# Describing linguistic information in a behavioural framework: possible or not?*

**Gert DE COOMAN**†
Universiteit Gent, Vakgroep Elektrische Energietechniek
Technologiepark – Zwijnaarde 9, 9052 Zwijnaarde, Belgium
E-mail: gert.decooman@rug.ac.be

## Abstract

The paper discusses important aspects of the representation of linguistic information, using imprecise probabilities with a behavioural interpretation. We define linguistic information as the information conveyed by statements in natural language, but restrict ourselves to simple affirmative statements of the type 'subject-is-predicate'. Taking the behavioural stance, as it is described in detail in [24], we investigate whether it is possible to give a mathematical model for this kind of information. In particular, we evaluate Zadeh's suggestion [28] that we should use possibility measures to this end. We come to the conclusion that, generally speaking, possibility measures are possible models for linguistic information, but that more work should be done in order to evaluate the suggestion that they may be the only ones.

## 1 LINGUISTIC INFORMATION

Natural language is often used to transmit information. We shall deal with the question whether the information conveyed by statements in natural language, from now on called *linguistic information*, can be given a mathematical representation. In doing so, we shall restrict ourselves to very simple affirmative statements of the type 'subject-is-predicate', or 'subject-satisfies-property', such as for instance 'John is old'. Such a statement indeed conveys information: if initially we know nothing about John's age, and we subsequently hear that he is old, our knowledge about John's age has increased, although it will not be sufficient to let us say exactly how old he is. In order to formalize this, we consider a *possibility space* $\Omega$ consisting of all the different and mutually exclusive values that a variable $\xi$ (such as John's age) may assume. We look for a mathematical representation of the linguistic information conveyed by the statement '$\xi$ is $p$' — and *only by this statement* —, where $p$ is a property of $\xi$ (such as old).

The property $p$ may be clear or vague, precise or imprecise. Informally, a property is called *vague* if there exist objects which satisfy it only partially, or to a certain extent. We call a property *clear*, or nonvague, if every object considered either completely satisfies it or completely does not. A *clear* property will be called *imprecise* if there is more than one object which satisfies it, and *precise* if there is only one. It is possible to ascribe imprecision to vague properties as well, and say that a vague property is imprecise if there is more than one object which satisfies it at least to a certain extent.

We want to make it very clear from the outset that this discussion is necessarily rather sketchy, due to the conflict between the complexity of the questions we address and the limitations of space a conference paper imposes on us. At the same time, we do not present these results as being final, but rather as a preliminary exploration of the mathematical and behavioural representation of linguistic information. Nevertheless, we believe the paper contains enough relevant and surprising results to make it interesting.

## 2 A BEHAVIOURAL MODEL

We want a mathematical model of linguistic information to be both meaningful and useful. This means that the notions which appear in the model must have a clear interpretation. Moreover, since statements in

* This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian state, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the author.

† Gert de Cooman is a Postdoctoral Fellow of the Belgian National Fund for Scientific Research (NFWO).

natural language convey information, and since humans often rely on this kind of information as a basis for choosing between actions, we want to be able to use the model in a decision-theoretic context. This explains why we want to represent linguistic information in a behavioural context, using the theory of imprecise probabilities as it has been convincingly put forward by Walley in [24]. Let us give a brief discussion of the various relevant notions that we shall need further on.

A *gamble* $X$ is a bounded real-valued mapping on $\Omega$. It can be interpreted as an uncertain reward in some linear utility. Note that $\sup X = \sup_{\omega\in\Omega} X(\omega)$ and $\inf X = \inf_{\omega\in\Omega} X(\omega)$ are finite real numbers. The set of the gambles on $\Omega$ is denoted by $\mathcal{L}(\Omega)$. A subset $A$ of $\Omega$ is called an *event*. Its indicator, or characteristic function, $\chi_A$ can be interpreted as a $0-1$-valued gamble, and we shall often identify $A$ with $\chi_A$, a convention which goes back to de Finetti [9]. In such cases it should be clear from the context whether $A$ denotes an event (a set) or a gamble (an indicator). The set of events is denoted by $\wp(\Omega)$.

An *upper prevision* $\overline{P}$ [24] can be formally defined as a real-valued function on a class of gambles $\mathcal{K} \subseteq \mathcal{L}(\Omega)$. In order to identify the possibility space and domain, we also denote $\overline{P}$ as $(\Omega, \mathcal{K}, \overline{P})$. The *conjugate lower prevision* $\underline{P}$ is defined on $-\mathcal{K} = \{-X \mid X \in \mathcal{K}\}$ by $\underline{P}(X) = -\overline{P}(-X)$. When the domain of an upper (lower) prevision is a class of events — indicators —, it will also be called an upper (lower) *probability*.

Upper and lower previsions and probabilities can be given the following behavioural interpretation[1]. If $\overline{P}(X)$ (respectively $\underline{P}(X)$) is Your[2] upper (lower) prevision for the gamble $X \in \mathcal{L}(\Omega)$, this means that You *state* that You are willing to sell (buy) the gamble $X$ for any price higher (lower) than $\overline{P}(X)$ (respectively $\underline{P}(X)$). Thus, informally $\overline{P}(X)$ is Your *infimum selling price* and $\underline{P}(X)$ Your *supremum buying price* for the gamble $X$.

Since upper and lower previsions have this interpretation in terms of behaviour, they are subject to criteria of rationality, both internal and external. Informally stated, external rationality implies that Your upper and lower previsions should be linked in some way to the available information, be it statistical, combinatorial, linguistic, ... Internal rationality means that Your choice of upper and lower previsions, regardless of how You come by it, should not lead You to a behaviour that is harmful to Yourself (criterion of avoiding sure loss) or internally contradictory (criterion of coherence).

Define, for any gamble $X$ on $\Omega$, $G(X) = X - \underline{P}(X)$, whence $G(-X) = \overline{P}(X) - X$. An upper prevision $(\Omega, \mathcal{K}, \overline{P})$ is said to *avoid sure loss* [24] iff for any $X_1$, ..., $X_n$ in $\mathcal{K}$, $n \geq 1$, $\sup \sum_{k=1}^{n} G(-X_k) \geq 0$. It is said to be *coherent* [24] iff for any $m \geq 0$, $n \geq 0$ and $X_o$, $X_1$, ..., $X_n$ in $\mathcal{K}$, $\sup(\sum_{k=1}^{n} G(-X_k) - mG(-X_o)) \geq 0$. If the upper prevision $(\Omega, \mathcal{K}, \overline{P})$ avoids sure loss, then its *natural extension* [24] $(\Omega, \mathcal{L}(\Omega), \overline{E})$ is the maximal coherent upper prevision on $\mathcal{L}(\Omega)$ that is dominated by $\overline{P}$ on $\mathcal{K}$. It carries the *minimal* behavioural assumptions implied by the initial $(\Omega, \mathcal{K}, \overline{P})$ and coherence.

Note that coherence implies that $\underline{P}(X) \leq \overline{P}(X)$. Your upper (and lower) prevision for $X$ is called *imprecise* iff $\underline{P}(X) < \overline{P}(X)$, and *precise* otherwise.

This discussion of upper and lower previsions, coherence and natural extension is necessarily very limited. For a detailed exposition of the theory of imprecise probabilities, we refer to Walley's book [24]. A good working knowledge of the material covered there is more or less required for a proper understanding of much of what follows.

## 3 THE BAYESIAN APPROACH

From the so-called Bayesian point of view [9], there is a third (internal) rationality criterion that must be satisfied besides avoiding sure loss and coherence: upper and lower prevision should coincide on their common domain[3] (self-conjugacy), and therefore be *precise*. Walley has argued that, whereas avoiding sure loss and coherence are perfectly acceptable as rationality criteria on the behavioural interpretation, it is not always realistic to impose self-conjugacy, especially if there is little available information [24].

An upper prevision which satisfies all three rationality criteria is called a *linear prevision*, and will be given the generic notation $P$. A linear prevision on the linear space $\mathcal{L}(\Omega)$ is a positive linear functional with unit norm, and a linear prevision on an arbitrary domain always coincides on this domain with such a linear functional. In particular, a linear prevision on a field of subsets of $\Omega$ is a (finitely additive) probability.

[1] The interpretation is minimal, because it is allowed that upper and lower previsions have other interpretations besides a behavioural one; how they are obtained is practically irrelevant, provided that they can (also) be interpreted as infimum selling prices and supremum buying prices. Similarly, upper and lower previsions need not be maximally precise. It is not excluded that a more careful analysis of the information available may lead to a more precise model. For a more detailed discussion of the interpretation of upper and lower previsions in the behavioural context, we refer to [24].

[2] We follow Good's convention in calling You the agent whose beliefs (based on the given information) we want to model.

[3] If their domains do not coincide, the situation is more involved, see [9, 24].

In this section, we briefly investigate whether the self-conjugacy criterion is warranted when we want to model the information conveyed by a statement such as '$\xi$ is $p$'. We only look at the most simple case, where $p$ is a clear property and $\Omega$ is finite. Let $A_p = \{\omega \in \Omega \mid \omega \text{ is } p\}$ be the *extension* of $p$, then the given information is equivalent to '$\xi \in A_p$'. It does not allow You to differentiate between the elements of $A_p$, and if You were a Bayesian, You would solve this by using the so-called *noninformative probability distribution* on $A_p$, i.e. by assigning equal probabilities to all the elements of $A_p$ and zero probability to elements outside $A_p$. This leads to the linear prevision $P_p$, given by

$$P_p(X) = \frac{1}{N_{A_p}} \sum_{\omega \in A_p} X(\omega), \quad X \in \mathcal{L}(\Omega), \qquad (1)$$

where $N_{A_p}$ is the number of elements in $A_p$.

However, failure to differentiate between the elements of $A_p$ must lead to equal probabilities only if You are *forced* to compare these elements (choose between them). We believe that it is more rational not to want to compare the elements of $A_p$ on the basis of the given information (which is relatively scarce if $N_{A_p} > 1$), since it *is not detailed enough for You to do so.* Refusing any comparison would lead You to choose the *vacuous* upper prevision $(\Omega, \mathcal{L}(\Omega), \overline{P}_{A_p})$ relative to $A_p$, defined as

$$\overline{P}_{A_p}(X) = \sup_{\omega \in A_p} X(\omega), \quad X \in \mathcal{L}(\Omega) \qquad (2)$$

as Your *upper* prevision based on the available information. We shall give a more detailed justification for this choice in Section 6. A much deeper discussion of why vacuous upper (and lower) previsions are better models for lack of information than noninformative probability distributions can be found in [24].

In conclusion, we are convinced that the *precise* Bayesian model $P_p$ given by (1) is not warranted unless the given information is precise, i.e. $N_{A_p} = 1$, in which case also $P_p = \overline{P}_{A_p}$. When the information — or in other words the property $p$ — is imprecise ($N_{A_p} > 1$), the imprecise upper prevision $\overline{P}_{A_p}$ given by (2) is preferred to $P_p$. When $p$ is vague besides being imprecise, the argument against choosing a linear prevision remains essentially the same.

## 4 POSSIBILITY MEASURES

Lotfi Zadeh was one of the first to recognize that the theory of additive probability is in general not very well suited for modelling linguistic information. In two interesting papers, he provided the basis for an alternative model, which he called the *theory of possibility* [28], and which is based on his notion of a *fuzzy set* [27]. Let us give a very brief overview of the most important ideas in his approach.

First of all, Zadeh recognized that in order to deal with the pervasive vagueness of properties in natural language, a mathematical model is needed in order to represent it. He suggested [27] associating with a vague property $p$ a *membership function* $\mu_p: \Omega \to [0, 1]$, in his view a generalization of the indicator of the extension of a clear property. He suggested the following (somewhat unclear) interpretation for $\mu_p$: for any $\omega$ in $\Omega$, $\mu_p(\omega)$ is the degree to which $\omega$ satisfies the property $p$, or the *degree of membership* of $\omega$ in the fuzzy set of elements satisfying $p$ — which could be called the fuzzy extension of $p$. At the time of writing this, it is still under debate how membership functions and degrees of membership should be interpreted or can be justified. Many suggestions have been given, some measurement-theoretic [20, 21, 22], some purely ordinal [6, 7], others based on random sets [18, 19] and still others based on likelihoods [11, 16]. A good overview can be found in [13]. A very promising behavioural and decision-theoretic justification of membership functions is due to Robin Giles [15], who associates membership degrees with net utility received for making typical assertions.

In order to explain how Zadeh uses membership functions to represent linguistic information, we first give a definition [3] of a possibility measure which is slightly more general than the one he has given [28]. We begin by introducing an *ample field* $\mathcal{R}$ on $\Omega$ as a class of subsets of $\Omega$ that is closed under arbitrary unions and complementation, and therefore also under arbitrary intersections. The *atom* $[\omega]_{\mathcal{R}}$ of $\mathcal{R}$ containing $\omega$ is defined as $[\omega]_{\mathcal{R}} = \bigcap\{A \in \mathcal{R} \mid \omega \in A\}$, $\omega \in \Omega$. The atoms of $\mathcal{R}$ make up a partition of $\Omega$. Note that for any subset $A$ of $\Omega$, $A \in \mathcal{R}$ iff $A = \bigcup_{\omega \in A} [\omega]_{\mathcal{R}}$. The intersection of an arbitrary family of ample fields is again an ample field. Given a subset $\mathcal{E}$ of $\wp(\Omega)$, we may therefore define the smallest ample field $\tau(\mathcal{E})$ including $\mathcal{E}$ as the intersection of all the ample fields which include $\mathcal{E}$.

A *possibility measure* $\Pi$ on $(\Omega, \mathcal{R})$ is a *complete join-morphism* [2] between the complete lattices $(\mathcal{R}, \subseteq)$ and $([0, 1], \leq)$. In other words, for any family $(A_j \mid j \in J)$ of elements of $\mathcal{R}$, $\Pi(\bigcup_{j \in J} A_j) = \sup_{j \in J} \Pi(A_j)$. It follows that $\Pi(\emptyset) = 0$. $\Pi$ is called *normal* iff $\Pi(\Omega) = 1$. A *distribution* for $\Pi$ is a $\Omega{-}[0, 1]$-mapping $\pi$ which is constant on the atoms of $\mathcal{R}$ ($\mathcal{R}$-measurable) and for which $\Pi(A) = \sup_{\omega \in A} \pi(\omega)$, $A \in \mathcal{R}$. Such a distribution is unique and completely determined by $\pi(\omega) = \Pi([\omega]_{\mathcal{R}})$, $\omega \in \Omega$. Conversely, a possibility measure is uniquely and completely determined by its distribution. The *conjugate* or *dual necessity measure*[4] N of a possibility

[4] Note that we shall further on interpret $\Pi$ as an upper

measure $\Pi$ on $(\Omega, \mathcal{R})$ is defined by $\mathrm{N}(A) = 1 - \Pi(\overline{A})$, where $\overline{A}$ is the set-theoretic complement of $A$, $A \in \mathcal{R}$. Note that $\mathrm{N}(A) \leq \Pi(A)$, $\max(\Pi(A), \Pi(\overline{A})) = 1$, $\min(\mathrm{N}(A), \mathrm{N}(\overline{A})) = 0$ and $\Pi(A) < 1 \Rightarrow \mathrm{N}(A) = 0$.

Zadeh [28] has advanced the thesis that the linguistic information conveyed by the statement '$\xi$ is $p$' should be modeled by a possibility measure $\Pi_p$ on $(\Omega, \wp(\Omega))$. More importantly, he has claimed that there exists a link between $\Pi_p$ and the fuzzy set $\mu_p$ representing $p$, through the so-called *possibility assignment equation*:

$$\pi_p = \mu_p,$$

where $\pi_p$ is the distribution of $\Pi_p$. For any $A \subseteq \Omega$, he called $\Pi_p(A) = \sup_{\omega \in A} \mu_p(\omega)$ the *possibility* of $A$, or the possibility that $\xi$ belongs to $A$, given the information '$\xi$ is $p$'.

In what follows, we intend to investigate the validity of these claims when the possibility and necessity measures involved are given the behavioural interpretation of upper and lower probabilities, i.e. are respectively interpreted as infimum selling prices and supremum buying prices for $0-1$-valued gambles.

## 5 INTERNAL RATIONALITY

Let us first investigate whether possibility measures, interpreted as upper probabilities, satisfy the criteria of internal rationality described in Section 2. In other words, we ask ourselves under what conditions possibility measures avoid sure loss and are coherent. At the same time, we study the relationship between possibility measures and natural extension. A more detailed treatment, with proofs of the results outlined here, will be given in a forthcoming paper [8].

Coherence is guaranteed if You specify Your beliefs in the form of a *normal* possibility measure on an ample field.

**Proposition 1** [8] *Let $\mathcal{R}$ be an ample field on $\Omega$, and let $\Pi$ be a possibility measure on $(\Omega, \mathcal{R})$. The following statements are equivalent: $\Pi$ is normal; $(\Omega, \mathcal{R}, \Pi)$ avoids sure loss; and $(\Omega, \mathcal{R}, \Pi)$ is coherent.*

Since a normal possibility measure is in particular 2-alternating (or submodular [10]) on its domain, its natural extension can be calculated using a Choquet integral [1, 23, 24], as is made clear in the following proposition. Interestingly, it also tells us that *natural extension preserves the possibilistic nature of an upper probability*: the natural extension to $\wp(\Omega)$ of a normal possibility measure defined on an ample field $\mathcal{R}$ on $\Omega$, is still a normal possibility measure, *with the same distribution.*

**Proposition 2** [8] *Let $\mathcal{R}$ be an ample field on $\Omega$, let $\Pi$ be a normal possibility measure on $(\Omega, \mathcal{R})$, and let $\overline{E}_\Pi$ be the natural extension of $\Pi$ to $\mathcal{L}(\Omega)$. Then for any $X$ in $\mathcal{L}(\Omega)$, $\overline{E}_\Pi(X) = \overline{E}_\Pi(\tilde{X})$, where*

$$\overline{E}_\Pi(\tilde{X}) = \inf \tilde{X} + \int_{\inf \tilde{X}}^{\sup \tilde{X}} \Pi(\{\omega \in \Omega \mid \tilde{X}(\omega) > x\}) \mathrm{d}x,$$

*and $\tilde{X}(\omega) = \sup_{\nu \in [\omega]_\mathcal{R}} X(\nu)$, $\omega \in \Omega$. In particular, for any $A \subseteq \Omega$, $\overline{E}_\Pi(A) = \Pi(\tilde{A}) = \Pi(\bigcup_{\omega \in A} [\omega]_\mathcal{R})$.*

Of course, specifying the upper probability of all members of an ample field may be a cumbersome task. Let us therefore assume that on the basis of the linguistic information '$\xi$ is $p$' given to You, You make an assessment by specifying an upper probability $(\Omega, \mathcal{E}, \overline{P}_p)$ defined on a set of events $\mathcal{E} \subseteq \wp(\Omega)$ that is not necessarily an ample field.

Since we are interested in possibility measures as upper probabilities, and moreover want to find out whether possibility measures can be used to represent the given linguistic information, it is natural to ask whether $\overline{P}_p$ may be extended to a possibility measure. In other words, does there exist a possibility measure $\Pi$ on[5] $(\Omega, \wp(\Omega))$ such that for all $A$ in $\mathcal{E}$, $\Pi(A) = \overline{P}_p(A)$? And if such indeed be the case, what is the relation between such a *possibilistic extension* $\Pi$, and the natural extension $\overline{E}_p$ of $\overline{P}_p$?

These questions are answered in the following proposition. In order to fix the terminology, call the upper probability $(\Omega, \mathcal{E}, \overline{P}_p)$ *P-consistent* [26] iff for any $A \in \mathcal{E}$ and any $\mathcal{F} \subseteq \mathcal{E}$,

$$A \subseteq \bigcup \mathcal{F} \Rightarrow \overline{P}_p(A) \leq \sup_{F \in \mathcal{F}} \overline{P}_p(F).$$

$(\Omega, \mathcal{E}, \overline{P}_p)$ is called sup-*normalizable* [8] iff

$$\sup_{\omega \in \Omega} \inf_{A \in \mathcal{E}, \omega \in A} \overline{P}_p(A) = 1.$$

Note that $(\Omega, \mathcal{E}, \overline{P}_p)$ is always sup-normalizable[6] if $\mathcal{E}$ does not cover $\Omega$. On the other hand, if $\mathcal{E}$ covers $\Omega$ a necessary condition for sup-normalizability is that for any $\mathcal{F} \subseteq \mathcal{E}$,

$$\bigcup \mathcal{F} = \Omega \Rightarrow \sup_{F \in \mathcal{F}} \overline{P}_p(F) = 1.$$

probability. N is therefore the conjugate lower probability of $\Pi$ [24].

[5] Or on $(\Omega, \mathcal{R})$, where $\mathcal{R}$ is an ample field on $\Omega$ including $\mathcal{E}$. The smallest such ample field is $\tau(\mathcal{E})$. If we can solve the problem for $\wp(\Omega)$, restriction gives us the solution for $\mathcal{R}$. Conversely, given a solution for $\mathcal{R}$, Proposition 2 shows us how to extend the solution to $\wp(\Omega)$. It is therefore to a certain extent irrelevant which $\mathcal{R}$ we choose, provided that $\tau(\mathcal{E}) \subseteq \mathcal{R}$.

[6] In the above definitions, sup and inf are to be taken in the bounded chain $([0, 1], \leq)$!

If $\overline{P}_p$ assumes only a finite number of values, a necessary and sufficient condition for sup-normalizability is that there exist an $\omega$ in $\Omega$ such that

$$(\forall A \in \mathcal{E})(\omega \in A \Rightarrow \overline{P}_p(A) = 1).$$

Both P-consistency and sup-normalizability are obviously rather strong requirements.

**Proposition 3** [8] *The upper probability* $(\Omega, \mathcal{E}, \overline{P}_p)$ *can be extended to a normal possibility measure iff it is P-consistent and* sup-*normalizable. In that case,* $(\Omega, \mathcal{E}, \overline{P}_p)$ *is coherent, and the greatest (least-committal) normal possibility measure* $\Pi^g$ *on* $(\Omega, \wp(\Omega))$ *that coincides with* $\overline{P}_p$ *on* $\mathcal{E}$ *has distribution* $\pi^g$*, given by* $\pi^g(\omega) = \inf_{A \in \mathcal{E}, \omega \in A} \overline{P}_p(A)$*,* $\omega \in \Omega$*. Moreover, for the natural extension* $\overline{E}_p$ *of* $\overline{P}_p$*, we have in that case that*

$$\overline{E}_p(\{\omega\}) = \overline{E}_p([\omega]_{\tau(\mathcal{E})}) = \pi^g(\omega),$$

*where* $\tau(\mathcal{E})$ *is the smallest ample field containing* $\mathcal{E}$*.*

Let us now take a cursory look at a number of special cases. For a more detailed discussion, we refer to [8]. If $\mathcal{E}$ is a *partition* of $\Omega$, then $(\Omega, \mathcal{E}, \overline{P}_p)$ is always P-consistent. It is sup-normalizable iff $\sup_{A \in \mathcal{E}} \overline{P}_p(A) = 1$. The natural extension $\overline{E}_p$ will in that case in general differ from $\Pi^g$, but coincide with it on the elements of the partition $\mathcal{E}$, which are precisely the atoms of $\tau(\mathcal{E})$.

If the elements of $\mathcal{E}$ are *nested*, or in other words constitute a subchain of the complete Boolean lattice $(\wp(\Omega), \subseteq)$, then under fairly general conditions, $(\Omega, \mathcal{E}, \overline{P}_p)$ will be P-consistent. If $\Omega \notin \mathcal{E}$, it is moreover automatically sup-normalizable. If $\Omega \in \mathcal{E}$, sup-normalizability is under the same conditions equivalent to $\overline{P}_p(\Omega) = 1$. It is interesting to note that under some additional continuity conditions imposed on $\overline{P}_p$, *possibilistic and natural extension coincide on all the subsets of* $\Omega$. As a corollary, it can be shown that any normal possibility measure $\Pi$ with distribution $\pi$ is the natural extension of its restriction to the nested family of sets $\{\omega \mid \pi(\omega) \leq x\}$, $x \in [0, 1]$.

An upper probability $(\Omega, \mathcal{E}, \overline{P}_p)$ can only be extended to a possibility measure under fairly restrictive conditions. Even when these are satisfied, the greatest possibilistic extension will in general be strictly dominated by the natural extension: there is only guaranteed equality on the elements of $\mathcal{E}$ and on those subsets of $\Omega$ that are included in the atoms of $\tau(\mathcal{E})$. On other events $A$, the natural extension $\overline{E}_p(A)$ may be strictly greater than the greatest possibilistic extension $\Pi^g(A)$. Since natural extension represents the minimal (least-committal) implications of the upper probability assessment $(\Omega, \mathcal{E}, \overline{P}_p)$, this means that if You accept $\Pi^g(A) < \overline{E}_p(A)$ as Your upper probability for $A$, You will be prepared to sell the gamble $A$ for a lower price than the one Your assessment $(\Omega, \mathcal{E}, \overline{P}_p)$ would lead You to demand, taking into account only the requirements imposed by coherence (internal rationality).

In conclusion, normal possibility measures satisfy the internal rationality criteria of avoiding sure loss and coherence. Nevertheless, only if the initial upper probability assessment $(\Omega, \mathcal{E}, \overline{P}_p)$ satisfies very restrictive criteria — P-consistency, sup-normalizability, natural extension coincides with possibilistic extension[7] — will possibility measures be the best choice for representing the given linguistic information.

# 6 EXTERNAL RATIONALITY

It might be argued that the course of reasoning given above is only based on internal rationality, and is therefore not sufficient to evaluate the claim that linguistic information can or should be modeled by possibility measures. We cannot *a priori* dismiss the hypothesis that if You made a deep and careful enough analysis of the information conveyed by the given statement '$\xi$ is $p$', You would be lead to only make assessments of a very special type, which would then lead to possibility measures as natural extensions. This would imply the existence of a more or less direct link between the given linguistic information and the upper previsions (or probabilities) used to model them, which brings us to considerations of external rationality.

In this section, we evaluate this hypothesis carefully, and point out a number of basic questions which should be dealt with in order to either dismiss or corroborate it. On the one hand, we show that acceptable assessments can be made that are closely linked to the meaning of the given linguistic information, and which nevertheless do not necessarily lead to possibility measures as the corresponding upper probabilities. On the other hand, we show that there also exist acceptable assessments that do necessarily lead to possibility measures as their natural extension.

In order to be able to link the information '$\xi$ is $p$' with upper probabilities, we need a mathematical model representing its meaning. As we mentioned before, an elegant behavioural and decision-theoretic justification for the introduction of Zadeh's membership functions for representing the meaning of the property $p$, has been developed by Giles [15]. Interestingly, Giles' approach takes into account that the meaning of statements in natural language is determined by the society

[7]Note that we have up to now identified only two cases where these conditions are satisfied, namely when $\mathcal{E}$ is an ample field or a collection of nested sets, the latter under some additional continuity assumptions.

in which the statement is used. Here, we use a simpler representation, which can be derived from Giles' model, and which uses only ordinal notions. For a more detailed discussion of this ordinal approach, we refer to [6, 7]. The first basic assumption is that elements in the possibility space $\Omega$ can be compared on the basis of their satisfying the property $p$. This allows us to introduce a binary *preference relation* $R_p$ on $\Omega$, as follows: $\omega R_p \nu$ iff $\omega$ satisfies the property $p$ at least as well as $\nu$ does. The second basic assumption is that $R_p$ is a total preorder on $\Omega$, i.e., is reflexive, transitive and complete[8]. Both assumptions are easily seen to be necessary in order to arrive at the notion of a membership function for $p$. In a standard way, this allows us to construct a partition $\mathcal{D}$ of $\Omega$, whose partition classes are made up of those elements of $\Omega$ which satisfy $p$ equally well, and a total order relation $\succeq_{\mathcal{D}}$ on $\mathcal{D}$, which is the canonical order associated with the preorder $R_p$. For $D_1$ and $D_2$ in $\mathcal{D}$, $D_1 \succeq_{\mathcal{D}} D_2$ means that any element in $D_1$ satisfies $p$ at least as well as any element in $D_2$ does. We shall use the obvious notation $D_1 \succ_{\mathcal{D}} D_2$ for $D_1 \succeq_{\mathcal{D}} D_2$ and $D_2 \not\succeq_{\mathcal{D}} D_1$. Clearly, $D_1 \succeq_{\mathcal{D}} D_2$ and $D_2 \succeq_{\mathcal{D}} D_1$ imply that $D_1 = D_2$.

The link with the membership function $\mu_p$ of the property $p$ is easily made. Since $\mu_p(\omega)$ is informally defined as the degree to which $\omega$ satisfies $p$ [27], we should have for any $\omega$ and $\nu$,

$$\mu_p(\omega) \geq \mu_p(\nu) \Leftrightarrow D(\omega) \succeq_{\mathcal{D}} D(\nu), \tag{3}$$

where $D(\omega)$ is the unique element $D$ of $\mathcal{D}$ for which $\omega \in D$, $\omega \in \Omega$.

Now, in order to make the link between this model and an upper probability assessment, let us look at the following example. Suppose You have been given the information that John is old, and consider the events $A = [0, 30]$ and $B = [40, 80]$ in the possibility space, say, $\Omega = [0, 150]$. Given the meaning of the statement 'John is old' it seems (externally) rational for You to have more confidence in $B$ than in $A$, since any element in $B$ satisfies the property 'old' better than any element in $A$ does. In other words, You are willing to give up the gamble $A$ in exchange of $B$. In Walley's theory this preference is modeled by the assessment $B \geqq A$, where $\geqq$ is a comparative probability ordering[9]. On the other hand, You will not be willing to give up $B$ for $A$, which permits You to make the even stronger assessment $B > A$, i.e., $B \geqq A$ and $A \not\geqq B$.

We may now formalize this as follows. The linguistic information '$\xi$ is $p$' leads You to make the following assessment for any $D_1$ and $D_2$ in $\mathcal{D}$:

$$D_1 \succeq_{\mathcal{D}} D_2 \Rightarrow D_1 \geqq D_2 \tag{4}$$

or the even stronger assessment

$$D_1 \succ_{\mathcal{D}} D_2 \Rightarrow D_1 > D_2. \tag{5}$$

In order not to complicate matters, it will from now on be assumed that the possibility space $\Omega$ is finite, which implies that $(\mathcal{D}, \succeq_{\mathcal{D}})$ is a finite, and therefore bounded, chain [2]. We want to find the answers to two questions. Firstly, does there exist a normal possibility measure $\Pi$ on[10] $(\Omega, \wp(\Omega))$ that is compatible with the given assessment in the sense that for any $D_1$ and $D_2$ in $\mathcal{D}$:

$$D_1 \succeq_{\mathcal{D}} D_2 \Rightarrow D_1 \geqq_{\Pi} D_2 \tag{6}$$

or

$$D_1 \succ_{\mathcal{D}} D_2 \Rightarrow D_1 >_{\Pi} D_2? \tag{7}$$

Here $\geqq_{\Pi}$ is the comparative probability ordering [24] associated with the possibility measure $\Pi$, and defined as follows, for any $A$ and $B$ in $\wp(\Omega)$:

$$\begin{aligned} A \geqq_{\Pi} B \quad &\Leftrightarrow \quad \overline{E}_{\Pi}(B - A) \leq 0 \\ &\Leftrightarrow \quad \mathrm{N}(A \cap \overline{B}) \geq \Pi(\overline{A} \cap B). \end{aligned}$$

In this expression, $\overline{E}_{\Pi}$ is the natural extension of $\Pi$ to $\mathcal{L}(\Omega)$, and N is the necessity measure associated with $\Pi$. If we can answer this question in the positive, we shall also look for the greatest (least-committal) such possibility measure.

Secondly, what is the natural extension[11] of the given comparative probability assessment, or in other words, what is the maximal (least-committal) coherent upper prevision $\overline{P}$ that is compatible with it in the sense that for any $D_1$ and $D_2$ in $\mathcal{D}$, $D_1 \succeq_{\mathcal{D}} D_2 \Rightarrow \overline{P}(D_2 - D_1) \leq 0$ or $D_1 \succ_{\mathcal{D}} D_2 \Rightarrow \overline{P}(D_2 - D_1) \leq 0$ and $\overline{P}(D_1 - D_2) > 0$?

Let us first look at the most simple case, where $p$ is a clear property. Then[12] $\mathcal{D} = \{A_p, \overline{A_p}\}$, where $A_p = \{\omega \in \Omega \mid \omega \text{ is } p\}$ is the extension of $p$, and $\succeq_{\mathcal{D}}$ is completely determined by $A_p \succ_{\mathcal{D}} \overline{A_p}$. For any normal possibility measure $\Pi$ on $(\Omega, \wp(\Omega))$, both (6) and (7) are equivalent to $\Pi(\overline{A_p}) \leq \frac{1}{2}$, so the greatest possibility

[8] The completeness is not absolutely essential here, but we shall impose it nevertheless, in order not to complicate the discussion.

[9] This is linked with upper and lower previsions as follows [24]. $A \geqq B$ means that You are willing give up the gamble $A$ in exchange of $B$, or in other words, that the gamble $B - A$ is desirable to You. A gamble is desirable if You are willing to accept it (i.e. buy it at zero cost), which is equivalent to $\underline{P}(B - A) \geq 0$, or alternatively, to $\overline{P}(A - B) \leq 0$.

[10] Or on $(\Omega, \tau(\mathcal{D}))$. As before this is of no consequence for the problem considered here.

[11] It is easily verified that the given comparative probability assessment avoids sure loss, so that its natural extension is a coherent upper prevision.

[12] Note that we restrict the discussion to the nontrivial case that $A_p \neq \emptyset$ and $A_p \neq \Omega$.

$\Pi$ measure compatible with the given assessment has distribution

$$\pi(\omega) = \begin{cases} 1 & ; \quad \omega \in A_p \\ \frac{1}{2} & ; \quad \omega \in \overline{A_p}. \end{cases}$$

For the natural extension $\overline{P}$ of the given assessment, we have, for any gamble $X$ in $\mathcal{L}(\Omega)$,

$$\overline{P}(X) = \frac{\overline{P}_{A_p}(X) + \overline{P}_{\Omega}(X)}{2},$$

which is also equal to $\overline{E}_{\Pi}(X)$, where $\overline{E}_{\Pi}$ is the natural extension of the greatest compatible possibility measure $\Pi$. In this expression, we used the notation $\overline{P}_C$ introduced in Section 3 for the vacuous prevision relative to a nonempty subset $C$ of $\Omega$:

$$\overline{P}_C(X) = \sup_{\omega \in C} X(\omega), \quad X \in \mathcal{L}(\Omega).$$

We shall have recourse to this notation further on as well.

If we add the assessment $A_p \approx \Omega$ (or equivalently, $\emptyset \approx \overline{A_p}$), which is reasonable if $p$ is clear, then we get an additional condition $\Pi(\overline{A_p}) = 0$, which yields $\pi = A_p$ for the distribution of the greatest possibility measure compatible with the given assessments. Interestingly, the natural extension is in this case given by $\overline{P}(X) = \overline{P}_{A_p}(X)$, $X \in \mathcal{L}(\Omega)$, and again coincides with the natural extension of the greatest compatible possibility measure. Note furthermore that $\pi = A_p$ is precisely Zadeh's *possibility assignment equation*, which he uses to associate a possibility measure with the given linguistic information '$\xi$ is $p$'. In summary, when $p$ is clear, we can corroborate both Zadeh's claim that the given linguistic information is best represented by a possibility measure, based on his possibility assignment equation.

The same conclusion can be reached using a different argument, based on conditioning. What does it mean when we say that You know that '$\xi$ is $p$', and nothing else? Before receiving the information, You are completely ignorant of the value which $\xi$ assumes in $\Omega$, and this can be modeled by the vacuous upper prevision $\overline{P}_{\Omega}$ relative to $\Omega$ [24]. After receiving the information, You know that $\xi \in A_p$, and since $A_p$ is a subset[13] of $\Omega$, You can model the updated information by a conditional (updated) upper prevision $\overline{P}_{\Omega}(\cdot \mid A_p)$ which should be coherent[14] with the unconditional upper prevision $\overline{P}_{\Omega}$. From Walley's coherent updating theorem [24] we deduce that, since $\overline{P}_{\Omega}$ is fully conglomerable, a conditional upper prevision $\overline{P}_{\Omega}(\cdot \mid B)$ coherent with $\overline{P}_{\Omega}$ is uniquely determined by the generalized Bayes' rule, whenever $\underline{P}_{\Omega}(B) > 0$ [24]. But since $\underline{P}_{\Omega}(B) = 0$ unless $B = \Omega$, we cannot use this result. Since, however, $\overline{P}_{\Omega}(B) > 0$ for any nonempty $B$, and since $\overline{P}_{\Omega}$ satisfies Walley's regularity condition, we can use the so-called *regular extension* as a conditional upper prevision which is coherent with $\overline{P}_{\Omega}$ (and in a sense uniquely so) [24]. This leads again to $\overline{P}_{\Omega}(\cdot \mid A_p) = \overline{P}_{A_p}$.

Let us now turn to the more involved case, where the property $p$ is vague. This implies that $\mathcal{D}$ has more than two elements, say $n$. We can always order the elements of $\mathcal{D}$ in such a way that

$$D_1 \succ_{\mathcal{D}} D_2 \succ_{\mathcal{D}} \ldots \succ_{\mathcal{D}} D_n.$$

The natural extension for the assessment (4), as well as for the assessment (5), is given by

$$\overline{P}(X) = \max_{m=1}^{n} \frac{1}{m} \sum_{k=1}^{m} \overline{P}_{D_k}(X), \quad X \in \mathcal{L}(\Omega), \tag{8}$$

and its restriction to $\wp(\Omega)$ is not a possibility measure when $n > 2$. Indeed, we have that $\overline{P}(D_k) = \frac{1}{k}$, $k = 1, \ldots, n$, and $\overline{P}(D_2 \cup D_3) = \frac{2}{3} \neq \max(\frac{1}{2}, \frac{1}{3})$.

In order to find out whether there is a possibility measure compatible with the given assessments, let us first look at the weaker assessment (6). For any normal possibility measure $\Pi$ on $(\Omega, \wp(\Omega))$, (6) is equivalent to

$$\mathrm{N}(D_k) \geq \Pi(D_{k+1}), \quad k = 1 \ldots, n-1,$$

where $\mathrm{N}$ is the necessity measure associated with $\Pi$. $\Pi$ therefore has a distribution of the type $\pi_\alpha$, $\alpha \in [0, \frac{1}{2}]$, given by

$$\pi_\alpha(\omega) = \begin{cases} 1 & ; \quad \omega \in D_1 \\ \alpha & ; \quad \omega \in D_2 \\ 0 & ; \quad \omega \in D_k, k > 2. \end{cases}$$

The greatest such possibility measure has distribution $\pi_{\frac{1}{2}}$, and does not coincide with the natural (least-committal) extension of the given assessments. For any $\alpha$ in $[0, \frac{1}{2}]$, a possibility measure $\Pi_\alpha$ is found that is moreover incompatible with Zadeh's possibility assignment equation: the distribution $\pi_\alpha$ can clearly not be equal to the fuzzy set $\mu_p$ associated with the property $p$, because $\pi_\alpha$ does not distinguish between elements of different $D_k$, $k > 2$, whereas $\mu_p$ should do so, taking into account (3).

By analogy with the case $n = 2$, You could consider adding the assessment $D_1 \approx \Omega$, which imposes the extra condition $\alpha = 0$ on $\pi_\alpha$. In this case, the natural extension of the given assessment is given by $\overline{P}(X) = \overline{P}_{D_1}(X)$, $X \in \mathcal{L}(\Omega)$, and coincides with the natural

[13] This can unfortunately not be done when $p$ is vague.

[14] For more information about conditional upper previsions and coherence, we refer to [24].

extension of the possibility measure with distribution $\pi_0$. But even so, this makes the incompatibility with Zadeh's possibility assignment equation even stronger, and it is not obvious that the extra assessment $D_1 \approx \Omega$ makes as much sense here as it does in the case of a clear $p$ ($n = 2$).

You may also add the extra assessment $\emptyset \approx D_n$. This yields the condition $\Pi(D_n) = 0$, and therefore imposes no additional restrictions on the $\pi_\alpha$, $\alpha \in [0, \frac{1}{2}]$. The natural extension of the given assessment is in this case obtained by replacing $n$ by $n-1$ in (8), and is therefore equal to the natural extension of the greatest possibility measure $\Pi_{\frac{1}{2}}$ only if $n = 3$. It is interesting to note that if $n = 3$, the given assessment $D_1 \succ_{\mathcal{D}} D_2 \succ_{\mathcal{D}} D_3 \approx \emptyset$ corresponds *qua* interpretation to the crude vagueness model given by Gentilhomme [6, 7, 14], which associates a *flou set* $(D_1, D_1 \cup D_2)$ with the vague property $p$ as follows: $D_1$ is the *certain region* (or region of minimal extension), containing those elements which perfectly satisfy $p$; $D_3$ is the *excluded region*, containing the elements which do not satisfy $p$; and $D_2$ is the *flou region*, containing the elements that neither perfectly satisfy nor perfectly do not satisfy $p$. $D_1 \cup D_2$ is called the *region of maximal extension*[15]. The membership function of this flou set is precisely given by $\pi_{\frac{1}{2}}$. We conclude that the use of possibility measures and Zadeh's possibility assignment equation can be corroborated when Gentilhomme's model of flou sets is used in order to represent the vagueness of the property $p$.

Next, we turn to the stronger assessment (7), which for any normal possibility measure $\Pi$ on $(\Omega, \wp(\Omega))$ is equivalent to

$$\left.\begin{array}{l} \mathrm{N}(D_k) \geq \Pi(D_{k+1}) \\ \Pi(D_k) > \mathrm{N}(D_{k+1}) \end{array}\right\}, \quad k = 1 \ldots, n-1.$$

It is easily verified that this condition cannot be satisfied for $n > 3$. If $n = 3$, the possibilistic solutions are again given by $\Pi_\alpha$ with distribution $\pi_\alpha$, $\alpha \in [0, \frac{1}{2}]$, so that the discussion in the previous paragraph applies here as well.

We may conclude that the assessments (6) and (7) are somewhat too strong to allow for a reasonable possibilistic model, unless $n \leq 3$. We can associate a weaker assessment with the given information '$\xi$ is $p$' as follows. Instead of the partition $\mathcal{D}$, we consider the set of nested sets $\mathcal{N} = \{N_1, \ldots, N_n\}$, defined by

$$N_k = \bigcup_{\ell=k}^{n} D_\ell, \quad k = 1, \ldots, n.$$

Note that $N_{k+1} \subseteq N_k$, $k = 1, \ldots, n-1$, and that $N_1 = \Omega$. It seems reasonable to make the following assessment[16]

$$N_1 > N_2 > \ldots > N_n, \tag{9}$$

since $N_k$ contains elements which satisfy $p$ strictly better than any element in $N_{k+1}$ does, $k = 1, \ldots, n-1$. For any normal possibility measure $\Pi$ on $(\Omega, \wp(\Omega))$, the condition $N_1 >_\Pi N_2 >_\Pi \ldots >_\Pi N_n$ is satisfied iff $\Pi(D_k) > 0$, $k = 1, \ldots, n$. The natural extension of the greatest such possibility measure is clearly given by the vacuous prevision $\overline{P}_\Omega$, and is therefore also the natural extension for the given assessment. The requirement (9) turns out to be very weak indeed! This alternative type of assessment leaves more room for using possibility measures and Zadeh's possibility assignment equation, but on the other hand, the chosen possibility measure will differ from the natural extension, unless $n = 2$ and $p$ is clear.

Adding the extra assessment $\emptyset \approx N_n$ yields the additional condition $\Pi(D_n) = 0$, which does not appreciably change the situation, unless $n = 2$, in which case we get the same result as for the assessment $D_1 > D_2 \approx \emptyset$, or $A_p > \overline{A_p} \approx \emptyset$, discussed above.

This might lead to the suggestion that the comparative probability assessment (9) is too weak to lead to possibility measures as the best choice for representing linguistic information. An obvious next step is to consider *numerical* upper probability assessments on nested sets. We have indeed mentioned in Section 5 that under some fairly unrestrictive conditions, the natural extension of a numerical upper probability assessment on nested sets is a possibility measure, and therefore the best choice for this type of assessment.

Moreover [8], suppose we have a membership function $\mu_p\colon \Omega \to [0,1]$ associated with the vague property $p$. Consider the family $\mathcal{N}$ of nested sets $D_x^{\mu_p} = \{\omega \in \Omega \mid \mu_p(\omega) \leq x\}$, $x \in [0,1]$, and the upper probability assessment

$$\overline{P}_p(D_x^{\mu_p}) = \sup_{\mu_p(\omega) \leq x} \mu_p(\omega), \quad x \in [0,1], \tag{10}$$

The natural extension of $(\Omega, \mathcal{N}, \overline{P}_p)$ coincides on $\wp(\Omega)$ with the possibility measure with distribution $\mu_p$, which corroborates both the use of possibility measures and Zadeh's possibility assignment equation for this type of assessment.

On the other hand [8], consider the partition $\mathcal{D}$, with as partition classes the nonempty sets among $L_x^{\mu_p} = \{\omega \in$

[15] We give English translations of the original French terms, which are, respectively, *ensemble flou*; *zone sûre*, *zone d'extension minimum*; *zone exclue*; *zone floue*; *zone d'extension maximum*.

[16] The weaker version involving $\geqq$ instead of $>$ does not impose any restriction, due to the nestedness of the elements of $\mathcal{N}$ and the monotonicity of coherent comparative probability orderings [24].

$\Omega \mid \mu_p(\omega) = x\}$, $x \in [0,1]$, and the upper probability assessment

$$\overline{P}_p(L_x^{\mu_p}) = x, \quad x \in [0,1] \text{ and } L_x^{\mu_p} \neq \emptyset.$$

The greatest (and the only) possibilistic extension is the possibility measure with distribution $\mu_p$, but it is generally different from the natural extension of $(\Omega, \mathcal{D}, \overline{P}_p)$, so that possibility measures and Zadeh's possibility assignment equation are acceptable, but do not seem to be the best choice for this type of assessment.

So anyone who wanted to argue that possibility measures, based on Zadeh's possibility assignment equation, are the best — or even stronger, the only — choice of upper probabilities if we want to represent the linguistic information '$\xi$ is $p$', would have to explain two things. Firstly, why are numerical assessments on nested sets to be preferred above other numerical assessments, or ordinal assessments? Secondly, assuming that numerical assessments on nested sets are indeed to be preferred, why should there be the particular link (10) between the numerical upper probability assessment and the membership function associated with the property $p$?

On the other hand, even if it is not claimed that possibility measures are the best choice always, and if it is conceded that, depending on the type of assessment, linguistic information may have many types of representations, the existing link (10) between upper probabilities and membership functions for nested set assessments remains, to say the least, very interesting. This brings us to the crucial — and we believe as yet unanswered — question whether a behavioural justification can be given for (10) — and therefore also for Zadeh's possibility assignment equation.

As a conclusion to this section, we may state that reasonable assessments reflecting the given linguistic information do not, unless in some special cases, as a rule lead to possibility measures as the only or best choice for representing the information. Moreover, we have seen that the best choice (natural extension) depends strongly on the type of assessment.

## 7 CONCLUSION

The results in this paper suggest that it is indeed possible to model linguistic information in a behavioural context, using the theory of imprecise probabilities. They also indicate that possibility measures can be suitable behavioural representations of linguistic information. They need, however, not be the only ones. The discussion above suggests that it would be better to allow the whole range of coherent upper previsions (and probabilities) as possible models for linguistic information, and not *a priori* restrict the choice to certain families of upper probabilities, be they possibility or probability measures. Even if we evade the difficult and controversial question whether and to what extent there are different *kinds* of uncertainty, the discussion above is somewhat in conflict with the often-heard claim that linguistic uncertainty (or information) is 'possibilistic rather than probabilistic in nature'.

Nevertheless, and depending on the type of assessment, possibility measures can provide a reasonable description of linguistic information, when they are given the behavioural interpretation of upper probabilities. At this point it remains to be investigated which notions and techniques from the theory of possibility [3, 4, 5, 12, 17, 28] — conditional possibility, possibilistic independence, product possibility measures, to name only a few — can be given a natural interpretation in the behavioural theory of imprecise probabilities, and are for instance compatible with the rationality criteria of avoiding sure loss and coherence, and with the technique of natural extension. Also, more attention should be given to statements in natural language which are more complicated than the simple affirmative sentences studied here. In particular, we think of statements with a qualified predicate (Wietse is not very tall), qualified statements (It is not very likely that Peter is older than 40), and conditional statements (If Linde is very young then it is unlikely that she is tall).

**Acknowledgements**

A few days before the deadline for this contribution, Peter Walley pointed out to me a recent and interesting paper of his [25], where he arrives at more or less similar conclusions in a different way. Unfortunately, I did not have the time nor place to include a discussion of his results.

I would like to thank Cliff Joslyn for the many stimulating discussions we had about subjects very much related to this paper, some of them in picturesque locations (including an art museum), others in more noisy surroundings (such as the Washington DC underground). I am particularly grateful to him for inviting me to present this paper in his workshop.

## References

[1] G. Choquet. Theory of capacities. *Ann. Inst. Fourier (U. Grenoble)*, 5:131–295, 1953–1954.

[2] B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order.* Cambridge University Press, Cambridge, 1990.

[3] G. de Cooman. Possibility theory I: The measure- and integral-theoretic groundwork. *International Journal of General Systems.* in press.

[4] G. de Cooman. Possibility theory II: Conditional possibility. *International Journal of General Systems.* in press.

[5] G. de Cooman. Possibility theory III: Possibilistic independence. *International Journal of General Systems.* in press.

[6] G. de Cooman. Evaluation sets: An order-theoretic approach to fuzzy set theory. *CC-AI*, 8:81–101, 1991.

[7] G. de Cooman. Evaluation sets and mappings — the order-theoretic aspect of the meaning of properties. In E. E. Kerre, editor, *Introduction to the Basic Principles of Fuzzy Set Theory and Some of Its Applications*, pages 159–213. Communication & Cognition, Ghent, 1991.

[8] G. de Cooman and D. Aeyels. Supremum preserving upper previsions. In preparation.

[9] B. de Finetti. *Theory of Probability.* John Wiley & Sons, Chichester, 1974. English Translation of *Teoria delle Probabilità.*

[10] D. Denneberg. *Non-additive measure and integral.* Kluwer Academic, Dordrecht, 1994.

[11] D. Dubois, S. Moral, and H. Prade. A semantics for possibility theory based on likelihoods. Technical Report IRIT/93-51-R, Institut de recherche en informatique de Toulouse, 1993.

[12] D. Dubois and H. Prade. *Théorie des possibilités.* Masson, Paris, 1985.

[13] D. Dubois and H. Prade. Fuzzy sets and probability: Misunderstandings, bridges and gaps. In *Proceedings of the Second IEEE Conference on Fuzzy Systems*, pages 1059–1068, 1993.

[14] Y. Gentilhomme. Les ensembles flous en linguistique. *Cahiers de linguistique théorique et appliquée*, pages 47–65, 1968. Editions de l'Académie de la République Socialiste de Roumanie.

[15] R. Giles. The concept of grade of membership. *Fuzzy Sets and Systems*, 25:297–323, 1988.

[16] E. Hisdal. Are grades of membership probabilities? *Fuzzy Sets and Systems*, 25:325–348, 1988.

[17] C. A. Joslyn. *Possibilistic processes for complex systems modelling.* PhD thesis, State University of New York at Binghamton, Binghamton, NY, 1994.

[18] J. Kampé de Fériet. Interpretation of membership functions of fuzzy sets in terms of plausibility and belief. In M. M. Gupta and E. Sanchez, editors, *Fuzzy Information and Decision Processes*, pages 93–98. North-Holland, Amsterdam, 1982.

[19] R. Kruse, J. Gebhardt, and F. Klawonn. *Foundations of Fuzzy Systems.* John Wiley & Sons, Chichester, 1994.

[20] A. M. Norwich and I. B. Turksen. The construction of membership functions. In R. R. Jager, editor, *Fuzzy Set and Possibility Theory: Recent Developments*, pages 61–67. Pergamon Press, New York, 1982.

[21] A. M. Norwich and I. B. Turksen. The fundamental measurement of fuzziness. In R. R. Jager, editor, *Fuzzy Set and Possibility Theory: Recent Developments*, pages 49–60. Pergamon Press, New York, 1982.

[22] A. M. Norwich and I. B. Turksen. Meaningfulness in fuzzy set theory. In R. R. Jager, editor, *Fuzzy Set and Possibility Theory: Recent Developments*, pages 68–74. Pergamon Press, New York, 1982.

[23] P. Walley. Coherent lower (and upper) probabilities. Statistics Research Report 22, University of Warwick, Coventry, 1981.

[24] P. Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London, 1991.

[25] P. Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83:1–58, 1996.

[26] Z. Wang and G. J. Klir. *Fuzzy Measure Theory.* Plenum Press, New York, 1992.

[27] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[28] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.