Decision Making Under Vagueness and Uncertainty *

Audun Jøsang
University of Oslo
Norway
josang@ifi.uio.no

***Abstract*—This paper describes how belief, uncertainty and vagueness can be represented by subjective opinions. Criteria for decision making can be articulated in terms of uncertainty and vagueness in addition to expected utility. The notions of uncertainty and vagueness provide a simple way of explaining the Ellsberg paradox in decision making.**

I. Introduction

Decision making is the process of identifying and choosing between alternative options based on beliefs about the different options and their associated utility gains or losses. The decision maker can be the analyst of the situation, or can act on advice produced by an analyst. In the following, we do not distinguish between the decision maker and the analyst, and use the term 'analyst' to cover both.

Opinions can form the basis of decisions, and this paper investigates how various aspects of an opinion should (rationally) determine the optimal decision. For this purpose, we introduce new belief, uncertainty and vagueness concepts that are described in SectionII. Decision criteria are defined in Section IV and applied to explain Ellsberg's paradox.

A. Subjective Opinions

Subjective logic is a formalism that represents uncertain probabilistic information in the form of *subjective opinions*, and that defines a variety of operations for subjective opinions. In this section we present in detail the concept of subjective opinion which is used for representing uncertain and vague arguments and which provide basis for decision making.

In subjective logic a *domain* is a state space consisting of two or more values. The values of the domain can e.g. be observable or hidden states, events, hypotheses or propositions, just like in traditional Bayesian modeling. Domains are typically specified to reflect realistic situations for the purpose of being practically analysed in some way.

The different values of the domain are assumed to be mutually exclusive and exhaustive, which means that the variable can take only one value at any time, and that all possible values of interest are included in the domain. For example, if the variable is the *WEATHER*, we can assume its domain to be the set $\{$*rainy, sunny, overcast*$\}$.

For a given variable of interest, the values of its domain are assumed to be the real possible states. A vague observation may indicate that the variable takes one of several possible states, but not which one in particular. For example, we might believe that the weather will be either *rainy* or *sunny*, but not which of them. For this reason it is useful to let subsets of the domain be possible values of the variable, i.e. to let the variable take values from the *hyperdomain* which contains the singleton values as well as the composite values. In this case we are talking about a *hypervariable* in contrast to a random variable. Belief mass can be assigned to the values of the hypervariable according to the available information.

A subjective opinion distributes a *belief mass* over the values of the hyperdomain. The sum of the belief masses is less than or equal to 1, and is complemented by the *uncertainty mass*. A subjective opinion also contains a *base rate* probability distribution expressing prior knowledge about the specific class of random variables, so that in case of significant uncertainty about a specific variable, the base rates provide a basis for default likelihoods.

Let X be a variable over a domain $\mathbb{X} = \{x_1, x_2, \ldots, x_k\}$ with cardinality k, where x_i $(1 \leq i \leq k)$ represents a specific value from the domain. Let $\mathscr{P}(\mathbb{X})$ be the powerset of $\mathbb{X}$. The *hyperdomain* is the reduced powerset of $\mathbb{X}$, denoted by $\mathscr{R}(\mathbb{X})$, and defined as follows:

$$\mathscr{R}(\mathbb{X}) = \mathscr{P}(\mathbb{X}) \setminus \{\{\mathbb{X}\}, \{\emptyset\}\}. \tag{1}$$

All proper subsets of $\mathbb{X}$ are values of $\mathscr{R}(\mathbb{X})$, but $\mathbb{X}$ and $\emptyset$ are not, since they are not considered possible observations to which we can assign beliefs. The hyperdomain has cardinality $2^k - 2$. We use the same notation for the values of the domain and the hyperdomain, and consider X a *hypervariable* when it takes values from the hyperdomain.

The composite set $\mathscr{C}(\mathbb{X})$ is the set of non-singleton values, expressed as

$$\mathscr{C}(\mathbb{X}) = \mathscr{R}(\mathbb{X}) \setminus \mathbb{X} \ . \tag{2}$$

Let A denote an *agent*. A's opinion on the variable X is a tuple denoted

$$\omega_X^A = (\boldsymbol{b}_X^A, u_X^A, \boldsymbol{a}_X^A), \tag{3}$$

where $\boldsymbol{b}_X^A : \mathscr{R}(\mathbb{X}) \rightarrow [0,1]$ is a belief mass distribution, the parameter $u_X^A \in [0,1]$ is an uncertainty mass, and $\boldsymbol{a}_X^A : \mathbb{X} \rightarrow [0,1]$ is a base rate probability distribution satisfying the following additivity constrains:

$$u_X^A + \sum_{x \in \mathscr{R}(\mathbb{X})} \boldsymbol{b}_X^A(x) = 1, \tag{4}$$

$$\sum_{x \in \mathbb{X}} \boldsymbol{a}_X^A(x) = 1 \ . \tag{5}$$

* In the proceedings of the 19th International Conference on Information Fusion (FUSION 2016), Heidelberg, July, 2016.

In the notation of the subjective opinion ω_X^A, the subscript denotes the variable X, the *object* of the opinion, while the superscript denotes the opinion owner A, the *subject* of the opinion. Explicitly expressing subjective ownership of opinions makes it possible to express that different agents have different opinions on the same variable. Indication of opinion ownership can be omitted when the subject is clear or irrelevant, for example, when there is only one agent in the modelled scenario.

The belief mass distribution $\boldsymbol{b}_X^A$ has $2^k - 2$ parameters, whereas the base rate distribution $\boldsymbol{a}_X^A$ only has k parameters. The uncertainty parameter u_X^A is a simple scalar. A general opinion thus contains $2^k + k - 1$ parameters. However, given that Eq.(4) and Eq.(5) remove one degree of freedom each, opinions over a domain of cardinality k only have $2^k + k - 3$ degrees of freedom.

A subjective opinion in which $u_X = 0$, i.e. an opinion without uncertainty, is called a *dogmatic opinion*. A dogmatic opinion for which $b_X(x) = 1$, for some x, is called an *absolute opinion*. In contrast, an opinion for which $u_X = 1$, and consequently, $b_X(x) = 0$, for every $x \in \mathscr{R}(\mathbb{X})$, i.e. an opinion with complete uncertainty, is called a *vacuous opinion*.

Every subjective opinion 'projects' to a probability distribution $\mathbf{P}_X$ over $\mathbb{X}$ defined through the following function:

$$\mathbf{P}_X(x_i) = \sum_{x_j \in \mathscr{R}(\mathbb{X})} \boldsymbol{a}_X(x_i|x_j)\, \boldsymbol{b}_X(x_j) + \boldsymbol{a}_X(x_i)\, u_X, \tag{6}$$

where $a_X(x_i|x_j)$ is the *relative base rate* of $x_i \in \mathbb{X}$ with respect to $x_j \in \mathscr{R}(\mathbb{X})$ defined as follows:

$$\boldsymbol{a}_X(x_i|x_j) = \frac{\boldsymbol{a}_X(x_i \cap x_j)}{\boldsymbol{a}_X(x_j)}\ , \tag{7}$$

where a_X is extended on $\mathscr{R}(\mathbb{X})$ additively. For the relative base rate to be always defined, it is enough to assume $a_X^A(x_i) > 0$, for every $x_i \in \mathbb{X}$. This means that everything we include in the domain has a non-zero probability of occurrence in general.

Binomial opinions apply to binary random variables where the belief mass is distributed over two values. Multinomial opinions apply to random variables in n-ary domains, and where the belief mass is distributed over the values of the domain. General opinions, also called *hyper-opinions*, apply to hypervariables where belief mass is distributed over values in hyperdomains obtained from n-ary domains. A binomial opinion is equivalent to a Beta probability density function, a multinomial opinion is equivalent to a Dirichlet probability density function, and a hyper-opinion is equivalent to a Dirichlet hyper-probability density function [1]. Binomial opinions thus represent the simplest opinion type, which can be generalised to multinomial opinions, which in turn can be generalised to hyper-opinions. Simple visualisations for binomial and trinomial opinions are based on barycentric coordinate systems as illustrated in Figures 1 below.

In general, a multinomial opinion can be represented as a point inside a regular simplex. In particular, a trinomial opinion can be represented inside a tetrahedron (a 4-axis barycentric system), as shown in Figure 1.

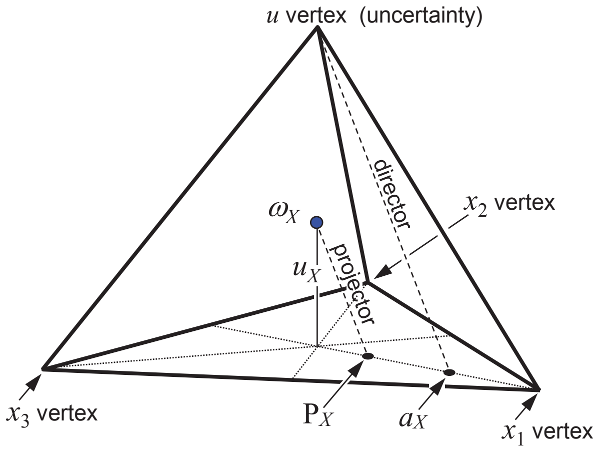

Figure 1. Visualisation of a trinomial opinion

Assume the random variable X on domain $\mathbb{X} = \{x_1, x_2, x_3\}$. Figure 1 shows multinomial opinion ω_X with belief mass distribution $\boldsymbol{b}_X = (0.20, 0.20, 0.20)$, uncertainty mass $u_X = 0.40$ and base rate distribution $\boldsymbol{a}_X = (0.750, 0.125, 0.125)$.

II. Aspects of Belief and Uncertainty in Opinions

The above section on opinions only distinguishes between belief mass and uncertainty mass. This section dissects belief into more granular types called *sharpness*, *vagueness* and *focal uncertainty*.

A. Sharpness

Belief mass that only supports a specific value is called *sharp belief mass*, because it sharply supports a single value and discriminates between values. Note that we also interpret belief mass on a composite value (and its subsets) to be sharp for that composite value, because it discriminates between that composite value and any other value which is not a subset of that value.

Definition 1 (Sharpness): Let $\mathbb{X}$ be a domain with hyperdomain $\mathscr{R}(\mathbb{X})$ and variable X. Given an opinion ω_X, the sharpness of value $x \in \mathscr{R}(\mathbb{X})$ is the function $\boldsymbol{b}_X^{\mathrm{S}} : \mathscr{R}(\mathbb{X}) \to [0,1]$ expressed as

$$\textit{Sharp belief mass:}\quad \boldsymbol{b}_X^{\mathrm{S}}(x) = \sum_{x_i \subseteq x} \boldsymbol{b}_X(x_i)\ . \tag{8}$$

It is useful to express sharpness of composite values in order to assist decision making in situations like the Ellsberg paradox described in Section V.

The total sharp belief mass denoted b_X^{TS} is simply the sum of all belief masses assigned to singletons, defined as follows.

Definition 2 (Total Sharp Belief Mass): Let $\mathbb{X}$ be a domain with variable X, and let ω_X be an opinion on $\mathbb{X}$. The total sharpness of the opinion ω_X is the function $b_X^{\mathrm{TS}} : X \to [0,1]$ expressed as

$$\textit{Total sharp belief mass:}\quad b_X^{\mathrm{TS}} = \sum_{x_i \in \mathbb{X}} \boldsymbol{b}_X(x_i)\ . \tag{9}$$

Sharpness represents the complement of the sum of vagueness and uncertainty, as described below.

B. Vagueness

Recall from Eq.(2) that the composite set $\mathscr{C}(\mathbb{X})$ is the set of all composite values from the hyperdomain. Belief mass assigned to a composite value expresses cognitive vagueness, because this type of belief mass supports the truth/presence of multiple singletons in $\mathbb{X}$ simultaneously, i.e. it does not discriminate between the singletons in the composite value. In the case of binary domains, there can be no vague belief mass, because there are no composite values. In the case of hyperdomains, composite values exist, and every singleton $x \in \mathbb{X}$ is a member of multiple composite values. The vagueness of a singleton $x \in \mathscr{R}(\mathbb{X})$ is defined as the weighted sum of belief masses on the composite values of which x is a member, where the weights are determined by the base rate distribution. The total amount of vague belief mass is simply the sum of belief masses on all composite values in the hyperdomain. The formal definitions of these concepts are given next.

Definition 3 (Vagueness): Let $\mathbb{X}$ be a domain and $\mathscr{R}(\mathbb{X})$ denote its hyperdomain. Let $\mathscr{C}(\mathbb{X})$ be the composite set of $\mathbb{X}$ according to Eq.(2). Let $x \in \mathscr{R}(\mathbb{X})$ denote a value in hyperdomain $\mathscr{R}(\mathbb{X})$, and let $x_j \in \mathscr{C}(\mathbb{X})$ denote a composite value in $\mathscr{C}(\mathbb{X})$. Given an opinion ω_X, the vagueness of value $x \in \mathscr{R}(\mathbb{X})$ is the function $\boldsymbol{b}^{\mathrm{V}}_X : \mathscr{R}(\mathbb{X}) \to [0,1]$ expressed as

$$\text{Vague belief mass: } \boldsymbol{b}^{\mathrm{V}}_X(x) = \sum_{\substack{x_j \in \mathscr{C}(\mathbb{X}) \\ x_j \not\subseteq x}} \boldsymbol{a}_X(x|x_j)\, \boldsymbol{b}_X(x_j) \ . \quad (10)$$

Note that Eq.(10) not only defines vagueness of singletons $x \in \mathbb{X}$, but also defines vagueness of composite values $x \in \mathscr{C}(\mathbb{X})$, i.e. of all values $x \in \mathscr{R}(\mathbb{X})$.

In case x is a composite value, then the belief mass $\boldsymbol{b}_X(x)$ does not contribute to the vagueness of x, despite $\boldsymbol{b}_X(x)$ representing vague belief mass for the whole opinion. The vague belief mass in an opinion ω_X is defined as the sum of belief masses on composite values $x_j \in \mathscr{C}(\mathbb{X})$, formally defined as follows.

Definition 4 (Total Vague Belief Mass): Let $\mathbb{X}$ be a domain with variable X, and let ω_X be an opinion on $\mathbb{X}$. The total vagueness of the opinion ω_X is the function $b^{\mathrm{TV}}_X : X \to [0,1]$ expressed as:

$$\text{Total vague belief mass: } b^{\mathrm{TV}}_X = \sum_{x_j \in \mathscr{C}(\mathbb{X})} \boldsymbol{b}_X(x_j) \ . \quad (11)$$

An opinion ω_X is dogmatic and vague when $b^{\mathrm{TV}}_X = 1$, and is partially vague when $0 < b^{\mathrm{TV}}_X < 1$. An opinion has mono-vagueness when only a single composite value has (vague) belief mass assigned to it. On the other hand, an opinion has pluri-vagueness when several composite values have (vague) belief mass assigned to them.

Note the difference between uncertainty and vagueness in subjective logic. Uncertainty reflects lack of evidence, whereas vagueness results from evidence that fails to discriminate between specific singletons. A vacuous (totally uncertain) opinion – by definition – does not contain any vagueness. Hyper-opinions can contain vagueness, whereas multinomial and binomial opinions never contain vagueness. The ability to express vagueness is thus the main aspect that makes hyper-opinions different from multinomial opinions.

Under the assumption that collected evidence never decays, uncertainty can only decrease over time, because accumulated evidence is never lost. As the natural complement, sharpness and vagueness can only increase. At the extreme, a dogmatic opinion where $b^{\mathrm{TV}}_X = 1$ expresses *dogmatic vagueness*. A dogmatic opinion where $b^{\mathrm{TS}}_X = 1$ expresses *dogmatic sharpness*, which is equivalent to a traditional probability distribution over a random variable.

Under the assumption that evidence decays e.g. as a function of time, uncertainty can increase over time because uncertainty increase is equivalent to the loss of evidence. Vagueness decreases in case new evidence is sharp, i.e. when the new evidence supports singletons, and old vague evidence decays. Vagueness increases in case new evidence is vague, i.e. when the new evidence supports composite values, and the old sharp evidence decays.

C. Dirichlet Visualisation of Opinion Vagueness

The vagueness of a trinomial opinion can not easily be visualised as such on the opinion tetrahedron. However, it can be visualised in the form of a hyper-Dirichlet PDF. Let us for example consider the ternary domain $\mathbb{X}$ with corresponding hyperdomain $\mathscr{R}(\mathbb{X})$ illustrated in Figure 2.

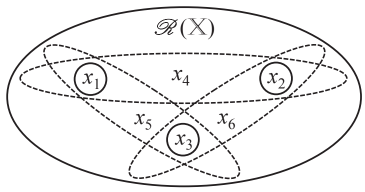

Figure 2. Hyperdomain for the example of vague belief mass

The singletons and composite values of $\mathscr{R}(\mathbb{X})$ are listed below.

$$\begin{cases} \text{Domain:} & \mathbb{X} & = \{x_1, x_2, x_3\}, \\ \text{Hyperdomain:} & \mathscr{R}(\mathbb{X}) & = \{x_1, x_2, x_3, x_4, x_5, x_6\}, \\ \text{Composite set:} & \mathscr{C}(\mathbb{X}) & = \{x_4, x_5, x_6\}, \end{cases}$$

$$\text{where} \quad \begin{cases} x_4 = \{x_1, x_2\}, \\ x_5 = \{x_1, x_3\}, \\ x_6 = \{x_2, x_3\}. \end{cases}$$

Let us further assume a hyper-opinion ω_X with belief mass distribution and base rate distribution specified in Eq.(12) below.

$$\begin{array}{ll} \text{Belief mass distribution} & \text{Base rate distribution} \\ \begin{cases} \boldsymbol{b}_X(x_6) & = 0.8, \\ u_X & = 0.2. \end{cases} & \begin{cases} \boldsymbol{a}_X(x_1) & = 0.33, \\ \boldsymbol{a}_X(x_2) & = 0.33, \\ \boldsymbol{a}_X(x_3) & = 0.33. \end{cases} \end{array} \quad (12)$$

Note that this opinion has mono-vagueness, because the vague belief mass is assigned to only one composite value.

The projected probability distribution on X computed with Eq.(6), and the vague belief mass computed with Eq.(10), are given in Eq.(13) below.

$$\text{Projected probability distribution} \begin{cases} \mathbf{P}_X(x_1) &= 0.066, \\ \mathbf{P}_X(x_2) &= 0.467, \\ \mathbf{P}_X(x_3) &= 0.467. \end{cases} \quad \text{Vague belief mass} \begin{cases} \boldsymbol{b}^{\mathrm{V}}_X(x_1) &= 0.0, \\ \boldsymbol{b}^{\mathrm{V}}_X(x_2) &= 0.4, \\ \boldsymbol{b}^{\mathrm{V}}_X(x_3) &= 0.4. \end{cases} \tag{13}$$

The hyper-Dirichlet PDF for this vague opinion is illustrated in Figure 3. Note how the probability density is spread out along the edge between the x_2 and x_3 vertices, which precisely indicates that the opinion expresses vagueness between x_2 and x_3. To be mindful of vague belief of this kind can be useful for an analyst, in the sense that it can exclude specific values from being plausible. A non-plausible value in this example is x_1.

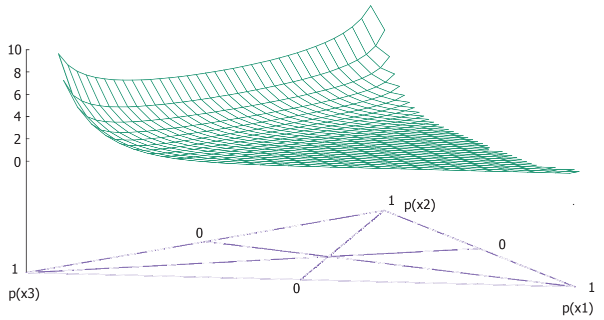

Figure 3. Hyper-Dirichlet PDF with vague belief

In the case of multinomial and hypernomial opinions larger than trinomial, it is challenging to design visualisations. A possible solution in case visualisation is required for opinions over large domains is to use partial visualisation over specific values of the domain that are of interest to the analyst.

D. Focal Uncertainty

When an opinion contains uncertainty, the simplest interpretation is to consider that the whole uncertainty mass is shared between all the values of the (hyper)domain. However, as indicated by the expressions for projected probability of e.g. Eq.(6), the uncertainty mass can be interpreted as being implicitly assigned to (hyper)values of the variable, as a function of the base rate distribution over the variable. This interpretation is captured by the definition of focal uncertainty mass.

Definition 5 (Focal Uncertainty): Let $\mathbb{X}$ be a domain and $\mathscr{R}(\mathbb{X})$ denote its hyperdomain. Given an opinion ω_X, the focal uncertainty mass of a value $x \in \mathscr{R}(\mathbb{X})$ is computed with the function $\boldsymbol{u}^{\mathrm{F}}_X : \mathscr{R}(\mathbb{X}) \to [0,1]$ defined as

$$\textit{Focal uncertainty mass:} \quad \boldsymbol{u}^{\mathrm{F}}_X(x) = \boldsymbol{a}_X(x)\, u_X \ . \tag{14}$$

E. Mass-Sum

The sharpness, vagueness and focal uncertainty concepts defined in the previous section are representative for each value by pulling belief and uncertainty mass proportionally across the belief masses and the uncertainty of the opinion. The concatenation of sharpness, vagueness and focal uncertainty is then called mass-sum, and similarly for total mass-sum.

The additivity properties of mass-sums are described next.

F. Mass-Sum for Sharpness, Vagueness and Focal Uncertainty

The sum of sharpness, vagueness and focal uncertainty of a value is equal to the value's projected probability, expressed as

$$\boldsymbol{b}^{\mathrm{S}}_X(x) + \boldsymbol{b}^{\mathrm{V}}_X(x) + \boldsymbol{u}^{\mathrm{F}}_X(x) = \mathbf{P}_X(x). \tag{15}$$

Eq.(15) shows that the projected probability can be split into three parts which are: i) sharpness, ii) vagueness, and iii) focal uncertainty. The composition of these three parts, called *mass-sum*, is the function denoted $\mathbf{M}_X(x)$. The concept of mass-sum is defined next.

Definition 6 (Mass-Sum): Let $\mathbb{X}$ be a domain with hyperdomain $\mathscr{R}(\mathbb{X})$, and assume that the opinion ω_X is specified. Consider a value $x \in \mathscr{R}(\mathbb{X})$ with sharpness $\boldsymbol{b}^{\mathrm{S}}_X(x)$, vagueness $\boldsymbol{b}^{\mathrm{V}}_X(x)$ and focal uncertainty $\boldsymbol{u}^{\mathrm{F}}_X(x)$. The mass-sum function of value x is the triplet denoted $\mathbf{M}_X(x)$ expressed as

$$\textit{Mass-sum:} \quad \mathbf{M}_X(x) = \left(\boldsymbol{b}^{\mathrm{S}}_X(x), \boldsymbol{b}^{\mathrm{V}}_X(x), \boldsymbol{u}^{\mathrm{F}}_X(x)\right). \tag{16}$$

Given an opinion ω_X, each value $x \in \mathscr{R}(\mathbb{X})$ has an associated mass-sum $\mathbf{M}_X(x)$ which is a function of the opinion ω_X. The term '*mass-sum*' means that the triplet of sharpness, vagueness and focal uncertainty has the additivity property of Eq.(15).

In order to visualise a mass-sum, consider the ternary domain $\mathbb{X} = \{x_1, x_2, x_3\}$ and hyperdomain $\mathscr{R}(\mathbb{X})$ illustrated in Figure 4, where the belief masses and uncertainty mass of opinion ω_X are indicated in the diagram.

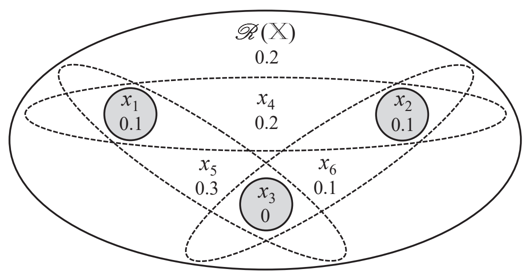

Figure 4. Hyperdomain with belief masses

Formally, the opinion ω_X is specified in Table I. The table includes the mass-sum values in terms of sharpness, vagueness and focal uncertainty. The table also shows the projected probability for every value $x \in \mathscr{R}(\mathbb{X})$.

The mass-sums from opinion ω_X listed in Table I are visualised as a *mass-sum diagram* in Figure 5. Mass-sum diagrams are useful for assisting decision making, because the

Table I
OPINION WITH SHARPNESS, VAGUENESS, FOCAL UNCERTAINTY AND PROJECTED PROBABILITY.

x	$\boldsymbol{b}_X(x)$ u_X	$\boldsymbol{a}_X(x)$	$\boldsymbol{b}^{\mathrm{S}}_X(x)$	$\boldsymbol{b}^{\mathrm{V}}_X(x)$	$\boldsymbol{u}^{\mathrm{F}}_X(x)$	$\mathbf{P}_X(x)$
x_1	0.10	0.20	0.10	0.16	0.04	0.30
x_2	0.10	0.30	0.10	0.16	0.06	0.32
x_3	0.00	0.50	0.00	0.28	0.10	0.38
x_4	0.20	0.50	0.40	0.12	0.10	0.62
x_5	0.30	0.70	0.40	0.14	0.14	0.68
x_6	0.10	0.80	0.20	0.34	0.16	0.70
$\mathbb{X}$	0.20					

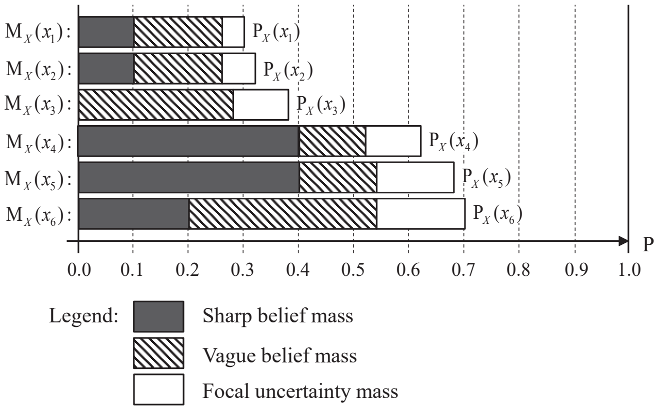

Figure 5. Mass-sum diagram for ω_X

degree of sharpness, vagueness and focal uncertainty can be clearly understood.

Mass-sum diagrams visualise the nature of beliefs in each value, and can also represent hyper-opinions in a way which scales to larger domains.

In Figure 5 it can be seen that x_3 has the greatest projected probability among the singletons, expressed as $\mathbf{P}_X(x_3) = 0.38$. However, the mass-sum of x_3 is void of sharpness, so its projected probability is solely based on vagueness and uncertainty, which affects decision making.

G. Total Mass-Sum

The belief mass of an opinion as a whole can be decomposed into sharpness which provides distinctive support for singletons, and vagueness which provides vague support for singletons. These two belief masses are then complementary to the uncertainty mass. For any opinion ω_X it can be verified that Eq.(17) holds:

$$b^{\mathrm{TS}}_X + b^{\mathrm{TV}}_X + u_X = 1 \ . \tag{17}$$

Eq.(17) shows that the belief and uncertainty mass can be split into the three parts of sharpness, vagueness and focal uncertainty. The composition of these three parts is called *total mass-sum*, denoted $\mathrm{M}^{\mathrm{T}}_X$, and is defined below.

Definition 7 (Total Mass-Sum): Let $\mathbb{X}$ be a domain with hyperdomain $\mathscr{R}(\mathbb{X})$, and assume that the opinion ω_X is specified. The sharpness $\boldsymbol{b}^{\mathrm{S}}_X(x)$, vagueness $\boldsymbol{b}^{\mathrm{V}}_X(x)$ and focal uncertainty $\boldsymbol{u}_X(x)$ can be combined as a triplet, which is then called the total mass-sum, denoted $\mathrm{M}^{\mathrm{T}}_X$ and expressed as

$$\textit{Total mass-sum:}\ \ \mathrm{M}^{\mathrm{T}}_X = \left(b^{\mathrm{TS}}_X, b^{\mathrm{TV}}_X, u_X\right). \tag{18}$$

The total mass-sum of opinion ω_X from Figure 4 and Table I is illustrated in Figure 6.

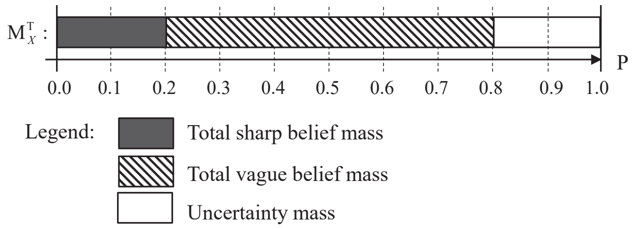

Figure 6. Visualising the total mass-sum from ω_X

III. UTILITY AND NORMALISATION

Utility for random variable is expressed by letting each value x have an associated utility $\boldsymbol{\lambda}_X(x)$ expressed on some scale such as monetary value, which can be positive or negative. Given utility $\boldsymbol{\lambda}_X(x)$ in case of outcome x, then the expected utility for x and total expected utility for X are

$$\text{Expected utility:}\ \ \mathbf{L}_X(x) = \boldsymbol{\lambda}_X(x)\mathbf{P}_X(x)\ , \tag{19}$$

$$\text{Total expected utility:}\ \ \mathrm{L}^{\mathrm{T}}_X = \sum_{x\in\mathbb{X}} \boldsymbol{\lambda}_X(x)\mathbf{P}_X(x)\ . \tag{20}$$

In classical utility theory, decisions are based on expected utility for possible options. Integrating utility into the probabilities for each option [2] produces a *utility-normalised probability vector*, which simplifies decision-making models, because every option then has a simple *utility-probability*.

Normalisation is needed when comparing decision options. The normalisation factor must be appropriate for all variables, so that the utility-normalised probability vectors are within a given range. Note that in case of negative utility for a specific outcome, the utility-normalised probability for that outcome is also negative. Utility-normalised probability is therefore a synthetic notion.

Let λ^+ denote the greatest absolute utility of all utilities in all vectors. Thus, if the greatest absolute utility is negative, then λ^+ takes its positive (absolute) value.

Definition 8 (Utility-Normalised Probability Vector): Let λ^+ denote the greatest absolute utility from $\boldsymbol{\lambda}_X$ and from other relevant utility vectors that will be considered in order to compare different options. The utility-normalised probability vector produced by $\mathbf{P}_X$, $\boldsymbol{\lambda}_X$ and λ^+ is expressed as

$$\mathbf{P}^{\mathrm{N}}_X(x) = \frac{\mathbf{L}_X(x)}{\lambda^+} = \frac{\boldsymbol{\lambda}_X(x)\mathbf{P}_X(x)}{\lambda^+},\ \ \forall x\in\mathbb{X}. \tag{21}$$

Note that the utility-normalised probability vector $\mathbf{P}^{\mathrm{N}}_X$ does not represent a probability distribution, and in general does not satisfy the additivity requirement of a probability distribution. Other utility-normalised notions are defined next.

Definition 9 (Utility-Normalised Masses): The utility-normalised sharpness, vagueness and focal uncertainty for $x \in \mathbb{X}$ are expressed as

$$\text{Utility-normalised sharpness:} \quad \boldsymbol{b}_X^{\mathrm{NS}}(x) = \frac{\boldsymbol{\lambda}_X(x)\boldsymbol{b}_X^{\mathrm{S}}(x)}{\lambda^+}\,, \tag{22}$$

$$\text{Utility-normalised vagueness:} \quad \boldsymbol{b}_X^{\mathrm{NV}}(x) = \frac{\boldsymbol{\lambda}_X(x)\boldsymbol{b}_X^{\mathrm{V}}(x)}{\lambda^+}\,, \tag{23}$$

$$\text{Utility-normalised focal uncertainty:} \quad \boldsymbol{u}_X^{\mathrm{NF}}(x) = \frac{\boldsymbol{\lambda}_X(x)\boldsymbol{u}_X^{\mathrm{F}}(x)}{\lambda^+}\,. \tag{24}$$

The *utility-normalised probability* is expressed in Eq.(25):

$$\text{Utility-normalised probability:} \quad \mathbf{P}_X^{\mathrm{N}}(x) = \boldsymbol{b}_X^{\mathrm{NS}}(x) + \boldsymbol{b}_X^{\mathrm{NV}}(x) + \boldsymbol{u}_X^{\mathrm{N}}(x). \tag{25}$$

Having defined utility-normalised probability, it is possible to directly compare options without involving utilities, because utilities are integrated into the utility-normalised probabilities.

The utility-normalised mass-sum is defined below.

Definition 10 (Utility-Normalised Mass-Sum): The utility-normalised mass-sum function of x is the triplet denoted $\mathbf{M}_X^{\mathrm{N}}(x)$ expressed as

$$\text{Utility-normalised mass-sum:} \quad \mathbf{M}_X^{\mathrm{N}}(x) = \Big(\boldsymbol{b}_X^{\mathrm{NS}}(x), \boldsymbol{b}_X^{\mathrm{NV}}(x), \boldsymbol{u}_X^{\mathrm{NF}}(x)\Big)\,. \tag{26}$$

Note that utility-normalised sharpness, vagueness and focal uncertainty do not represent realistic measures, and must be considered as purely synthetic.

As an example of applying utility-normalised probability, consider two urns named X and Y that both contain 100 red and black balls, where you will be asked to draw a ball at random from one of the urns. The possible outcomes are named $x_1 =$ 'Red' and $x_2 =$ 'Black' for urn X, and are similarly named $y_1 =$ 'Red' and $y_2 =$ 'Black' for urn Y.

You are told that urn X contains 70 red balls, 10 black balls and 20 balls that are either red or black. The corresponding opinion ω_X is expressed as

$$\text{Opinion } \omega_X = \begin{pmatrix} \boldsymbol{b}_X(x_1) = 7/10, & \boldsymbol{a}_X(x_1) = 1/2, \\ \boldsymbol{b}_X(x_2) = 1/10, & \boldsymbol{a}_X(x_2) = 1/2, \\ u_X \quad\; = 2/10. & \end{pmatrix} \tag{27}$$

You are told that urn Y contains 40 red balls, 20 black balls and 40 balls that are either red or black. The corresponding opinion ω_Y is expressed as

$$\text{Opinion } \omega_Y = \begin{pmatrix} \boldsymbol{b}_Y(y_1) = 4/10, & \boldsymbol{a}_Y(y_1) = 1/2, \\ \boldsymbol{b}_Y(y_2) = 2/10, & \boldsymbol{a}_Y(y_2) = 1/2, \\ u_Y \quad\; = 4/10. & \end{pmatrix} \tag{28}$$

Imagine that you must select one ball at random, from either urn X or Y, and you are asked to make a choice about which urn to draw it from in a single betting game. With option X, you receive \$1000 if you draw 'Black' from urn X (i.e. if you draw x_2). With option Y, you receive \$500 if you draw 'Black' from urn Y (i.e. if you draw y_2). You receive nothing if you draw 'Red' in either option. Table II summarises the options in this game.

Table II
BETTING OPTIONS IN SITUATION INVOLVING UTILITIES

	Red	Black
Option X, draw from urn X:	0	\$1000
Option Y, draw from urn Y:	0	\$500

The mass-sums for drawing 'Black' are different for options X and Y. However, the utility-normalised mass-sums are equal, as illustrated in Figure 7. The normalisation factor used in this example is $\lambda^+ = 1000$, since \$1000 is the greatest absolute utility.

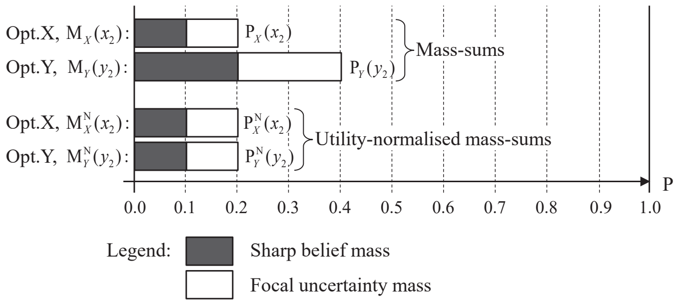

Figure 7. Diagram for mass-sums and for utility-normalised mass-sums

In general, the option with the greatest utility-normalised probability should be chosen. Note that the utility-normalised probability is equal for options X and Y, expressed as $\mathbf{P}_X^{\mathrm{N}}(x_2) = \mathbf{P}_Y^{\mathrm{N}}(y_2)$. Hence, utility-normalised probability alone is insufficient for determining the best option in this example. The decision in this case must be based on the sharpness, which is greatest for option Y, expressed as $\boldsymbol{b}_Y^{\mathrm{S}}(y_2) > \boldsymbol{b}_X^{\mathrm{S}}(x_2)$.

Note that is is not be meaningful to consider utility-normalised sharpness for choosing between options. This is explained in detail in Section IV.

In case of equal utilities for all options, then normalisation is not needed, or it can simply be observed that utility-normalised mass-sums are equal to the corresponding non-normalised mass-sums, as expressed below.

$$\text{When all options have equal utility:} \quad \begin{cases} \text{Projected probability:} & \mathbf{P}_X^{\mathrm{N}} \;\; = \mathbf{P}_X\,, \\ \text{Sharpness:} & \boldsymbol{b}_X^{\mathrm{NS}} \;\; = \boldsymbol{b}_X^{\mathrm{S}}\,, \\ \text{Vagueness:} & \boldsymbol{b}_X^{\mathrm{NV}} \;\; = \boldsymbol{b}_X^{\mathrm{V}}\,, \\ \text{Focal uncertainty:} & \boldsymbol{u}_X^{\mathrm{NF}} \;\; = \boldsymbol{u}_X^{\mathrm{F}}\,, \\ \text{Mass-sum:} & \mathbf{M}_X^{\mathrm{N}} \;\; = \mathbf{M}_X\,. \end{cases}$$

In the examples below, utilities for all options are equal, so for convenience, the diagrams show simple mass-sums, which are equal to the corresponding utility-normalised mass-sums.

IV. Decision Criteria

The decision criteria follow the indicated order of priority.

1) The option with the highest utility-normalised probability is the best choice.

2) Given equal maximal utility-normalised probability among multiple options, the option with the greatest sharpness is the best choice.
3) Given equal maximal utility-normalised probability as well as equal maximal sharpness among multiple options, the option with the least focal uncertainty (and therefore with the greatest vagueness, whenever relevant) is the best option.

The above criteria predict the choice of the majority of participants in the Ellsberg experiment described below.

V. The Ellsberg Paradox

The Ellsberg paradox [3] results from an experiment which shows how traditional probability theory is unable to explain typical human decision-making behaviour.

In the Ellsberg experiment you are shown an urn with 90 balls in it, and you are told that 30 balls are red, and that the remaining 60 balls are either black or yellow. One ball is going to be selected at random, and you are asked to make a choice in two separate betting games. Figure 8 shows the hyperdomain of the Ellsberg paradox.

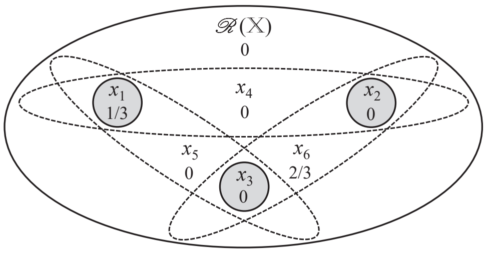

Figure 8. Hyperdomain and belief mass distribution in the Ellsberg paradox

The domain $\mathbb{X}$ and its hyper-opinion are then expressed as

$$\text{Hyperdomain}\quad \mathscr{R}(\mathbb{X}) = \left\{ \begin{array}{ll} x_1: & \text{Red}, \\ x_2: & \text{Black}, \\ x_3: & \text{Yellow}, \\ x_4: & \text{Red or Black}, \\ x_5: & \text{Red or Yellow}, \\ x_6: & \text{Black or Yellow}. \end{array} \right\} \tag{29}$$

$$\text{Hyper-opinion}\quad \omega_X = \left(\begin{array}{ll} \boldsymbol{b}_X(x_1) = 1/3, & \boldsymbol{a}_X(x_1) = 1/3, \\ \boldsymbol{b}_X(x_2) = 0, & \boldsymbol{a}_X(x_2) = 1/3, \\ \boldsymbol{b}_X(x_3) = 0, & \boldsymbol{a}_X(x_3) = 1/3, \\ \boldsymbol{b}_X(x_4) = 0, & \boldsymbol{a}_X(x_4) = 2/3, \\ \boldsymbol{b}_X(x_5) = 0, & \boldsymbol{a}_X(x_5) = 2/3, \\ \boldsymbol{b}_X(x_6) = 2/3, & \boldsymbol{a}_X(x_6) = 2/3, \\ u_X \quad = 0. & \end{array} \right) \tag{30}$$

A quick look at ω_X reveals that it contains some sharp belief mass, some vague belief mass and no uncertainty mass, so it is a dogmatic and partially vague opinion.

In betting game 1 you must choose between option 1A and 1B. With option 1A you receive \$100 if 'Red' is drawn, and you receive nothing if either 'Black' or 'Yellow' is drawn. With option 1B you receive \$100 if 'Black' is drawn, and you receive nothing if either 'Red' or 'Yellow' is drawn. Table III summarises the options in game 1.

Table III
Game 1: Pair of betting options

	Red	Black	Yellow
Option 1A:	\$100	0	0
Option 1B:	0	\$100	0

Make a note of your choice from betting game 1, and then proceed to game 2 where you are asked to choose between two new options based on the same random draw of a single ball from the same urn. With option 2A you receive \$100 if either 'Red' or 'Yellow' is drawn, and you receive nothing if 'Black' is drawn. With option 2B you receive \$100 if either 'Black' or 'Yellow' is drawn, and you receive nothing if 'Red' is drawn. Table IV summarises the options in game 2.

Table IV
Game 2: Pair of betting options

	Red	Black	Yellow
Option 2A:	\$100	0	\$100
Option 2B:	0	\$100	\$100

Would you choose option 2A or 2B?

Ellsberg reports that, when presented with these pairs of choices, most people select options 1A and 2B. Adopting the approach of expected utility theory, this reveals a clear inconsistency in probability assessments. On this interpretation, when a person chooses option 1A over option 1B, he or she is revealing a higher subjective probability assessment of picking 'Red' than of picking 'Black'.

However, when the same person prefers option 2B over option 2A, he or she reveals that his or her subjective probability assessment of picking 'Black' or 'Yellow' is higher than of picking 'Red' or 'Yellow', which implies that picking 'Black' has a higher probability assessment than that of picking 'Red'. This seems to contradict the probability assessment of game 1, which therefore represents a paradox.

When explicitly expressing the vagueness of the opinions, the majority's preference for choices 1A and 2B becomes perfectly rational, as explained next.

The utilities for options 1A and 1B are equal (\$100), so there is no difference between the utility-normalised probabilities and the projected probabilities which are used for decision modelling below. Projected probabilities are computed with Eq.(6) which for convenience is repeated below:

$$\mathbf{P}_X(x) = \sum_{x_j \in \mathscr{R}(\mathbb{X})} \boldsymbol{a}_X(x|x_j)\, \boldsymbol{b}_X(x_j) + \boldsymbol{a}_X(x)\, u_X. \tag{31}$$

Relative base rates are computed with Eq.(7) which for convenience is repeated below:

$$\boldsymbol{a}_X(x|x_j) = \frac{\boldsymbol{a}_X(x \cap x_j)}{\boldsymbol{a}_X(x_j)}. \tag{32}$$

The projected probabilities of x_1 and x_2 in game 1 are then

$$\begin{aligned} &\text{Option 1A:} \quad \mathbf{P}_X(x_1) &= \boldsymbol{a}_X(x_1|x_1)\, \boldsymbol{b}_X(x_1) &= \tfrac{1}{3}\ . \\ &\text{Option 1B:} \quad \mathbf{P}_X(x_2) &= \boldsymbol{a}_X(x_2|x_6)\, \boldsymbol{b}_X(x_2) &= \tfrac{1}{3}\ . \end{aligned} \tag{33}$$

Note that $\mathbf{P}_X(x_1) = \mathbf{P}_X(x_2)$, which makes the the options equal from a purely first-order probability point of view. However they are affected by different vague belief mass as shown below.

Vague belief mass of x, denoted $\boldsymbol{b}^{\mathrm{V}}_X(x)$, is computed with Eq.(10) which for convenience is repeated below:

$$\boldsymbol{b}^{\mathrm{V}}_X(x) = \sum_{\substack{x_j \in \mathscr{C}(\mathbb{X}) \\ x_j \not\subseteq x}} \boldsymbol{a}_X(x|x_j)\, \boldsymbol{b}_X(x_j)\ . \tag{34}$$

The vague belief masses of x_1 and x_2 in game 1 are then

$$\begin{aligned} &\text{Option 1A:} \quad \boldsymbol{b}^{\mathrm{V}}_X(x_1) &= 0\ , \\ &\text{Option 1B:} \quad \boldsymbol{b}^{\mathrm{V}}_X(x_2) &= \boldsymbol{a}_X(x_2|x_6)\, \boldsymbol{b}_X(x_6) \quad = \tfrac{1}{3}\ . \end{aligned} \tag{35}$$

Given the absence of uncertainty, the additivity property of Eq.(15) allows us to compute the sharpnesses as $\boldsymbol{b}^{\mathrm{S}}_X(x_1) = 1/3$ and $\boldsymbol{b}^{\mathrm{S}}_X(x_2) = 0$.

The mass-sum diagram of the options in Ellsberg betting game 1 is illustrated in Figure 9.

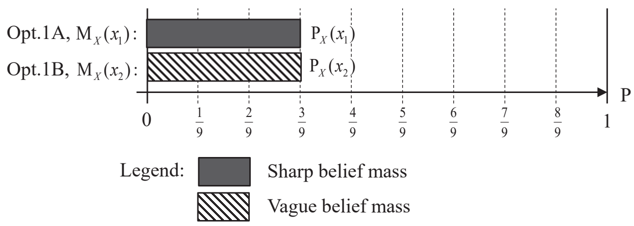

Figure 9. Mass-sum diagram for game 1 in the Ellsberg paradox

The difference between options 1A (x_1) and 1B (x_2) emerges with their different sharp and vague belief masses. People clearly prefer choice 1A because it only has sharpness and no vagueness, whereas choice 1B is affected by vagueness.

We now turn to betting game 2, where Option 2A (x_5) and Option 2B (x_6) have the the projected probabilities

$$\text{Opt.2A: } \mathbf{P}_X(x_5) = \boldsymbol{a}_X(x_1|x_1)\, \boldsymbol{b}_X(x_1) + \boldsymbol{a}_X(x_3|x_6)\, \boldsymbol{b}_X(x_6) = 2/3,$$

$$\text{Opt.2B: } \mathbf{P}_X(x_6) = \boldsymbol{a}_X(x_2|x_6)\, \boldsymbol{b}_X(x_6) + \boldsymbol{a}_X(x_3|x_6)\, \boldsymbol{b}_X(x_6) = 2/3.$$

Note that $\mathbf{P}_X(x_5) = \mathbf{P}_X(x_6)$, which makes the the options equal from a first-order probability point of view. However they have different vague belief masses, as shown below. Vague belief mass is computed with Eq.(10).

The vagueness of x_5 and x_6 in game 2 are

$$\text{Option 2A: } \boldsymbol{b}^{\mathrm{V}}_X(x_5) = \boldsymbol{a}_X(x_3|x_6)\, \boldsymbol{b}_X(x_6) = \tfrac{1}{3}\ ,$$

$$\text{Option 2B: } \boldsymbol{b}^{\mathrm{V}}_X(x_6) = 0\ .$$

Given the absence of uncertainty, the additivity property of Eq.(15) allows us to compute the sharpnesses as $\boldsymbol{b}^{\mathrm{S}}_X(x_5) = 1/3$ and $\boldsymbol{b}^{\mathrm{S}}_X(x_6) = 2/3$.

The mass-sum diagram of the options in Ellsberg betting game 2 is illustrated in Figure 10.

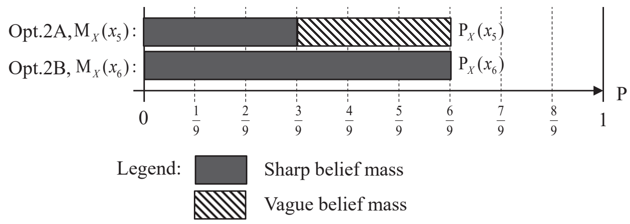

Figure 10. Mass-sum diagram for game 2 in the Ellsberg paradox

The difference between options 2A and 2B emerges with their different sharpness and vagueness. People clearly prefer choice 2B (x_6), because it has no vagueness, whereas choice 2A (x_5) is affected by its vagueness of $1/3$.

We have shown that preferring option 1A over option 1B, and that preferring option 2B over option 2A, is perfectly rational, and therefore does not represent a paradox within the opinion model.

Other models of uncertain probabilities are also able to explain the Ellsberg paradox, such as e.g. Choquet capacities (Choquet 1953 [4], Chateauneuf 1991 [5]). However, the Ellsberg paradox only involves vagueness, not uncertainty. In fact, the Ellsberg paradox is quite simple and does not put into play the whole spectre of sharpness, vagueness and focal uncertainty of opinions which can be relevant for decision making. More complex examples than the Ellsberg paradox can easily be articulated.

VI. Conclusion

The expressiveness of subjective opinions can distinguish between uncertainty and vagueness. This can be combined with utility which produces a powerful framework for decision making.

References

[1] A. Jøsang and R. Hankin, "Interpretation and fusion of hyper-opinions in subjective logic." in *Proceedings of the 15th International Conference on Information Fusion (FUSION 2012)*. IEEE, Los Alamitos, Singapore, July 2012.

[2] E. Castagnoli and M. LiCalzi, "Expected utility without utility," *Theory and Decision*, vol. 41, no. 3, pp. 281–301, November 1996.

[3] D. Ellsberg, "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Ecomonics*, vol. 75, pp. 643–669, 1961.

[4] G. Choquet, "Theory of capacities," *Annales de l'Institut Fourier*, vol. 5, pp. 131–295, 1953.

[5] A. Chateauneuf, "On the use of capacities in modelling uncertainty aversion and risk aversion," *Journal of Mathematical Economics*, vol. 20, no. 4, pp. 343–369, 1991.