\mu \models f)$. □

An example of a *true* formula is given by the property of additivity. The reader can check that the following diagram commutes because for every probability measure μ it holds that: $\mu(L \vee M) = \mu(L) + \mu(M) - \mu(L \wedge M)$.



Therefore, the formula of the internal language

$$P(L) + P(M) - P(L \wedge M) = P(L \vee M)$$

is *true* as an internal symmetry of **U**–*Sets*, without ad hoc axioms.

Due to the fact that every axiom of AX^P is true in every model we have that:

Theorem 1 *All the formulae of* **U**–*Sets obtained as translations of the axioms* AX^P *are true in* **U**–*Sets*

Therefore I can assume as axioms the set $AX^U = \{T(f)|f \in AX^P\}$. Due to the fact that the proof system of the internal language of a topos is valid, i.e. from *true* formulae we derive *true* formulae, we have that if $T(g)$ can be proved from the set of axioms AX^U then $T(g)$ is true. This proves validity:

Theorem 2 *If* $T(f)$ *is provable in* **U**–*Sets using* AX^U *as proper axioms, then* $(\forall\mu)(\mu \models f)$.

As I have pointed out, the formulae of $\mathcal{L}^P$ behave classically as the formulae of $\mathcal{L}^Q$. Moreover, all the negated atomic formulae of $\mathcal{L}^Q$ can be written as positive inequality (i.e. $\neg(l(L) \geq s)$ can be written as $-l(L) > -s$). This property holds also in **U**–*Sets*:

Fact 2 $\neg(P(L) \geq r) = P(L) < r = -P(L) > -r$

Proof For all A it holds that: $\neg(P(L) \geq r)_A(\star) = \{B \xrightarrow{f} A|\Omega(f)(P(L) \geq r)_A(\star)) = \emptyset\} = \{B \xrightarrow{f} A|\neg(\exists C \xrightarrow{g} B)(f \circ g \in (P(L) \geq r)_A(\star)\} = \{B \xrightarrow{f} A|(\forall\mu \in B)(\neg(\mu(L) \geq r))\} = \{B \xrightarrow{f} A|(\forall\mu \in B)(\mu(L) < r)\} = (P(L) < r)_A(\star)$. The second equality follows easily. □

For completeness it remains to prove that the proof system of **U**–*Sets* restricted to $\mathcal{L}^P$, the part of the internal language that represents $\mathcal{L}^Q$, is as strong as the proof system of $\mathcal{L}^Q$. I have proved that all the axioms are *true*, it remains to prove that MP can be represented as a valid rule. The proof of a formula $\vdash \beta$ from $\vdash \alpha$ and $\vdash \alpha \rightarrow \beta$ is an easy exercise in intuitionistic sequent calculus. What remains to see is that the translation is faithful w.r.t. implication. This is a consequence of the fact that the formulae of $\mathcal{L}^P$ behave classically but it can also be checked directly:

Fact 3 $T(f \rightarrow g) = T(\neg(f \wedge \neg g)) = \neg(T(f) \wedge \neg(T(g))) = T(f) \rightarrow T(g)$.

Proof For all A it holds by definition that $(T(f) \rightarrow T(g))_A(\star) = \{f \in max^A|\Omega(f)((T(f))_A(\star)) \subseteq \Omega(f)((T(g))_A(\star))\}$ while $(\neg(T(f) \wedge \neg T(g)))_A(\star) = \{f \in max^A|\Omega(f)((T(f))_A(\star) \cap (T(g))_A(\star)) = \emptyset\}$. From this using the fact that if $S_1, S_2 \in \Omega(A)$ then $\Omega(f)(S_1 \cap S_2) = \Omega(f)(S_1) \cap \Omega(f)(S_2)$ we obtain that $(\neg(T(f) \wedge \neg T(g)))_A(\star) = \{f \in max^A|\Omega(f)((T(f))_A(\star)) \cap \Omega(f)((\neg T(g))_A(\star)) = \emptyset\}$. Note that $\Omega(f)((\neg T(g))_A(\star)) = \{k \in max^B|f \circ k \notin (T(g))_A(\star)\}$, from this the fact follows. □

By fact 3 and from the fact that the rule MP is provable in intuitionistic sequent calculus, we obtain that the translation preserves also the (only) deductive rule of $\mathcal{L}^Q$. Assume that $(\forall\mu)(\mu \models f)$. Then by completeness of $\mathcal{L}^Q$ we have only two possibilities: f is an axiom and therefore $T(f)$ is provable in **U**–*Sets* or it is obtained using MP and also in this case $T(f)$ is provable in **U**–*Sets*. Therefore we obtain completeness for the translation.

References

[1] J. L. Bell. *Toposes and local set theories.* Clarendon Press, Oxford, 1988.

[2] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[3] R. Fagin, J. Y. Halpern, and M. Megiddo. A logic for reasoning about probabilities. *Information and Computation*, 87:78–128, 1990.

[4] J. Lambek and P. J. Scott. *Introduction to higher order categorical logic*. Cambridge University Press, Cambridge, 1986.

[5] S. Mac Lane and I. Moerdijk. *Sheaves in geometry and logic: a first introduction to topos theory*. Springer-Verlag, New York, 1992.

[6] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, NJ, 1976.

[7] C. Sossai. U-sets as a possibilistic set theory. *Fuzzy Sets and Systems*, 144:193–212, 2004.

[8] C. Sossai. Reasoning with uncertainty and context-dependent languages. In *Fourth Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-LFA 2005)*, pages 769–774, Barcelona, Spain, 2005.

[9] P. Walley. *Statistical reasoning with imprecise probability*. Chapman and Hall, 1991.