Have Fuzzy Sets Anything to Do with Vagueness?

DIDIER DUBOIS[1]

1 Introduction

From their inception, fuzzy sets were introduced by Zadeh [55] with a view to formalize human knowledge in engineering problems. This implies fuzzy sets had somehow to come to grip with some aspects of natural language modeling, and in particular, with the concept of vagueness, i.e., the idea that the extension of some natural language predicates lacks clear truth conditions. The claim that fuzzy sets are a basic tool for addressing vagueness of linguistic terms has been around for a long time. For instance, Novák [38] insists that fuzzy logic is tailored for vagueness and he opposes vagueness to uncertainty.

Nevertheless, in the last thirty years, the literature dealing with vagueness has grown significantly, and much of it is far from agreeing on the central role played by fuzzy sets in this phenomenon. Following Keefe & Smith [42], vague concepts in natural language display three features:

- **The existence of borderline cases**: That is, there are some objects such that neither a concept nor its negation can be applied to them. For a borderline object, it is difficult to make a firm decision as to the truth or the falsity of a proposition containing a vague predicate applied to this object, even if a precise description of the latter is available. The existence of borderline cases is sometimes seen as a violation of the law of excluded middle.
- **Unsharp boundaries**: The extent to which a vague concept applies to an object is supposed to be a matter of degree, not an all-or-nothing decision. It is relevant for predicates referring to continuous scales, like *tall, old*, etc. This idea can be viewed as a specialisation of the former, if we regard as borderline cases objects for which a proposition is neither totally true nor totally false. In the following, we shall speak of "gradualness" to describe such a feature. Using degrees of appropriateness of concepts to objects as truth degrees of statements involving these concepts goes against the Boolean tradition of classical logic.
- **Susceptibility to Sorites paradoxes.** This is the idea that the presence of vague propositions make long inference chains inappropriate, yielding debatable results. The well-known examples deal with heaps of sand (whereby, since adding a grain

[1]The author wishes to thank the referees for their careful reading of the preliminary drafts, for their suggestions and additional references.

of sand to a small heap keeps its small, all heaps of sand should be considered small), young persons getting older from one day to the next, bald persons that are added one hair, etc.

Since their inception, fuzzy sets have been controversial for philosophers, many of whom are reluctant to consider the possibility of non-Boolean predicates, as it questions the usual view of truth as an absolute entity. A disagreement opposes those who, like Williamson, claim a vague predicate has a standard, though ill-known, extension [53], to those who, like Kit Fine, deny the existence of a decision threshold and just speak of a truth value gap [19]. However, the two latter views reject the idea of gradual truth, and concur on the point that fuzzy sets do not propose a good model for vague predicates. One of the reasons for the misunderstanding between fuzzy sets and the philosophy of vagueness may lie in the fact that Zadeh was trained in engineering mathematics, not in the area of philosophy. In particular, vagueness is often understood as a defect of natural language (since it is not appropriate for devising formal proofs, it questions usual rational forms of reasoning). Actually, the vagueness of linguistic terms was considered as a logical nightmare for early 20th century philosophers. In contrast, for Zadeh, going from Boolean logic to fuzzy logic is viewed as a positive move: it captures tolerance to errors (softening blunt threshold effects in algorithms) and may account for the flexible use of words by people [54]. It also allows for information summarisation: detailed descriptions are sometimes hard to make sense of, while summaries, even if imprecise, are easier to grasp [56].

However, the epistemological situation of fuzzy set theory itself is far from being clear. Fuzzy sets and their extensions have been understood in various ways in the literature: there are several notions that are appealed to in connection with fuzzy sets, like similarity, uncertainty and preference [17]. It is indeed natural to represent incomplete knowledge by sets (e.g. of possible models of a knowledge base or error intervals). This is also connected to modal logic accounts of reasoning about knowledge [29], and to possibility theory [58, 15]. So not only does vagueness interact with fuzzy sets, but both interact with notions of uncertainty as well, even within fuzzy set theory, to wit:

- A fuzzy set may account for epistemic uncertainty since it extends the notion of a set of possible values understood as an epistemic state.
- Epistemic uncertainty is gradual since belief is often a matter of degree.
- Sometimes, membership functions may account for an ill-known crisp boundary (the random set view of fuzzy sets [25]) and can then be seen as modeling vagueness in agreement with the excluded middle law.
- Higher order fuzzy sets, such as interval-valued or type 2 fuzzy sets (see [12] for a bibliography) are supposed to capture ill-known membership functions of linguistic categories. This seems to refer again, perhaps more convincingly, to a form of vagueness (this will become clearer in Section 4).

Nevertheless, in his works, Zadeh insists that fuzziness is not vagueness. The term fuzzy is restricted to sets where the transition between membership and non-membership is gradual rather than abrupt. Zadeh [54] argues as follows:

Although the terms fuzzy and vague are frequently used interchangeably in the literature, there is, in fact, a significant difference between them. Specifically, a proposition, p, is fuzzy if it contains words which are labels of fuzzy sets; and p is vague if it is both fuzzy and insufficiently specific for a particular purpose. For example, "Bob will be back in a few minutes" is fuzzy, while "Bob will be back sometime" is vague if it is insufficiently informative as a basis for a decision. Thus, the vagueness of a proposition is a decision-dependent characteristic whereas its fuzziness is not.

Of course, the distinction made by Zadeh may not be so strict as he claims. While "in a few minutes" is more specific than "sometime" and sounds less vague, one may argue that there is some residual vagueness in the former, and that the latter does not sound very crisp after all.

The basic aim of this paper is to tentatively clarify the positioning of fuzzy sets in studies about vagueness. Our thesis is that the issue of gradual membership has little to do with the issue of vagueness of words in natural language. Vagueness refers to some uncertainty of meaning, but the fact that the extension of a predicate is not crisp is distinct from the idea that this predicate is tainted with vagueness. We only admit that a gradual predicate is more likely to be vague than a Boolean one, simply because providing a formal model and a protocol to compute gradual truth values precisely is more difficult than when only truth or falsity is to be decided. In fact we can use fuzzy sets (gradual membership functions) or non-dichotomous representations of sets in contexts where natural language is absent. Membership functions can be constructed from fuzzy clustering procedures, from imprecise statistics, or they can represent utility functions, without requesting interpretations in natural languages.

The rest of the paper is organised as follows: first we revisit the information-oriented view of non-dichotomous representations of sets, previously devised with Esteva, Godo and Prade [11], that enables some form of classification of situations where properties are not perceived to be all-or-nothing. In Section 3, we point out that the gradualness of predicates is perhaps not intrinsic. It may depend on whether we take the point of view of agents asserting statements involving non-Boolean predicates, or receiving them. This kind of contextualisation may be relevant for the study of vagueness. In Section 4, we shall then propose a tentative solution to some controversies about vagueness and the role of fuzzy sets, considering that the gradualness of predicates is distinct from, even if sometimes related to, the issue of uncertainty of meaning. So, epistemic and truth value gap approaches apply as much to membership functions of gradual predicates as to Boolean concepts. We also discuss the situation of Nick Smith's blurry sets and fuzzy plurivaluationism in this framework.

2 An information-oriented setting for non-dichotomous representations of sets

We consider the issue of describing objects by means of properties, a core issue in information sciences. There are a number of situations in which a property or a set of properties lead to a non-dichotomous partition of a set of objects [11]. These situations interfere with the issue of vagueness, but not all of them are motivated by it. Notations adopted are as follows. Consider:

- A finite set of objects or entities, denoted by $\mathscr{O}$
- A finite set $\mathscr{A}$ of attributes $a : \mathscr{O} \to D_a$ each with domain D_a
- A property or predicate P referring to attribute a

For binary attributes and in the case of a clear-cut property P, $D_a = \{y_a, n_a\}$, and the extension of P is $Ext(P) = \{o \in \mathscr{O}, a(o) = y_a\}$. More generally, for a many-valued attribute domain, there is a non-empty subset $Y_P \subset D_a$, called positive domain of P, such that $Ext(P) = \{o \in \mathscr{O}, a(o) \in Y_P\}$.

If $\neg P$ denotes the opposite property let $Ext(\neg P) = \{o \in \mathscr{O}, a(o) \in N_P\}$ for some subset N_P of D_a. Then a property is classical if the two following laws hold:

- the excluded-middle law (EML): $Y_P \cup N_P = D_a$, so that $Ext(P) \cup Ext(\neg P) = \mathscr{O}$;
- the non-contradiction law (NCL): $Y_P \cap N_P = \emptyset$, so that $Ext(P) \cap Ext(\neg P) = \emptyset$.

We consider non-classical properties where EML does not seem to apply (nor possibly NCL). Six scenarii have been considered where properties share the set of objects under concern into three subsets [11]. Here, we refresh this classification into three categories:

1. Gradual properties: dropping the bivalence assumption;
2. Bivalent views of the vagueness of linguistic terms where truth conditions are ill-defined (creating truth value gaps) or ill-known (due to partial ignorance on the position of the threshold between true and false);
3. Limited perception of the human mind whereby some objects and/or attribute values are indiscernible.

For each scenario, we consider whether vagueness is at stake or not.

2.1 Gradual properties

Many properties in natural languages P like *tall, young*, etc. and concepts as well (like *bird, chair*) seem to define an implicit (sometimes complete) ordering on the attribute domain D_a and /or the set of objects they pertain to. No matter how narrow the considered context, there does not seem to exist an arbitrarily precise threshold dictating whether a male human height corresponds to the expression *tall man* or not. As people can be tall *to some extent*, it always seems to be a matter of degree. A gradual property P is defined by a pair $(D_a^P, \geq_P)$, where $D_a^P \subset D_a$ is the support of P, and $\geq_P$ is a partial order on D_a such that

- $a(o) \geq_P a(o')$ means that P applies to o is at least as much as to o';
- $u = a(o) \in D_a^P$ means that o is somewhat P: $\forall u \in D_a^P, v \notin D_a^P, u >_P v$
- $u = a(o) \notin D_a^P$ means that P is clearly false for o, and $\forall u, v \notin D_a^P, u =_P v$

An opposite ordering relation $\geq_{\neg P}$ on D_a referring to the negation of P can be defined by $u \geq_{\neg P} v$ if and only if $v \geq_P u$, where $D_a^{\neg P} = D_a \setminus \{u, \nexists u' >_P u\}$ (the prototypical values for $a(o)$ relative to P are ruled out for $\neg P$). It is not clear whether the ordering $\geq_P$ should be total. Some objects may be incomparable in terms of property P. This is especially true with complex concepts that underlie several dimensions. For instance, if P means *comfortable*, two chairs may be somewhat comfortable to some extent for different reasons (one has a soft back but is too low, the other has a perfect seat height but has a hard back), without leading to preferring one to the other.

Nevertheless, the role of the membership function μ_P of P is to provide a representation of this possibly partial ordering on a totally ordered scale (incomparabilities may be solved by a final choice to buy one of the two chairs). A membership function is a mapping from the attribute scale D_a to a bounded totally ordered set L (with top 1 and bottom 0), $\mu_P(u)$ representing the degree to which object o such that $a(o) = u$ satisfies P. In other words, $a(o) >_P a(o')$ implies $\mu_P(a(o)) > \mu_P(a(o'))$. In particular, D_a^P is the support of the fuzzy set defined by μ_P: $D_a^P = \{u, \mu_P(u) > 0\}$. An early example of membership function was suggested by the American philosopher Max Black in 1937 [1], who called them "consistency profiles" in order to "characterize vague symbols." The generalization of the traditional binary characteristic function has been first considered by H. Weyl [52], who explicitly replaces it by a continuous characteristic function to the unit interval. The same kind of generalization was further proposed in 1951 by Kaplan and Schott [31]. See Dubois et al. [13] for details on these works pioneering the notion of fuzzy sets.

There are some reasons for the presence of gradualness in natural language:

- Some predicates refer to an underlying continuous measurement scale D_a for the attribute (*tall*: height; *young*: age). Such terms behave as if there were no threshold on the real line separating the P's from the non$P$'s (if there were such a threshold, it would be too precise to be cognitively relevant). A natural test for detecting this kind of predicates is to check whether the phrase *very P* makes sense. The property is gradual when the answer is yes. The hedge *very* sounds odd when applied to Boolean properties, like *single* or *major* applied to a person. This hedge test clearly makes sense for predicates referring to the extremities of a measurement scale (like *tall* and *small*) but may fail on predicates like *medium*.[2] Gradual predicates of that kind are simple in the sense that the underlying measurement scale is crystal clear. Then, the use of the unit interval as a truth set L is a just a way of rescaling the attribute domain D_a according to the meaning of P. Truth-functionality for complex propositions involving such fuzzy predicates is mathematically consistent (thus yielding algebraic structures different from a Boolean algebra, as seen in mathematical fuzzy logic [26]) even if not compulsory.

- Some concepts like *Bird, Chair*, underlie a typicality ordering on the set of objects. For instance, a penguin is a less typical bird than a swallow. So when someone speaks of birds in general, people do not think that penguins are being referred to in the first place. In this case, the corresponding sets of attributes is less clear,

[2] As pointed out by a referee. But for those terms, perhaps another hedge like *more or less*, may be used.

and there is no obvious numerical measurement scale that can account for them. So it is much more difficult to come up with a numerical membership function. A mere ordering relation $\geq_P$ may make sense, or, at best, a coarse classification of objects into typical, borderline and clearly unsuitable items. Such concepts are better described by a list of (possibly Boolean) properties, the satisfaction of part of which ensures partial membership of an object to the concept extension. These properties may have levels of priority, some being more important than others for defining the concept. This kind of framework comes close to Formal Concept Analysis [21] and has been studied in detail by Freund [20] who proposes a methodology for deriving a partial ordering $\geq_P$ from the knowledge of a set Δ_P of more or less important properties that make up a concept. A similar idea is outlined in [8], where importance degrees belong to a totally ordered scale. Freund actually considers two orderings of objects relative to a concept: a membership ordering, and a typicality ordering where typicality is viewed as stronger than mere membership and relies on a subset of more or less characteristic properties in Δ_P.

How can the extension of P be defined in such framework?

- One may assume that there is no Boolean extension but a gradual one $\tilde{E}xt(P)$ with membership function μ_P.
- Or one may define $Ext(P)$ to be the set of prototypes of P, i.e. $Prot(P) = \{o \in \mathscr{O} \mid a(o) \text{ maximal according to } \geq_P\}$.
- Or yet one may define $Ext(P) = \{o \in \mathscr{O}, a(o) \in D_a^P\}$, only excluding the clearly irrelevant objects for P.

A precise boundary separating objects such that P holds from those where $\neg P$ holds *does not exist* under the gradual view. Depending on the choice of connectives, EML and CL may hold or not [26]. The two last options lead to a trichotomy of the set of objects. In the second one, there are the prototypes of P, the prototypes of $\neg P$, i.e., $Prot(\neg P) = \{o \in \mathscr{O}, a(o) \notin D_a^P\}$ and the set of borderline cases $\mathscr{O} \setminus (Prot(P) \cup Prot(\neg P))$. In the third option, the borderline cases are objects that are both P and $\neg P$.

The third option can be obtained when the domain D_a is equipped with a distance function $d_a : D_a \times D_a \to [0, +\infty)$. In that case, there is a notion of similarity between objects that leads to a form of gradualness understood as limited deviation from full truth [11]. A similarity relation expressing closeness between objects can be defined as a mapping $S : \mathscr{O} \times \mathscr{O} \to [0,1]$ such that $S(o,o) = 1$ (reflexivity), $S(o,o') = S(o',o)$ (symmetry) and $S(o',o) = 1 \implies o = o'$ (separability). An example is of the form $S(o,o') = \frac{1}{1+d_a(a(o),a(o'))}$. $S(o_1,o_2) > S(o_1,o_3)$ means that o_1 is more similar to o_2 than to o_3. Given a Boolean predicate P pertaining to attribute a, one can propose the following computation of membership degrees: $\mu_P(a(o))$ is the extent to which o is close or similar to some object in $Ext(P)$ (Ruspini [43]). By means of the similarity relation, it is possible to declare a gradual property $\tilde{P}$ all the more satisfied by objects as they are closer to being P according to S. We can proceed similarly for $Ext(\neg P)$ (which is equal to $Ext(P)^c$, since P is Boolean). More precisely:

$$\mu_{\tilde{P}}(a(o)) = \sup_{o' \in Ext(P)} S(o,o') \qquad \mu_{\widetilde{\neg P}}(a(o)) = \sup_{o' \notin Ext(P)} S(o,o').$$

By construction $\mu_{\tilde{P}}(a(o)) = 1, \forall o \in Ext(P)$ and $\mu_{\widetilde{\neg P}}(a(o)) = 1, \forall o \notin Ext(P)$. Noticing that $D_a^{\tilde{P}} = \{u, \mu_{\tilde{P}}(u) > 0\}$, it is clear that $\{o, a(o) \in D_a^{\tilde{P}} \cap D_a^{\widetilde{\neg P}}\}$ forms a generally non-empty set of borderline cases, i.e. the law of contradiction fails while the excluded middle law holds. This view may be related to Weston [51]'s idea of approximate truth as reflecting a distance between a statement and the ideal truth. It is also related to the notion of truthlikeness of Niiniluoto [37] and of similarity-based reasoning as surveyed in [22]. It also emphasizes a prototype-based view of fuzzy sets (considering the membership function of a gradual predicate as induced by a set of prototypes and a similarity relation. See Osherson and Smith [39] for a critical discussion and Zadeh's [59] reply. This view of fuzzy sets is not truth-functional since $\widetilde{P \cap Q}$ will generally differ from $\tilde{P} \cap \tilde{Q}$ (for instance, if $P \cap Q = \emptyset, \widetilde{P \cap Q} = \emptyset$ but $\tilde{P} \cap \tilde{Q}$ may be non-empty).

Under the above view, gradualness, not vagueness specifically, is caused by closeness (the presence of a distance between attribute values making objects with very close descriptions possible). This view conflicts with Smith's claim that closeness is the essence of vagueness [47]. The present thesis is that closeness is a natural source of gradualness (related to a continuous measurement scale for the concerned attribute), and that the increased measurement difficulty for numerical membership grades, compared with Boolean ones, results in a higher propensity of gradual predicates to being perceived as vague. As pointed out earlier, typicality is another source of (ordinal) gradualness not especially accounted for by metric structures, even if relative closeness between objects with respect to their adequacy to a gradual predicate could be rendered by means of preference-difference measurement techniques [2].

Even if the presence of intermediate truth values is considered to be a feature of vague propositions, gradualness understood as above clearly does not cover all issues debated about vagueness. Especially, insofar as a precise membership function is obtained for the property and one admits that there is no underlying ill-known boundary, the gradual extension is perfectly defined and there is no uncertainty about the meaning of property P. In that sense a membership function is a more accurate description of a gradual predicate than a regular characteristic function. While vagueness is a defect, gradualness is an enrichment of the Boolean representation. Likewise a metric space is a richer description than a mere set, so that the use of similarity-based degrees of truth makes a logical description finer: it does not create an anomaly nor a defect, contrary to what vagueness is supposed to do, according to many philosophers. Especially the argument against membership functions, whereby it is a paradox to use precise membership grades, works if the gradual model of the non-Boolean concept under concern is meant to account for uncertainty about membership. Insofar as gradual concepts just display gradualness, the membership function model makes sense. However, the latter is of course an ideal view tailored for gradual properties defined on simple linear measurement scales, and many have pointed out the difficulty to actually come up with the precise membership function, if any, of a gradual property. But the use of a membership function is often a good enough working assumption for information engineers, with no pretence to address philosophical issues, but for the suggestion that for some propositions, truth may intrinsically be a matter of degree. This point also indicates that there is more to vagueness than gradualness.

2.2 Ignorance and truth value gaps

Another point of view on non-classical properties is to admit that there are some objects that for sure satisfy them, others that for sure don't, and still other objects for which it is hard or even impossible to assign a clear-cut truth value. This situation, that is considered as a feature of vagueness, also leads to partition the set of objects into three subsets: $\mathscr{C}(P), \mathscr{C}(\neg P), \mathscr{B}(P)$ forming a partition of $\mathscr{O}$, where $\mathscr{C}(P)$ consists of objects for which P is definitely true, $\mathscr{C}(\neg P)$ of objects for which P is definitely false, and the set $\mathscr{B}(P)$ consists of borderlines cases.

As it seems there are two main views of this kind of paradigm: the truth value gap view of Fine [19] and the epistemic view of Williamson [53].

According to Fine, a proposition *o is P* is said to be supertrue if it is true in all ways of making *P* classical (sharpenings, or precisiations of *P*). This approach is non-classical in the sense that a proposition can be neither supertrue nor superfalse. A precisiation of P is a clear-cut property $P^{\mathscr{S}}$ with extension $\mathscr{S} \subset \mathscr{O}$, in agreement with P on all objects that clearly satisfy P and all objects that clearly falsify it. Hence $Ext(P^{\mathscr{S}})$ is a subset $\mathscr{S}$ of objects such that $\mathscr{C}(P) \subseteq \mathscr{S} \subseteq \mathscr{C}(P) \cup \mathscr{B}(P)$. So *o is P* is super-true if *o is* $P^{\mathscr{S}}$ is true for all $\mathscr{S}$ in this family. Clearly it is equivalent to requiring that *o is* $P^{\mathscr{C}(P)}$ is true.

According to Williamson, a proposition *o is P* is either true or false in borderline cases, it is just that we are not in a position to find out which truth value the proposition takes. The main difference between Fine and Williamson seems to be in some sense metaphysical: whether or not there is a true classical extension. For Fine, this true extension does not exist, and this is precisely the characteristic of vagueness: if o is borderline, i.e., $o \in \mathscr{B}(P)$, there is no truth value for the proposition *o is P*. There is a truth value gap. In contrast, for Williamson, the true classical extension $Ext(P)$ that provides the precise meaning of P exists, but it is ill-known. Vagueness thus consists in this uncertainty of meaning. All that is known is that $\mathscr{C}(P) \subseteq Ext(P) \subseteq \mathscr{C}(P) \cup \mathscr{B}(P)$.

There is in fact yet another view in the same vein called *plurivaluationism* [4]. It accepts Boolean representations of vague concepts, but contends that any sharp definition of the extension of P lying between $\mathscr{C}(P)$ and $\mathscr{C}(P) \cup \mathscr{B}(P)$ is equally good to represent P. Under this view the thick boundary between P accounts for the idea that in practice, there is no need to bother being more precise.

Kit Fine's truth value gap approach is based on so-called "*supervaluations*", first proposed by van Fraassen [49] in the context of incomplete information logics, whereby *supertrue* means true in all complete information states compatible with the available knowledge. Van Fraassen criticizes the loss of the excluded middle and contradiction law in logics of incomplete information like partial logic, where interpretations are changed into partial interpretations. This view is closely related to possibility theory (as discussed in [9]). Define a Boolean possibility distribution $\pi : 2^{\mathscr{O}} \to \{0,1\}$ over possible extensions of P, such that $\pi(\mathscr{S}) = 1$ if and only if $\mathscr{C}(P) \subseteq \mathscr{S} \subseteq \mathscr{C}(P) \cup \mathscr{B}(P)$. The level of certainty that *o is P* is $N(o \in Ext(P)) = \inf_{\mathscr{S}: o \notin \mathscr{S}} 1 - \pi(\mathscr{S})$ [15]. Then, P being supertrue is equated to *o is P* having full certainty, that is, $N(o \in Ext(P)) = 1$. Indeed the latter reduces to

$$\{\mathscr{S}, \mathscr{C}(P) \subseteq \mathscr{S} \subseteq \mathscr{C}(P) \cup \mathscr{B}(P)\} \subseteq \{\mathscr{S}, o \in \mathscr{S}\},$$

which holds only if *o is P* is supertrue (i.e., again $o \in \mathscr{C}(P)$).

Supervaluation was introduced to cope with incomplete information, not vagueness. As a complete state of information refers to a precise description of the actual world, it is clear that in the original supervaluation framework, the correct complete state of information exists, so that it sounds coherent with Williamson's ideas. Interestingly, Kit Fine seems to borrow the supervaluation machinery while doing away with this assumption.

These approaches, including plurivaluationism, may be viewed as attempts to reconcile vagueness and the laws of classical logic. For supervaluationists and plurivaluationists, the latter hold for each sharpening of *P*; for epistemicists, they hold for the "real" extension of *P*. By construction, the proposition *o is P or not P* is, accordingly, supertrue or certainly true as, whatever the choice of $\mathscr{S}$ as a substitute for $Ext(P)$, $\mathscr{S} \cup \neg\mathscr{S} = \mathscr{O}$. Likewise, *o is P and not P* is superfalse or certainly false.

This kind of view can go along with the presence of a membership function, even if its meaning will be different from the case of the gradual setting. For instance one can explain the truth value gap by the disagreement between people as to what is the true extension of *P*.

Suppose we get different crisp representations of *P* provided by a set of *n* agents. Each agent *i* provides a partition (Y_P^i, N_P^i) of D_a. Then the trichotomy $\mathscr{C}(P), \mathscr{C}(\neg P)$, $\mathscr{B}(P)$ is retrieved letting

- $\mathscr{C}(P) = \{o \in \mathscr{O}, a(o) \in \cap_{i=1,\ldots,n} Y_P^i\}$
- $\mathscr{C}(\neg P) = \{o \in \mathscr{O}, a(o) \in \cap_{i=1,\ldots,n} N_P^i\}$

"*o is P*" is super-true (false) if it is true (false) for all agents. Otherwise the Boolean truth value is not defined. Instead, one may define a membership function of the form:

$$\mu_P(o) = \frac{|\{i, a(o) \in Y_P^i\}|}{n}.$$

This kind of fuzzy sets clearly expresses variability across agents and will not be truth-functional. This membership function is the one-point-coverage function of a random set, and is not meant to express gradualness.

The above protocol may be hard to follow as people may be reluctant to provide clear-cut subsets of the attribute scale, or even to exhaustively describe the extension of a predicate. There have been some more realistic experiments carried out to identify membership functions of gradual or vague predicates by asking individuals in a group to classify objects as being *P* or not, enforcing a clear-cut reply. The membership grade $\mu_P(u)$ is then interpreted as a conditional probability [28]:

$$\mu_P(u) = Prob(\text{asserting "}o \text{ is } P\text{"} | o(a) = u).$$

Again it represents variability, not gradualness [30]. Nevertheless, if a property is intrinsically gradual, and a clear-cut decision is artificially enforced in the experiment, one may assume that the more *o is P* (in the gradual view) the more likely (in the probability sense) *o* will be classified as *P* by agents. At the theoretical level, this is also the path followed by Scozzafava and his group [3] to interpret membership functions. The two above approaches are compatible if each agent *i* declares "*o is P*" whenever $u \in Y_P^i$.

Recently, Lawry [34] has built an extensive theory of what he calls "label semantics", based on the voting approach and the bivalence assumption that for each voting individual a label P (the name of a property on an attribute domain) applies or not to a given object. It also obviates the need to use attribute domains and extensions. Given an object o and a set of labels Λ, a mass function m_o in the sense of the transferable belief model of Smets [45] is defined on Λ. The mass $m_o(T)$ of a set of labels $T \subseteq \Lambda$ is the proportion of individuals that consider the object o to be properly qualified by the set of labels T. The appropriateness (membership grade) of label P to object o is then defined as $\sum_{P \in T} m_o(T)$. This view of vagueness is clearly in agreement with Williamson's idea of an unknown crisp extension of vague predicates, but it is agreed that this crisp description may vary across individuals and is not a presupposed objective entity. More recently, Lawry and Gonzalez-Rodriguez [35] have extended the label semantics, moving from binary to three-valued truth (individuals decide between true, false and borderline) while keeping the same experimental setting. Their work emphasises the point that the three truth values are a matter of representation convention and should not be confused with the uncertainty due to vagueness, the latter being expressed by a hesitation between the three truth values in the three-valued framework.

2.3 Limited perception

Numerical measurement scales and continuous mathematical models often provide formal representation settings that are much more refined than what human perception can handle. Measurement scales are often assumed to be arbitrarily precise, and the unit interval chosen as the set of possible membership grades is a good example such an excessively refined modeling choice. The continuity assumption is in contradiction with the limited perception capabilities of the human mind. Hence, for gradual properties with continuous attribute scales, one may argue that people cannot distinguish between very close attribute values, hence between objects whose description is almost the same. Some philosophers like R. Parikh [40] argue that one reason for vagueness maybe the difficulty to perceive the difference between close values in D_a: if $d_a(u,v) \leq \varepsilon$ then u is perceived as being the same value as v. Even if there is an actual standard extension of property P, two objects o and o' such that $o \in Ext(P), o' \in Ext(\neg P)$ will be perceived as borderline for P whenever $d_a(a(o),a(o')) \leq \varepsilon$, where ε is the perception threshold in D_a. Even if the boundary of the extension of P exists, it will be perceived as a thick area of borderline objects (of width 2ε).

In other words, there is a reflexive and symmetric indiscernibility relation I on D_a hence on $\mathscr{O}$, defined by uIv if and only if $d_a(u,v) \leq \varepsilon$. So the continuous scale D_a is not the proper space for describing properties pertaining to attribute a. Each element $u \in D_a$ is actually perceived as the subset $G_a(u) = \{v, uIv\}$ of elements that cannot be told apart from u. Each subset $G_a(u)$ is called a granule around u. The perceived attribute scale is a subset of the set of possible granules of D_a, that is a family $\mathscr{G}_a \subset \{G_a(u), u \in D_a\}$, that forms a covering of D_a (and oftentimes, just a partition).

This perception limitation induces indiscernibility on the set $\mathscr{O}$ of objects. Any object o cannot be told apart from objects in $[o]_a = \{o', a(o') \in G_a(a(o))\}$, and even, using a granular scale, $\mathscr{G}_a$, from objects whose attribute value belongs to the same granule in $\mathscr{G}_a$. As a consequence any standard predicate on D_a corresponds to an ill-defined subset of objects in $\mathscr{O}$, that can be modelled as follows:

- The set $\mathscr{C}(P) = Ext(P)_* = \{o \in \mathscr{O}, [o]_a \subseteq Ext(P)\}$ is the set of objects that are clearly P.
- The set $\mathscr{P}(P) = Ext(P)^* = \{o \in \mathscr{O}, [o]_a \cap Ext(P) \neq \emptyset\} \supseteq \mathscr{C}(P)$ is the set of objects that are possibly P.
- The set $\mathscr{B}(P) = \mathscr{P}(P) \setminus \mathscr{C}(P)$ is the set of borderline objects for P.
- The set $\mathscr{C}(\neg P) = Ext(\neg P)_* = (Ext(P)^c)_* = (Ext(P)^*)^c$, i.e., the complement of $Ext(P)^*$ is the set of objects that are clearly not P.

This approach was formalised by Williamson [53, appendix] in his logic of clarity. It is also a special case of imprecise information handling in the representation of objects [14, 5]. Namely, one may define a multivalued mapping Γ from $\mathscr{O}$ to D_a, namely, $\Gamma(o) = G_a(a(o))$, and it is easy to see that $\mathscr{C}(P)$ is the lower inverse image, via Γ, of the positive domain of P in the sense of Dempster [7]: $\mathscr{C}(P) = \{o \in \mathscr{O}, \Gamma(o) \subseteq Y_P\}$, where $\Gamma(o)$ is the set of possible values of o. The set $\Gamma(o)$ represents what is known about the attribute value $a(o)$. Here the lack of knowledge is due to limited perception.

Note that in the case of modeling vagueness due to indiscernibility, one can have access to $\mathscr{C}(P)$ and $\mathscr{C}(\neg P)$ (individuals can point out objects that are definitely P or not), but not to the real extension of P (nor the positive domain Y_P). Moreover in this interpretive setting, there is a reductio ad infinitum effect because it is clear that the crisp boundary of $\mathscr{C}(P)$ cannot be precisely perceived either, since $(Ext(P)_*)_* \subset Ext(P)_*$, generally. This is because the reflexive and symmetric relation I is not transitive. It is transitive if the granulation of the attribute scale D_a is done via quantization, that is using a partition of D_a whose elements are perceived as distinct. This is a very common method for interfacing numerical attribute scales and symbolic ones. Then I is an equivalence relation, and the approximation pair $(Ext(P)_*, Ext(P)^*)$ defines a rough set [41], such that $(Ext(P)_*)_* = Ext(P)_*$. Halpern [27] tries to reconcile, in a modal logic framework, the intransitivity of reported perceptions (by sensors), and the transitivity of subjective perceptions, noticing the former are based on measurements of the real state of the world, while the perceived appropriateness of vague predicates depends on the state of the perceiving agent at the moment of the perception experiment.

The limited perception approach can actually go along with gradual truth if we admit that two attribute values may be distinguishable to a degree, all the higher as the values are far away from one another. In this case, the indiscernibility relation on the attribute scale is a fuzzy symmetric and reflexive relation, an example of which is: $I(u,v) = \min(1, \frac{\varepsilon}{d_a(u,v)})$. Note that the separability property of similarity relations used in Section 2.1 does not hold since the stress here is on limited perception. See Klawonn [33] for a detailed overview of this approach. In that case, indistiguishability granules become gradual, with membership values $\mu_{G_a(u)}(v) = I(u,v)$, and the set $\mathscr{P}(P)$ becomes the gradual extensional hull of the crisp extension $Ext(P)$: its membership function becomes $\mu_{\mathscr{P}(P)}(a(o)) = \sup_{o' : a(o') \in Ext(P)} I(a(o), a(o'))$.

2.4 Connectives

Boolean connectives can be extended to the gradual setting and remain truth-functional, extending truth tables to more than two truth values, including the unit interval. The price paid is clearly a loss of properties with respect to the Boolean algebra [26]. When the gradualness of properties is due to some tolerance, modelled by a distance function, and applied to basically Boolean properties, the crisp sets are the "real" things while the fuzzy sets are their relaxed representations. One can observe a lack of truth-functionality as indicated in Section 2.1. This is because in this case there are not more so-generated fuzzy sets than crisp sets, so that the underlying algebra is Boolean, and truth-functionality on a continuous truth-set is at odds with the Boolean structure. Connectives are not truth-functional either under the epistemic and truth value gap or plurivaluationist semantics, but the Boolean nature of extensions (or precisiations) is retained. This is not surprising because truth-functionality is always lost under partial ignorance (see [9] for a detailed discussion). The third approach based on limited perception is a also a case of incomplete information due to indiscernibility between close values. The truth-functionality is then again lost. In other words truth-functionality is maintained only if the extension of properties is considered intrinsically gradual and no reference is made to an underlying Boolean property either made flexible so as to cope with closeness, or blurred due to limited perception. Most models of vagueness pointed out above lead to a loss of truth-functionality, but for the pure gradualness situation.

2.5 How to tell vagueness from gradualness?

Even though the above discussion points out a distinction, and possibly some independence, between the idea of vagueness and the idea of gradualness of concepts, it is noticeable that the three features of vagueness considered by Keefe and Smith and recalled in the introduction are of little avail to tell one from the other. Clearly, among these three key-features, the first one, i.e., the presence of borderline cases, seems to suggest a three-valued logic; besides, admitting gradual truth leads to consider as borderline cases situations where fuzzy propositions take truth values different from true and false. The second property seems to directly propose that truth might come by degrees. But several approaches, recalled above, view borderline cases as an effect of the partial knowledge of a crisp boundary, or derive numerical membership functions that represent uncertainty about Boolean truth, and not intrinsic gradualness of concepts. The presence of the Sorites paradox does not seem to solve the dispute between bivalence and multivalence. Indeed, the Sorites paradox has been given plausible explanations both in the setting of fuzzy logic (originally by Goguen [24]) via a graceful degradation of the truth values of the conclusion, and by the epistemic, yet Boolean, approach to vagueness (as discussed by Williamson [53]) as a result of uncertainty pervading the meaning of concepts. One way out of this difficulty consists in acknowledging a difference of nature between vagueness and gradual truth. The gradualness of concepts may perhaps be observed (a person can rank objects in terms of their appropriateness as instances of a concept), but gradual truth is a representation convention: we decide to define a property as liable to have two or more truth values when applied to object for a reason of modeling convenience. Gradual truth looks natural and easy to quantify for some concepts (like *tall*), natural but more difficult to quantify for other ones (like *beautiful* or *bird*) and very debatable for yet others (like *single*). In contrast, vagueness is a phenomenon observed in the way people use language, and is characterized by, as Halpern [27] says, variability in the use of some concepts both between and within speakers. It may be that one cause of such variability is the gradual perception of some concepts or some words in natural language. However, one should consider gradual truth more as a modeling assumption than as an actually observable phenomenon (it makes no sense to ask people for numerical membership grades, or truth degrees).

3 Is there a threshold underlying a gradual concept?

At this point we are led to the question whether some properties are intrinsically gradual so as to lead to a notion of gradual truth for statements involving such properties. The advocates of the truth value gap and the epistemic view deny the possibility of gradual predicates, and consider their approaches as undermining the arguments in favor of many-valued truth. Yet, it is hard to believe that, talking about the height of people there is an infinitely precise (but unknown) threshold dividing the height range between tall heights and non-tall heights (assuming the context is made precise enough). Instead of taking sides, it sounds more reasonable to try and reconcile the epistemic and the gradual views on fuzzy concepts. In fact, what is embarrassing about the claims made on all sides is that they seem to consider the meaning of vague properties and the existence or not of classical extensions *in abstracto*, regardless of the way people use such properties. We suggest that the choice between an ill-known crisp extension and a gradual extension of a vague property (say *tall*) depends on the role of the agent with respect to a statement containing this vague property (say *Jon is tall*). It is interesting to consider two very distinct situations:

- **Asserting a gradual statement**: the case when an agent declares "*Jon is tall*". This claim is unambiguous so the utterer must use an implicit decision threshold.
- **Receiving a gradual statement**: the case when an agent receives a piece of information of the form "*Jon is tall*". There is no need for any threshold if the gradualness of the meaning is acknowledged by the receiver.

3.1 Asserting a gradual statement

The fact of asserting a statement "*o is P*" is a Boolean event (the agent asserts it or not) whether the statement is vague, gradual, or not. In the case where P is a gradual property, everything occurs as if the decision of asserting P were made by means of a clear-cut positive domain $Y_P \subset D_a$, whereby "*o is P*" is asserted because $a(o) \in Y_P$. For instance, if an agent declares that *Jon is tall*, and *tall* is a gradual predicate on the human height scale, say a real interval D_a, then there must be a threshold α in $[0,1]$ such that statement "*Jon is tall*" was asserted because $\mu_P(Jon) \geq \alpha$. This is equivalent to claim that everything happens as if there were a threshold $\theta \in D_a$ such that $height(Jon) \geq \theta$. We say "everything happens as if" because we certainly do not have access to this threshold when we are told that *Jon is tall* (if we know the actual height $height(Jon) = h$, all we know is that $\theta \leq h$). Even more, it is not clear whether the utterer is aware of this threshold. It may vary with time (the next time the same individual utters the same statement, another threshold may be used), let alone with the identity of the utterer, and some other kind of circumstances (the people that were seen before Jon showed up) that makes the perception of tallness a subjective matter, and not only a function of height [27]. In any case this threshold (that makes a gradual property P temporarily crisp) is ill-known, utterer-dependent, possibly time-dependent. It is not intrinsic.

This model somewhat goes along with the epistemic view, while putting it in a very pragmatic setting. The threshold view seems to have been adopted by linguists like Kennedy [32]. The latter studies adjectives gradable on a numerical scale and comments at length on the nature of such a threshold. In the basic model, it is supposed to reflect "the average degree to which the objects in the comparison class possess the property", and the author considers more general kinds of context-sensitive thresholds acting as comparison standards justifying the utterance of gradual propositions. Kennedy makes the distinction between absolute gradable adjectives (like *open* and *closed* for a door) that admit minimum and maximal standards, and relative ones (like *tall* for a man). When asserting *the door is open* we just need a minimal aperture of the door; when asserting *the door is closed* we generally mean fully closed. Representing *open* and *closed* by membership functions, it is clear that the decision threshold in these cases are respectively such that $\mu_{open}(\theta) = 0$ (the statement is uttered for doors inside the support of the fuzzy set), resp. $\mu_{close}(\theta) = 1$ (resp. inside the core). For those adjectives, the similarity-based explanation of gradual membership degrees (tolerance with respect to a norm) looks plausible. Kennedy indicates that absolute gradable adjectives do not trigger the Sorites paradox. The threshold generally corresponds to another cut level $\mu_P(\theta) = \alpha \neq 0, 1$ for relative gradable adjectives, and the variability of this threshold explains the perceived vagueness of such adjectives.

In practice, when asserting a gradual statement, gradual truth can be temporarily dispensed with, and membership degrees may then just reflect the probability that a label is appropriate to an object for the agent (like in the voting paradigm) [34, 3]. Note that the Boolean quality of the asserted proposition is only a convention adopted by the utterer. We may ask for a more refined convention, asking the agent to choose between true, false and borderline, like Lawry and Gonzalez-Rodriguez [35]. However, as said earlier, it does not look sensible to ask questions referring to a finer classification, let alone to request numbers between 0 and 1.

The implicit threshold involved in the utterance of a statement involving a gradual predicate like *tall* is more likely to be subject to variability among and within agents than the threshold used to define a crisp predicate, as the latter is often part of the definition of the concept (for instance *major* is defined by a legal age threshold). This variability is characteristic of gradual predicates understood as being vague.

3.2 Receiving a gradual statement

Generally people have a good idea whether a property is gradual or not, as suggested in Section 2.1. The predicate *tall* is clearly gradual since an individual can be more or less *tall*. Hence, taking the point of view of the receiver, the latter, upon hearing that *Jon is tall* has been asserted, already knows that this property P is gradual. Its acceptance as a valid information presupposes that the receiver believes the utterer, but it does not require any decision, on the side of the receiver, pertaining to an underlying threshold separating *tall* and *non-tall*. The only decision made by the receiver is that of accepting the statement as reliable enough (if any threshold is involved in the receiver's decision to take the utterer's statement for granted, it is a reliability threshold pertaining to the source of information). Since the receiver acknowledges the gradualness of the property, one may assume, insofar as he knows the meaning of P and the attribute scale is clear and simple (like the height scale), that this meaning can be approximated by a membership function. It can be done by pointing out prototypes of P and $\neg P$ and a simple interpolation is often enough for practical purposes. Another assumption is needed for the receiver to understand the information conveyed by the utterer properly: the receiver must assume that his/her membership function of P is close enough to the one of the emitter (which means, of course, the receiver is aware of the context in which the statement is uttered).

If these assumptions are met, the information about the height of Jon boils down to a set of more or less possible values modelled by the membership function μ_P on D_a. This membership function is interpreted as a possibility distribution π, whereby $\pi(u)$, the degree of possibility that $height(Jon) = u$ is equated to the membership grade $\mu_P(u)$ [58]. This possibility distribution represents the epistemic state of the receiver only knowing that *Jon is tall*. In this view, saying that the statement *Jon is tall* is true for the receiver means that the latter considers it as a totally safe assertion that can be fully trusted. It does not mean that for the receiver, $\mu_P(height(Jon)) = 1$; it means $\pi = \mu_P$. In case the receiver does not trust the utterer, one may model this kind of unreliability by means of certainty-qualification [15], for instance $\pi = \max(\mu_P, 1-r)$, where $r \in [0,1]$ is the degree of the receiver's certainty that the information is reliable. This is similar to the discounting of testimonies in Shafer's theory of evidence [44].

This piece of information may be used to expand or revise the prior epistemic state of the receiver [16]. Note that even the absolute gradable adjectives in the sense of Kennedy [32] can be interpreted in a gradual way, by introducing some tolerance. The major interest for the receiver to interpret statements of the utterer in a gradual way could be to eventually solve conflicts in the information thus collected: conflicts between gradual representations can be gradual too, and lead to a compromise solution, while inconsistent bodies of crisp representations of gradual propositions are harder to handle. In other words, the use of membership functions of fuzzy sets to represent gradual properties, as proposed by Zadeh [58], makes sense and is even quite useful from the point of view of receiving and processing information, not so much from the point of view of asserting propositions involving gradual predicates. Also, note the opposite nature of the respective situations of the emitter and the receiver. The emitter has enough knowledge about the height of Jon to be able to assert *Jon is tall*, making this statement true in the Boolean sense. In contrast, the information item *Jon is tall* maybe the only one obtained so far by the receiver (who thus has no prior idea of the actual height of Jon nor about the threshold used by the utterer to make his statement). Viewed from the receiver, the "truth" of the utterer's statement means that he is ready to adopt the membership function μ_P as a possibility distribution representing his epistemic state about the height of Jon. So, the receiver does not even handle degrees of truth, but only possibility degrees: $\mu_P(u)$ is the degree to which the receiver, accepting the information item, considers *Height(Jon)* = *u* plausible.

One way to cope with some controversies between fuzzy set theory advocates and vagueness philosophers is to consider they do not study the same problem at all. The latter are interested in how people decide to use vague predicates when they speak, the former are concerned with the modeling and the storing of information coming from humans. In that sense, the vagueness and the gradualness of natural language terms can be viewed as orthogonal concerns.

4 Toward a reconciliation of some views of vagueness

In this section, we suggest another reason why the epistemic and truth value gap approaches to vagueness should be seen as compatible with a non-bivalent view of some properties or predicates. Indeed, there seems to be a strong historical tradition for Bivalence in logic. The status of truth in philosophy is so prominent that it is taken as an objective notion, whose perfection cannot go along with shades of truth. Especially, the existence of a decision threshold in the epistemic view sounds like a "realistic" point of view a la Plato. Yet, as pointed out by De Finetti [6] in an early paper discussing Łukasiewicz logic, the bivalence of propositions can be viewed as a representational convention, not at all a matter of actual fact:[3]

> Propositions are assigned two values, true or false, and no other, not because there "exists" an a priori truth called "excluded middle law", but because we call "propositions" logical entities built in such a way that only a yes/no answer is possible

Hence, we may infer that more generally, truth values are always a matter of convention. Gradual truth is another convention, different from the bivalent one. It is instrumental in the faithful representation of the meaning of some terms in natural language pertaining to numerical attribute scales, in order to process information. So adopting gradual truth as the proper convention for representing predicates like $tall, old$ and the like, one may again consider the issue of vagueness as the presence of cases where truth values can hardly be assigned to propositions applied to some objects. But instead of considering borderline cases as those where it is not clear whether a vague proposition is true of false, one may consider that in borderline cases, not only truth is gradual but the assignment of a gradual truth value is difficult or impossible.

4.1 Ill-known fuzzy sets and gradual truth value gaps

Adopting this stance, the existence of intrinsically gradual properties no longer contradicts the epistemic view of vagueness. The epistemic thesis assumes that vague predicates have clear-cut extensions even if there is some limited knowledge about them. Consider the claim made by Williamson [53, p. 201]:

> Bivalence holds in borderline cases. It is just that we are in no position to find out which truth value the vague utterance has.

The last part of the claim may hold for gradual properties, without requesting bivalence as a prerequisite. Even if a gradual predicate should rather be modelled by a fuzzy set

[3] Our translation from the French.

than a crisp set, its membership function has little chance to be well-known. A *gradual epistemic view* could postulate that the membership function of a gradual predicate exists but one is partially ignorant about it, hence the vagueness phenomenon.

As it looks much more difficult to define membership functions of gradual predicates precisely than crisp extensions of clearly bivalent ones, it is natural that most gradual concepts sound more usually vague than crisp ones. In other words, the vagueness of a fuzzy concept could be modelled via intervals of truth values, or even fuzzy sets thereof, the latter representing knowledge about an albeit precise but ill-known gradual truth value [57]. In a nutshell, we could argue that vagueness is due to partial ignorance about the meaning of categories irrespective of their being considered as gradual or not. This thesis disentangles the issue of gradual propositions from the problem of vagueness.

The way interval-valued fuzzy sets (IVFs, for short) are used in the fuzzy set community ([50], for instance) seems to be at odds with the epistemic view of vagueness, even with the above non-bivalent stance. An interval-valued fuzzy set *IF* is defined by an interval-valued membership function: $IF(u) = [\mu_*(u), \mu^*(u)], \forall u \in D_a$. Under the epistemic view, there exists a real membership function $\mu \in IF$, i.e., $\mu_*(u) \leq \mu(u) \leq \mu^*(u)$.

However, interval-valued fuzzy sets are construed truth-functionally. The union, intersection and complementation of IVF's are obtained by canonically extending fuzzy set-theoretic operations to interval-valued operands in the sense of *interval arithmetics*. For instance, restricting to the most commonly used connectives, with $IF(u) = [\mu_*(u), \mu^*(u)], IG(u) = [\nu_*(u), \nu^*(u)]$:

$$IF \cap IG(u) = [\min(\mu_*(u), \nu_*(u)), \min(\mu^*(u), \nu^*(u))];$$

$$IF \cup IG(u) = [\max(\mu_*(u), \nu_*(u)), \max(\mu^*(u), \nu^*(u))];$$

$$IF^c(u) = [1 - \mu^*(u), 1 - \mu_*(u)].$$

IVFs are then viewed as special case of L-fuzzy sets in the sense of Goguen [23] where L is a set of intervals on [0, 1]. Hence, interval-valued fuzzy sets have a weaker structure than the fuzzy set algebra of precise values they extend. However, just as the epistemic view on vague predicates insists that they remain bivalent whether truth or falsity can be decided or not, so that the properties of classical logic should be preserved, the epistemic view on vague gradual predicates maintains that the algebraic structure of precise truth values should be preserved even if ill-known. For instance, if the conjunction of fuzzy sets is performed using the minimum, the weak form of the contradiction law ($\min(\mu(u), 1 - \mu(u)) \leq 0.5$) should hold for gradual propositions, while it does not hold for IVFs with the above definition of intersection and negation of interval-valued fuzzy sets (see [10] for a more detailed discussion). In this approach, truth values are precise but ill-known. The price for preserving the tautologies of the underlying many-valued logic is as usual a loss of truth-functionality. This would allow a non-truth-functional epistemic approach to gradual concepts, such that propositions have precise gradual truth values, however pervaded with uncertainty. A logical approach to this view is described by Lehmke [36], for instance.

Likewise, the supervaluationism of Kit Fine could be accommodated in a gradual setting: for instance we could define a vague statement to be super-α-true if it is at least

α-true in all of its gradual precisiations (precise membership functions). Restricting the truth set to three values, one may likewise consider propositions that are "super-borderline" (i.e., they are neither totally true nor totally false in all three-valued precisiations of the membership function). In this approach, no assumption is made about the existence of a "true" membership function. However, in the same way as in the classical approach, where all classical models compatible with the non-classical supervaluationist model must be used to check super-truth, all membership functions $\mu \in IF$ are to be used for checking super-α-truth, even if none of them is the true one. In the interval-fuzzy set setting, one may say that *o is F* is

- super-α-true when $\mu_*(u) \geq \alpha$,
- super-α-false when $\mu^*(u) \leq 1-\alpha$.

The supervaluationistic approach to vagueness proposed by Kit Fine could thus extend to gradual truth values, making it closer to Smith's fuzzy plurivaluationism [48]. In the latter view though, in contrast with the former, all precise membership functions are equally acceptable, and one lacks reason for choosing between them due to semantic indecision, or as Smith puts it "the meaning-determining facts" preventing us from choosing a unique intended model. Both fuzzy supervaluationism and fuzzy plurivaluationism would again lead to a non-truth-functional calculus of IVFs in order to preserve the algebraic properties of the underlying set of precise membership functions.

4.2 Blurry sets

At this point, it is interesting to examine the positioning of the so-called blurry set approach to vagueness proposed by Nick Smith [46]. Smith advocates the idea that if a property is gradual, propositions referring to it should have gradual truth values. However, like in this paper, he considers simple membership functions valued on the unit interval are insufficient to account for vagueness. More precisely, for him, numerical values are only approximations of actual truth values. Smith proposes to represent truth values by means of so-called degree functions, which are probability-like kinds of fractal constructs. The idea is that if someone says *Jon is tall is 0.6 true*, we call assertion A_1, then we should allow degrees of truth for assertion A_1, say an assertion A_2 of the form A_1 *is 0.3 true* to make sense, and so on, ad infinitum. However, rather than seeing this construction as a hierarchy of assignments of simple numerical truth values, Smith views it as the *single assignment* of a complex truth value consisting of a so-called *degree function*.

Roughly speaking a degree function is a mapping f from arbitrary sequences of real values in the unit interval to the unit interval. For instance if f pertains to the statement *Jon is tall*, $f(0.6, 0.3) = 0.5$ means: *It is 0.5 true that it is 0.3 true that it is 0.6 true that Jon is tall.* Moreover each of these degrees is viewed as a scalar approximation of a more complex entity, degrees at lower levels being fixed. Namely, there is a density function $\delta_{f(a_1,\ldots,a_n)}(x)$ on [0, 1] representing the blurry region of more or less appropriate values of the degree of truth at level $n+1$, whose approximation is then understood as the mean-value of the probability measure having this density. In the paper [46], the density is supposed to be normal, with a variance small enough to fit the narrow gauge of the unit interval. Under these restrictions, a degree function is an infinite sequence of elements of the unit interval, the constant sequence of 1's standing for *utterly true*, the constant sequence of 0's standing for *utterly false*. Then, standard fuzzy set connectives can be extended to degree functions, by applying them componentwise to sequences of numbers in the unit interval.

This extension of fuzzy sets, and the context in which it is devised calls for several comments:

- First, Smith seems to endorse an objectivist view of degrees of truth, whereby, in his words [46, p. 169], "each vague sentence is assigned a unique degree function as its unique truth value". This is a first point of disagreement with the positioning of fuzzy sets with respect to the vagueness problem adopted here. In the view advocated in the present discussion, there is no such thing as the actual truth value of a vague statement. The use of membership function (and the unit interval) and of a precise truth value is viewed as a pure convention that helps representing knowledge. The reason why a statement like *Jon is tall is 0.6 true* may be debatable is not because, as Smith says "it is a first approximation of the actual truth value of the vague statement". Fair enough, a membership function is an approximate rendering of a meaning. But what is problematic in Smith's construct is to assume that any sensible person will ever make a statement of the form *Jon is tall is 0.6 true*. The utterer may declare at best *Jon is more or less tall*, for instance, and this statement is considered (fully) true by the receiver who takes it for granted. The 0.6 degree plays a role in the emitter-receiver framework outlined above. However it may be seen as follows: what the receiver models is a membership function for *tall*, which is supposed to be a good enough representation. Then, if the height of Jon eventually gets to be known by the receiver, say 1.7 meters high, the degree of membership $\mu_P(1.7) = 0.6$ (say) can be obtained, and acts as an encoding of the idea that John is tall to some extent. However this figure is a mathematical representation and cannot be naturally produced (let alone interpreted) in communication between persons (even if people can easily outline whole membership functions on simple continuous measurement scales).
- Another difficulty is the extreme mathematical complexity of what Smith considers to be an actual truth value, and the fact that at the same time the author has to resort to ad hoc tricks like approximately fitting a Gaussian density to the unit interval. One may admit such approximations if empirical tests are made to generate sequences of numbers in the unit interval prescribed by the theory. Given the lack of intuition of what a degree of truth can be for lay-people (and even very educated ones), as opposed to other concepts such as utility, probability, cost or similarity, it is unlikely that this theory can be empirically tested. Now if the notion of degree function is to be taken as the basis of a mathematical theory of vagueness, then it could be useful to see if the theoretical results obtained are still valid beyond Gaussians squeezed on the unit interval, i.e. whether this ad hoc restriction is needed at all.

- Finally it is surprizing to see probability density functions and averages playing a key-role in this construction, while never being given any uncertainty-driven interpretation. The author says (p. 172) that the degree of truth of a statement should be represented as a blurry region stretching between 0 and 1. The idea is accommodated by regarding the curve as the graph of a density function. What this density function stands for is not very clear: does it account for variability of the truth value people would use (if they were forced to)? is it a representation of subjective belief about the truth value? In contrast, if a precise value cannot be assigned due to a lack of information, why not use a possibility distribution? Why should the blurry region have a symmetric shape at all (especially for small and for large membership values?). The use of averages also looks partially debatable: if the precise approximation of the truth value corresponds to the one that stands out in the blurry region, a modal value looks more plausible than an average value. In a nutshell, the extreme sophistication of the representation seems to go along with a number of degrees of freedom in the choice of definitions and parameters, both at the interpretive and the mathematical level, which deserves more scrutiny.

Overall, beyond their impressive mathematical and conceptual construction, blurry sets seem to be an idealistic view of gradual vagueness, whose adequacy to concrete data looks very difficult to validate. We suggest here that the use of simple membership functions cannot by itself account for the vagueness phenomenon (and we claim it was not the original intention of the founder of fuzzy sets either, let alone the one of the many users of fuzzy set theory since then). From this point of view, we are not better off by making the notion of truth value very complex, and considering it as a real entity. One reason is that it does not directly fit with the point that vagueness is to a large extent due to uncertainty of meaning. It is the human incapacity to represent gradual predicates by precise membership functions that should be modelled (be it due to their contextual nature, lack of knowledge or lack of definiteness). Instead, the theory of blurry sets seems to bypass this kind of uncertainty by means of a new kind of higher order truth value. The fact that blurry sets are compositional just like fuzzy sets (and interval-valued fuzzy sets) should act as a warning on the fact that this construction is not tailored for uncertainty due to vagueness. But then, what do these blurry regions of the unit interval represent? The fuzzy plurivaluationistic approach later developed by Smith [48] looks simpler and more convincing, even if the present paper suggests, contrary to Smith, that the vagueness phenomenon is related to the fact that several membership functions are possible, and not to the gradual nature of propositions considered vague.

5 Conclusion

It should be clear from the above discussion that fuzzy sets, as explained by Zadeh, have no ambition to grasp the philosophical issue of vagueness, and that gradualness does not always imply the presence of uncertainty: some membership functions only encode the idea of degrees as a substitute to Boolean representations. In fact, some gradual functions do not even encode fuzzy sets [18]. Fuzzy sets refer to sets with gradual boundaries, while vagueness results in a lack of capability to assign truth values to linguistic statements (whether modelled in a bivalent setting or not). If this separation between gradualness and vagueness is taken for granted, vagueness appears as a form of uncertainty in meaning, hence does not lend itself to compositionality. It seems that many controversies between fuzzy set theory scholars and philosophers of vagueness come from the presupposition that fuzzy set theory is a full-fledged approach to vagueness, which turns out not to be the case. Fuzzy sets can be useful in many areas not concerned with the study of natural language, and the vagueness phenomenon is at best facilitated by the presence of gradual predicates, since gradual representations look more cognitively demanding than Boolean ones. We also claim that the study of vagueness may benefit from the choice of the point of view in a dialogue: whether the vague statement is asserted or received seems to matter. Moreover, just as Smith did for plurivaluationism, we argue that neither the epistemic view to vagueness nor supervaluations are incompatible with the idea of gradual predicates.

BIBLIOGRAPHY

[1] M. Black. Vagueness. *Philosophi of Science*, 4:427–455, 1937.

[2] D. Bouyssou and M. Pirlot: Following the traces: An introduction to conjoint measurement without transitivity and additivity. *European Journal of Operational Research*, 163(2):287–337, 2005

[3] G. Coletti and R. Scozzafava. Conditional probability, fuzzy sets and possibility: a unifying view. *Fuzzy Sets and Systems*, 144:227–249, 2004.

[4] R.T. Cook Vagueness and mathematical precision. *Mind*, 111:225–247, 2002.

[5] I. Couso and D. Dubois. Rough sets, coverings and incomplete information. *Fundamenta Informaticae*, 108:223–247, 2011.

[6] B. De Finetti. La logique de la probabilité. In *Actes Congrès Int. de Philos. Scient., Paris*, pages 1–9, Hermann et Cie Editions, Paris, 1936.

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

[8] Y. Djouadi, D. Dubois, and H. Prade. Graduality, uncertainty and typicality in formal concept analysis. In C. Cornelis, G. Deschrijver, M. Nachtegael, and S. Schockaert, editors, *35 years of Fuzzy Sets Theory*, pages 127–147. Springer, Berlin, 2010.

[9] D. Dubois. On ignorance and contradiction considered as truth-values. *Logic Journal of the IGPL*, 16(2):195–216, 2008.

[10] D. Dubois. Degrees of truth, ill-known sets and contradiction. In B. Bouchon-Meunier, L. Magdalena, M. Ojeda-Aciego, and J.-L. Verdegay, editors, *Foundations of Reasoning under Uncertainty,*, volume 249 of *Studies in Fuzziness and Soft Computing*, pages 65–83. Springer-Verlag, 2010.

[11] D. Dubois, F. Esteva, L. Godo, and H. Prade. An information-based discussion of vagueness: six scenarios leading to vagueness (Chapter 40) . In H. Cohen and C. Lefebvre, editors, *Handbook of categorization in cognitive science* , pages 891–909. Elsevier, Amsterdam, 2005.

[12] D. Dubois, S. Gottwald, P. Hájek, J. Kacprzyk, and H. Prade. Terminological difficulties in fuzzy set theory - the case of "intuitionistic fuzzy sets". *Fuzzy Sets and Systems*, 156(3):485–491, 2005.

[13] D. Dubois, W. Ostasiewicz, and H. Prade. Fuzzy sets: History and basic notions. In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy Sets* , The Handbooks of Fuzzy Sets Series, pages 21–124. Kluwer Academic Publishers, Boston, 2000.

[14] D. Dubois and H. Prade. Twofold fuzzy sets and rough sets: Some issues in knowledge representation. *Fuzzy Sets and Systems*, 23:3–18, 1987.

[15] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.

[16] D. Dubois and H. Prade. Belief change and possibility theory. In P Gärdenfors, editor, *Belief Revision*, pages 142–182. Cambridge University Press, Cambridge, UK, 1992.

[17] D. Dubois and H. Prade. The three semantics of fuzzy sets. *Fuzzy Sets and Systems*, 90:141–150, 1997.

[18] D. Dubois and H. Prade. Gradual elements in a fuzzy set. *Soft Computing*, 12:165–175, 2008.
[19] K. Fine. Vagueness, truth and logic. *Synthese*, 30:265–300, 1975.
[20] M. Freund. On the notion of concept I. *Artificial Intelligence*, 172(4–5):570–590, 2008.
[21] B. Ganter and R. Wille. *Formal Concept Analysis, Mathematical Foundations*. Springer-Verlag, 1999.
[22] L. Godo and R. O. Rodriguez. Logical approaches to fuzzy similarity-based reasoning: an overview. In *Preferences and Similarities*, volume 504 of *CISM Courses and Lectures*, pages 75–128. Springer, 2008.
[23] J. A. Goguen. *L*-fuzzy sets. *Journal of Mathematical Analysis and Applications*, 18:145–174, 1967.
[24] J. A. Goguen. The logic of inexact concepts. *Synthèse*, 19:325–373, 1969.
[25] I. R. Goodman. Some new results concerning random sets and fuzzy sets. *Information Sciences*, 34:93–113, 1984.
[26] P. Hájek. *Metamathematics of Fuzzy Logic*. Kluwer, Dordrecht, 1998.
[27] J. Halpern. Intransitivity and vagueness. In D. Dubois, C.A. Welty, and M.-A. Williams, editors, *Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning*, pages 121–129. AAAI Press, 2004.
[28] H.M. Hersh, A. Caramazza, and H.H. Brownell. Effects of context on fuzzy membership functions. In M.M.Gupta, R.R. Ragade, and Yager R.R., editors, *Advances in Fuzzy Set Theory and Applications*, pages 389–408. North-Holland, Amsterdam, 1979.
[29] J. Hintikka. *Knowledge and Belief.* Cornell University Press, Ithaca, 1962.
[30] E. Hisdal. Are grades of membership probabilities? *Fuzzy Sets and Systems*, 25:325–348, 1988.
[31] A. Kaplan and H.F. Schott. A calculus for empirical classes. *Methods*, III:165–188, 1951.
[32] C. Kennedy. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30:1–45, 2007.
[33] F. Klawonn. Fuzzy points, fuzzy relations and fuzzy functions. In V. Novák and I. Perfilieva, editors, *Discovering the World with Fuzzy Logic*, pages 431–453. Springer, Berlin, 2000.
[34] J. Lawry. *Modelling and Reasoning with Vague Concepts*, volume 12 of *Studies in Computational Intelligence*. Springer, 2006.
[35] J. Lawry and I. González Rodríguez. Generalised label semantics as a model of epistemic vagueness. In Claudio Sossai and Gaetano Chemello, editors, *ECSQARU*, volume 5590 of *Lecture Notes in Computer Science*, pages 626–637. Springer, 2009.
[36] S. Lehmke. Degrees of truth and degrees of validity. In *Discovering the World with Fuzzy Logic*, pages 192–237. Physica Verlag, Heidelberg, 2001.
[37] I. Niiniluoto. *Truthlikeness*. D. Reidel, Dordrecht, 1987.
[38] V. Novák. Are fuzzy sets a reasonable tool for modeling vague phenomena? *Fuzzy Sets and Systems*, 156(3):341–348, 2005.
[39] D.N. Osherson and E.E. Smith. On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9:59–72, 1981.
[40] R. Parikh. The problem of vague predicates. In R.S. Cohen and M.W. Wartopsky, editors, *Language, Logic and Method*, pages 241–261. D. Reidel, Dordrecht, The Netherlands, 1983.
[41] Z. Pawlak. *Rough Sets - Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publ., Dordrecht, The Netherlands, 1991.
[42] P. Smith and R. Keefe. *Vagueness: A Reader*. MIT Press, Cambridge, MA, 1997.
[43] E.H. Ruspini. On the semantics of fuzzy logic. *International Journal of Approximate Reasoning*, 5:45–88, 1991.
[44] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.
[45] Ph. Smets. The transferable belief model for quantified belief representation. In D.M. Gabbay and Ph. Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 1, pages 267–301. Kluwer, Doordrecht, The Netherlands, 1998.
[46] N.J.J. Smith. Vagueness and blurry sets. *Journal of Philosophical Logic*, 33:165–235, 2004.
[47] N.J.J. Smith. Vagueness and closeness. *Australasian Journal of Philosophy*, 83:157–183, 2005.
[48] N.J.J. Smith. Vagueness and Degrees of Truth. Oxford University Press, Oxford, UK, 2008.
[49] B.C. van Fraassen. Singular terms, truth-value gaps, and free logic. *Journal of Philosophy*, 63:481–495, 1966.
[50] B. van Gasse, C. Cornelis, G. Deschrijver, and E.E. Kerre. Triangle algebras: A formal logic approach to interval-valued residuated lattices. *Fuzzy Sets and Systems*, 159(9):1042–1060, 2008.
[51] T. Weston. Approximate truth. *Journal of Philosophical Logic*, 16:203–227, 1987.

[52] H. Weyl. The ghost of modality. In *Philosophical essays in memory of Edmund Husserl*, pages 278–303. Cambridge, MA, 1940.
[53] T. Williamson. *Vagueness*. Routledge, London, 1994.
[54] L.A. Zadeh. PRUF - a meaning representation language for natural languages. *International Journal of Man-Machine Studies*, 10:395–460, 1978.
[55] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
[56] L.A. Zadeh. Outline of a new approach to the analysis of complex system and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, 3:28–44, 1973.
[57] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning, part I. *Information Sciences*, 8:199–249, 1975.
[58] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:1–28, 1978.
[59] L.A. Zadeh. A note on prototype theory and fuzzy sets. *Cognition*, 12:291–297, 1982.

Didier Dubois
IRIT – Institut de Recherche en Informatique de Toulouse
CNRS and University of Toulouse
118 route de Narbonne
31062 Toulouse Cedex 09, France
Email: dubois@irit.fr