Handling Borderline Cases Using Degrees: An Information Processing Perspective

HENRI PRADE AND STEVEN SCHOCKAERT[1]

Overview

Although linguistic vagueness is generally related to the existence of borderline cases, the main theories of vagueness are not primarily aimed at handling borderline cases explicitly. As a result, these theories are able to avoid the use of truth degrees, e.g. by focusing on truth value gaps or by taking the epistemic point of view that there is an unknown but sharp boundary between truth and falsity. We argue that, on the other hand, in information processing settings, borderline cases usually are a matter of degree, or at least of ranking, although such degrees may serve different purposes. To support our claim, this paper discusses a number of information processing scenarios, in relation with their use of degrees to handle borderline cases. First, we consider the use of fuzzy labels, contrasting the role of (graded) borderline cases in three different situations: (i) allowing for flexibility when specifying and evaluating fuzzy requests, (ii) allowing for abstraction when describing precise information, and (iii) allowing for uncertainty when stating imprecise information. Second, we discuss the idea of degrees of typicality in the setting of formal concept analysis, seeing e.g. penguins or kiwis as borderline cases of birds. Finally, we illustrate how degrees of similarity may be useful for maintaining consistency, taking advantage of a flexible understanding of linguistic terms, seeing e.g. civil unions as borderline cases of marriages.

1 Introduction

Vagueness already has a long history in modern philosophy, and there are different, somewhat rival, views of vagueness [24, 23, 41, 43]. The supervaluation view [44, 17], for instance, prefers to admit a truth value *gap*: borderline statements have no truth value, but a compound statement involving vague terms may be true if it is true for every possible way in which these vague terms can be precisified (where a vague term is thus viewed as the collection of all its possible sharpened versions). In contrast, the epistemic view [48] rather presupposes the existence of a unique, precise border between truth and falsity, but considers "vagueness as a kind of ignorance", inasmuch that "there really is a grain of sand whose removal turns a heap into a non-heap, but we

[1]This paper develops the content of an extended abstract entitled "About vagueness, typicality and similarity" by the first author, presented at the Logical Models of Reasoning with Vague Information Conference (LoMoReVI), on Sept. 14–17, 2009 at Čejkovice, Czech Republic. Steven Schockaert is a postdoctoral fellow of the Research Foundation – Flanders (FWO).

cannot know which one it is". However, despite the lack of a unified view on vagueness, it is generally agreed that "a term is vague to the extent that it has borderline cases" [42], hence vagueness might be equated with the idea that a vague concept partitions the universe of discourse (sometimes implicitly) into more than two parts; see e.g. [9] for a detailed discussion on the kind of informational scenarios that give rise to such a situation. Nonetheless, the study of vagueness has mainly concentrated on the difficulty of reasoning with vague statements in general, rather than on the explicit handling of borderline cases in practical applications.

In contrast, while the representation of vagueness has not been so much of an issue in Artificial Intelligence (AI) and Pattern Recognition, these fields have progressively built methods that allow to explicitly handle borderline cases in different information processing settings, using degrees in one way or another to distinguish between different levels of being borderline. The field of AI, for instance, has focused on the representation of incomplete and uncertain information, or on the introduction of flexibility when dealing with soft constraints. In that respect, the development of fuzzy logic [49]—assigning intermediary truth values to borderline statements—has primarily been motivated by the need to handle gradual properties in approximate reasoning, rather than offering an alternative theory of vagueness, even if a treatment of sorites using fuzzy logic has been proposed very early [20] (see also [40]). Another example is the development of nonmonotonic reasoning approaches for handling exceptions under incomplete information [25], which often rely on plausibility orderings between interpretations to determine what is true in the most normal worlds. For instance, given that typical birds can fly while penguins cannot, penguins are seen as exceptional birds, and are in this sense borderline. Along the same lines, AI uses orderings to represent user preferences, or to encode priorities when revising knowledge bases, while Pattern Recognition has found it advantageous to allow elements to have a graded membership in a given cluster [4].

Thus there seems to be a gap between the philosophical views on vagueness (e.g. supervaluation semantics), and practical methods for handling borderline cases (e.g. based on fuzzy set theory). In this respect, it is interesting to note that a refinement of supervaluation semantics, called standpoint semantics [3], has recently been proposed which seems to bridge this gap to a large extent. Essentially, standpoint semantics adds structure to supervaluation semantics by restricting the possible precisifications of a vague term to those that correspond to a *standpoint*, i.e. a given choice of threshold values for a fixed set of parameters. A fuzzy set can then be seen as an approximate representation of the possible standpoints that can be taken regarding some vague term, sacrificing some structure for better tractability. Note that this view on fuzzy sets is similar to the view of a fuzzy set as the approximation of a random set. The transition from supervaluation semantics to standpoint semantics also seems to support the view that in practical applications, handling borderline cases often requires the use of degrees in one form or another, including intermediary truth values, rankings, or threshold values (i.e. standpoints).

In this paper, we further elaborate on the role of degrees when dealing with borderline cases. By zooming in on a number of basic situations in which borderline cases are encountered, we illustrate how degrees may serve different purposes, and why they are important in applications. In Section 2, we discuss the different uses of fuzzy labels, where degrees are used to model gradualness, allowing for a continuous transition from objects that perfectly satisfy some criterion to objects that do not satisfy it at all. We especially contrast three different uses of fuzzy labels. First, fuzzy labels may be used to express requests, in which case the degrees associated with borderline cases express levels of flexibility. Second, fuzzy labels may be used to describe precise information in a more abstract way (e.g. categorization), in which case degrees express levels of compatibility. Third, fuzzy labels may be used for the imprecise description of ill-known cases, in which case degrees express levels of uncertainty. Subsequently, Section 3 considers statements such as "Tweety is a bird". While "bird" does not correspond to a fuzzy label, birds which lack some typical properties such as "being able to fly" or "having feathers" (e.g. kiwis) may be considered as borderline. Note that tacitly assuming that "Tweety is a typical bird" when all we know is that "Tweety is a bird" is often done in commonsense reasoning. Borderline cases of birds may be graded by asserting that birds are typical (i.e. not borderline) to the extent that they satisfy all the important properties. We discuss the practical use of such degrees of typicality in the context of formal concept analysis. In particular, we focus on the question of how to assess the degree to which some known entity is a typical bird, as well as on the question of how to model the uncertainty that arises when all we know is that "Tweety is a typical bird". Finally, in Section 4, we consider assertions such as "John is married", which may require a more permissive understanding than what the term "married" actually means. For instance, when a source tells us that John is married, it is still somewhat plausible that John is, in fact, in a civil union. Concepts which are similar to marriages, such as a civil union, may thus be seen as borderline cases, even if civil unions do not correspond to special cases of marriages. This view is close to the idea of verisimilitude, i.e. the idea that some false statements are closer to the truth than others. Indeed, the assertion "John is married" is intuitively less false when John is in a civil union than when he is single. We discuss how such a similarity-based treatment of statements can be a useful tool for maintaining consistency in a knowledge representation setting.

In some sense, the information processing setting illustrates the difference between a vague term and the vague understanding of a term. Indeed, the use of fuzzy labels allows for a non-vague understanding of vague terms, while Sections 3 and 4 are concerned with the vague understanding of non-vague terms.

In all of the considered scenarios, we find that degrees play a key role in the practical handling of borderline cases. At first glance, this observation seems at odds with the limited attention that has been given to degrees in most philosophical treatments of vagueness. However, one may argue that, despite the focus on borderline cases, strictly speaking, none of the scenarios we discuss is really concerned with vagueness as it is usually understood, but rather with related notions such as fuzziness, typicality and similarity. Indeed, borderline cases between compatible and incompatible; typical and atypical; similar and dissimilar; etc., are different in nature from borderline cases between truth and falsity. It is tempting to speculate that this subtle difference may be responsible for much of the controversy on the use of degrees in theories of vagueness.

2 On the different uses of fuzzy labels

A fuzzy set A is a mapping from some universe of discourse U to the unit interval $[0,1]$. For any u in U, $A(u)$ is called the membership degree of u in A, and reflects the extent to which object u satisfies some underlying property. This property is usually described by means of a possibly vague, linguistic label. We use the term fuzzy label to denote such a linguistic description whose interpretation is explicitly specified as a fuzzy set. It is important to note that by providing a fuzzy set as the interpretation of a linguistic description, any vagueness that may have resided in the linguistic description is resolved. What remains is the idea of gradualness, captured by the associated fuzzy set representation: some objects are more compatible than others with a fuzzy label. Indeed, Zadeh himself already considered that vagueness should not be confused with fuzziness. He wrote (in a footnote p. 396 in [49]): "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." So he associates the idea of vagueness with situations where the information at hand is insufficient for some particular use. This may be viewed as putting vagueness closer to generality, or even to ambiguity, two other notions that are classically distinguished from vagueness (see, e.g., [42]). More interestingly, it is noticeable that in Zadeh's view, vagueness is not a feature which is intrinsically attached to a statement (as fuzziness is), but is rather associated to a proposition in the context of a particular purpose. In fact, acknowledging the use of fuzzy labels when modeling vagueness would for instance amount to adapt the supervaluation view such that the possible delineations of a vague term may be fuzzy sets rather than classical sets. This is beyond the scope of this paper.

We may think of very different circumstances in which fuzzy labels or fuzzy sets may be used. There are uses of fuzzy sets as a device for compact encoding that are motivated by technical needs, e.g. using Łukasiewicz multi-valued logic to encode linear optimization problems [21] or to deal with aspects of coding theory [5], using fuzzy description logics for handling data on continuous domains in multimedia information retrieval [31], and using fuzzy answer set programming to find strong Nash equilibria [37]. In this paper, however, we focus on the cognitive uses. Fuzzy sets are useful when expressing desires or requests, when categorizing situations, or when interpreting received information. These three cognitive tasks, which are discussed in more detail below, favor different representation capabilities of fuzzy sets: respectively the encoding of a preference ordering, the embedding of a notion of similarity, and the expression of an uncertainty distribution [11].

Fuzzy labels expressing preference

Consider the situation in which some user is looking for an apartment to let, and attempts to express her desires as a query to some database. In such a case, it may be difficult to come up with a good query which only uses hard constraints, because there is a risk to over-constrain (in which case there would be no matches) or under-constrain (in which case there would be many suboptimal matches) the request. To cope with this, the user may come up with an under-constrained query, together with additional information about her preferences, which could be used to rank the objects matching the actual query. This can easily be accomplished using fuzzy labels. For instance, the user may ask for an apartment that is "rather cheap" and "close" to downtown. This request may be translated into a query by assuming a very liberal understanding of the terms "rather cheap" and "close", while a fuzzy set representation of these terms may be used to subsequently rank the matching apartments.

When expressing desires, fuzzy sets are thus used to encode which are the more or less acceptable values, the fully acceptable values, and the fully excluded values of the relevant attributes. In case the underlying attribute values are ordered on a continuum, processing the query basically amounts to building a membership function representation for each linguistic term, and choosing appropriate operations on fuzzy sets to combine the elementary requirements (e.g. generalizations of set complementation, union and intersection). If the underlying attribute domain is discrete and unordered, a direct assessment of acceptability levels would be required for every attribute value. In such a case, it may be appropriate to use a *scale* for the membership degrees which is different from the real unit interval, the latter being appropriate especially for numerical attributes.

Note that, in principle, fuzzy sets representing preferences are not necessarily associated with a label. Labels are useful here only if the user needs to communicate about what she is looking for, or if the queries are expressed in a prescribed fuzzy vocabulary that has been defined in advance. Since all linguistic terms in such a fuzzy vocabulary have been given a precise meaning, through the specification of fuzzy sets, there is no vagueness in the query itself. However, the vagueness of the associated linguistic terms helps us to get an intuitive grasp of the kind of flexibility that is offered by a fuzzy request. Hence, instead of being a culprit, the vagueness of language is actually advantageous here.

Fuzzy labels expressing similarity

For humans, it is often straightforward to provide natural language descriptions of numerical information. By abstracting away from the actual numbers, the main conclusions that can be drawn from available data may then be easier to see. There is a discrepancy, however, between the continuum of values for attributes such as "age", and the finite (and generally small) number of natural language terms that are at our disposal to describe them. When interfacing numerical domains with natural language terms, it may be beneficial to represent these terms as fuzzy sets. The gradualness of predicates such as "young", viewed as a fuzzy label, then enables us to avoid the use of precise threshold values, which would to a large extent be arbitrary. In other words, the purpose of using fuzzy labels is to allow for flexible threshold values. The decision of which is the most appropriate label to describe a given situation or object can then be postponed, which may allow us to take additional information into account. Note that the final label that is used to describe some situation may also be a compound label (e.g. "old but not very old"), which is obtained using linguistic connectives (e.g. "but") and hedges (e.g. "very"). Note that this use of fuzzy labels is rather different from its use in the specification of queries. Here, the underlying fuzzy set representations essentially convey a notion of similarity: the ages that are somewhat compatible with the predicate "young" are those that are *similar* to ages which are perfectly compatible with it. The need to qualify a particular situation or object, using a linguistic label, may arise for instance in categorization, as is done by fuzzy symbolic sensors [29]. Another application is when a natural language generation system has to build a minimal description designating (in a non-ambiguous manner) an object in a scene, using a limited vocabulary referring for instance to size, location, or color (see [15] for a study of this problem).

The qualification of a particular case, as described above, should be contrasted with the description of *generic* situations. Given a database (whose attribute values are precisely known), one may for instance be interested in providing some partial summaries of its content through aggregates and association rules. Using fuzzy labels, we may state assertions such as "the average salary of *young* people in the company is *about* 1500 euros" or "more than 90% of *young* people have a salary less than ...". Without fuzzy labels, precise thresholds need to be used, as in "the average salary of people below 30 ...", or "93% of people below 35 have a salary less than ...". The main advantage of using fuzzy labels, here, is that they may cover a larger portion of the data. The flexible thresholds that are introduced by using fuzzy labels thus work at our advantage: as several labels may be more or less acceptable to describe a given situation, there is more freedom to come up with informative and useful rules, while staying sufficiently close to what the data actually conveys. However, in order for such statements with fuzzy categories to make sense, it is desirable that for any acceptable precisification of these categories, the value of the average (or any other considered measure) does not vary too much, and remains in a prescribed (fuzzy) range. See [10] and [14] for database aggregates and association rules respectively, on this issue.

Fuzzy labels expressing uncertainty

In the previous two settings, fuzzy labels were used w.r.t. perfectly known information, resp. to convey desires and to provide linguistic descriptions. Here we consider the setting where one receives a piece of information which is expressed linguistically, such as "Bob is young". In this case, we do not have access to any precise information, and we thus remain uncertain about the actual age of Bob. Such a statement of the form X *is* A, where X refers to a single-valued attribute (here the age) is *fuzzy* if A is represented as a fuzzy set. Similarly, the statement X *is* A is *imprecise* as soon as the extension of the set (or fuzzy set) representing A is not a singleton of the universe of discourse U, as in the statement "John is between 20 and 30 years old" (assuming for instance that U is some subset of the set of positive integers). In both cases, we are *uncertain* about the precise age of the person, although we do know that some ages are impossible, because they are not compatible at all with the given (fuzzy) label. Moreover, given a fuzzy label such as "young", it is clear that the more a given age can be considered as young, the more it is compatible with the asserted information, and hence the more it should be considered as a possible value. This leads us to take the *degree of possibility* that $age(Bob)$ is 25 to be the degree of membership of 25 in the fuzzy set representing "young" [49]. More generally, it is postulated that

$$\pi_X(u) = \mu_A(u),\ \forall u \in U$$

where $\pi_X(u)$ represents the degree to which u is considered as a possible value of X and $\mu_A(u)$ represents the membership degree of u in the fuzzy set corresponding with the label A. This means that a value u is all the more possible ($\pi_X(u)$ is all the higher) as u is considered to be more consistent with the fuzzy meaning conveyed by A. *But*, this does not mean that the predicate "young" is vague, or involve any uncertainty in *itself*, at least as long as there is no disagreement about the membership function representing this predicate in the considered context. What may be uncertain is not the piece of information in itself (in fact, it is fully certain that “Bob is young”, inasmuch as the source that provides the information is reliable), but e.g. the truth of a statement such as “Bob is less than 25”, when all we know is that “Bob is young”.

The three uses of fuzzy labels discussed in this section illustrate the role which is played by graded borderline cases in settings where vague linguistic terms are precisiated using fuzzy sets. As we illustrate in the next two sections, graded borderline cases may also play a central role when there is no vagueness, strictly speaking, in the considered descriptions. Indeed, even for a term with a well-defined meaning, in some situations, it is of interest to consider a more restrictive, or on the contrary, a more liberal view of the extension of the term, as we are going to see in Sections 3 and 4 respectively.

3 Introducing typicality in formal concept analysis

The idea of a concept, for which there exist different views, is difficult to define, although it is clear that concepts play a crucial role in categorization tasks [28, 30]. Classically, categorical membership differs from typicality, and both notions may even refer to different sets of features. Moreover typicality is not directly linked with vagueness, while categorical membership is. Still, it is interesting to examine how these aspects interact in the setting of formal concept analysis, even if formal concepts are an idealized and simplified view of the idea of a concept. In this section, we provide an illustration of the use of degrees for handling borderline cases with respect to the notion of a formal concept.

Formal concept analysis (FCA) [47] is a mathematical theory that defines a formal concept as a pair of two sets: i) a set of objects, called the extension of the formal concept, and ii) a set of properties, called its intension. Thus, for instance, a set of animal species (playing the role of the objects) is considered together with a set of properties (e.g. ‘laying eggs’, ‘flying’, ‘having fur’, etc.). In particular, FCA starts from the notion of a (formal) context, which is represented by a binary relation $R \subseteq Obj \times Prop$ that encodes which objects satisfy which properties. In other words, the pair (x, y) is an element of R if and only if object x has property y. Then, the concept ‘bird’ can be defined in a context which links animal species to relevant properties for describing animals.

What makes FCA theory attractive is its ability, when objects have known properties, to jointly identify the extension and the intension of formal concepts. Moreover, the set of formal concepts induced by a formal context is organized in a double lattice structure (w.r.t. objects and properties), which is exploited by data mining algorithms [34]. Despite the general attractiveness of FCA, however, typicality has not been considered in this setting, with the exception of a very recent proposal [7] that we discuss now, after recalling some basic notions.

Given a context, the extension and intension of a concept are supposed to mutually determine each other. We denote by $R(x)$ the set of properties possessed by object x, and by $R^{-1}(y)$ the set of objects having properties y. One can define

- the set of objects $R^{-1}(Y)$ having all the properties of some subset $Y \subseteq Prop$ as $R^{-1}(Y) = \{x \in Obj \mid Y \subseteq R(x)\}$, and dually,
- the set of properties $R(X)$ possessed by all the objects in some subset $X \subseteq Obj$ as $R(X) = \{y \in Prop \mid X \subseteq R^{-1}(y)\}$.

Then, a formal concept is an (extension, intension)-pair (X,Y), such that

(1) $X = R^{-1}(Y)$ and $Y = R(X)$

with respect to a context R. Thus, in a formal concept, the objects in the extension have all the properties in the intension, and conversely the properties in the intension are possessed by all the objects in the extension.

Table 1 provides a toy example of a formal (sub)context, where the objects and properties relate to birds:

$$Obj = \{sparrow, parrot, penguin, kiwi\}$$
$$Prop = \{\text{'laying eggs'}, \text{'having two legs'}, \text{'flying'}, \text{'having feathers'}\}$$

The symbol $+$ in a cell of Table 1 indicates that the corresponding object has the corresponding property. Both of the following pairs are examples of formal concepts:

$$(\{sparrow, parrot\}, \{\text{'laying eggs'}, \text{'having two legs'}, \text{'flying'}, \text{'having feathers'}\})$$
$$(\{sparrow, parrot, penguin, kiwi\}, \{\text{'laying eggs'}, \text{'having two legs'}\})$$

Note that since $(Obj, Prop)$ does not constitute a formal concept in itself, either some properties or some objects need to be left out. The first case corresponds to the (unique) formal concept whose intention contains all properties. Likewise, the second case corresponds to the (unique) formal concept whose extension is the set of all objects. Generally speaking, it can be shown that a formal concept (X,Y) is such that $X \times Y \subseteq R$ holds. Note that this Cartesian product visually gives birth to a rectangle made of $+$ (maybe after a proper reordering of the lines and/or the columns) in the table.

Degrees can be introduced in two different ways in FCA. First, we may consider the relation describing the formal context to be fuzzy [1], with the aim of acknowledging the gradualness of properties. Then, the extent to which (x,y) is an element of R is a matter of degree that reflects to what extent object x has property y, and $R^{-1}(y)$ is the fuzzy set of objects having property y to some degree. By extending the definition of a formal concept in a natural way, formal fuzzy concepts can then be introduced. However, such fuzzy concepts have nothing to do with typicality.

The second type of degrees that may be considered in FCA are obtained by keeping R crisp (i.e. binary-valued), and by relating degrees of typicality of objects to degrees of importance of properties, as we are going to see. Let us consider a particular subcontext, corresponding to a set of objects $X^{sc} \subseteq Obj$ and a set of associated properties $Y^{sc} \subseteq Prop$. Then, the basic idea is to equip Y^{sc} with degrees of importance and X^{sc} with degrees of typicality, and to put the *important* properties of the subcontext in relation with its *typical* objects, via a mutual characterization similar to the one provided by Equation (1). This can be summarized by the two following principles:

(A) An object x is all the more normal (or typical) w.r.t. a set of properties Y^{sc} as it has all the properties $y \in Y^{sc}$ that are sufficiently important;

(B) A property y is all the more important w.r.t. a set of objects X^{sc} as all the objects $x \in X^{sc}$ that are sufficiently normal have it.

Let us illustrate this idea on the example of Table 1, and let us assume the following typicality levels: $X^{sc}_{typ}(sparrow) = X^{sc}_{typ}(parrot) = 1$, $X^{sc}_{typ}(penguin) = \alpha < 1$ (since penguins do not fly), and $X^{sc}_{typ}(kiwi) = \beta$, with $1 > \alpha > \beta$ (since kiwis do not fly and have no feathers). Note that the levels used here are purely symbolic, and are just supposed to belong to an ordinal scale.

Table 1.

R	*eggs*	*2 legs*	*feather*	*fly*
sparrow	+	+	+	+
parrot	+	+	+	+
penguin	+	+	+	
kiwi	+	+		

To compute the fuzzy set of important properties according to principle (B) above, we need to evaluate the degree of inclusion of a fuzzy set of typical objects into the set of objects having property y. This can be accomplished using a multiple-valued implication connective $\rightarrow$ as follows:

(2) $\quad Y^{sc}_{imp}(y) = min_x(X^{sc}_{typ}(x) \rightarrow R(x,y))$

where $Y^{sc}_{imp}(y)$ is the degree of importance of property y. Moreover, as R is a crisp relation, we either have $R(x,y) = 1$ (when x has property y) or $R(x,y) = 0$ (otherwise). The implication connective satisfies $a \rightarrow 1 = 1$ as any multiple-valued implication, and should be chosen such that $a \rightarrow 0 = 1 - a$. This choice expresses the idea that the more a bird is considered typical, the less the properties that it does not have are assumed to be important in the definition of the concept *bird*. Note that in the case where (X^{sc}, Y^{sc}) would be a classical formal concept, $Y^{sc}(y) = min_x X^{sc}(x) \rightarrow R(x,y)$ is nothing but $Y^{sc} = R(X^{sc})$.

Let us now compute the fuzzy set (with membership function μ) of typical objects according to principle (A) above by

(3) $\quad \mu(x) = min_y(Y^{sc}_{imp}(y) \rightarrow R(x,y))$

We get $\mu(sparrow) = \mu(parrot) = 1$; $\mu(penguin) = \alpha$; $\mu(kiwi) = \beta$ (since $(1-a) \rightarrow 0 = a$). As can be observed, we have $\forall x$, $\mu(x)$=$X^{sc}_{typ}(x)$. In fact, the equations (2)–(3) are the counterparts of the definition of a formal concept, given in (1), taking into account the graded interrelation between important properties and typical objects [7]. Thus, in the example of Table 1, the pair of fuzzy sets $(X^{sc}_{typ}, Y^{sc}_{imp})$ constitutes a generalized formal concept, since the fuzzy sets of important properties and of typical objects are in agreement.

More generally, typicality has been addressed in the setting of prototype theory. The idea, noticed by Wittgenstein, that things covered by a term often share a family resemblance may be seen as being at the basis of prototype theory [36], where categorization is understood in terms of similarity (see [28]). Then, a sparrow appears to be a more typical bird than a kiwi, inasmuch as a sparrow has more of the constituent properties of a bird prototype than a kiwi. This is also acknowledged in the approach presented above.

In this context, it is useful to refer to the critical discussion of a fuzzy set-based approach to prototype theory made by Osherson and Smith [33] when dealing with conceptual combination (see Zadeh [50] for a reply). This discussion is centered around the problem of compound concepts, noticing e.g. that the set of all typical red birds is not the intersection of the set of typical red animals with the set of typical birds. In our setting, this phenomenon can be accounted for noting that typical red birds would be defined on the basis of an explicit representation of important properties for red birds, which cannot be obtained from the important properties for red animals and those for birds. As a result, a fuzzy set of typical red birds, in our setting, will not necessarily correspond to the fuzzy set combination of a fuzzy set of typical red animals and a fuzzy set of typical birds.

This approach enables us to relate a typicality preordering among objects to an importance preordering among properties. It has some common flavor with the approach proposed, in a different perspective, by Freund, Desclés, Pascu and Cardot [16], which starts from the commonly accepted idea that a (proto)typical object in a category is an object that satisfies any *normally expected* property from the category (e.g., "birds fly"), and uses a *non-monotonic* reasoning view for performing contextual inferences about typical objects. However, our focus here is not inference, and we have rather tried to embed typicality in the setting of formal concept analysis. Besides, Freund [18, 19] proposes a qualitative model where the typicality associated with a concept is described using an ordering that takes into account the characteristic features of the concept. The approach uses a "salience" (partial) order between features, just as we use an importance ordering between properties. However, the salience ordering is not associated with a typicality ordering in the way proposed here, even if the relation with FCA is also discussed in [18] (when typicality is not graded).

Interestingly enough, the view proposed here leads to a representation of statements such as "Tweety is a bird" in terms of the certainty with which Tweety has each of the considered properties, just as in possibilistic logic [13] (which also provides a way for processing non-monotonic reasoning [2]).

Indeed, representing a statement such as "Tweety is a bird" amounts to state that i) it is fully possible that Tweety has the (Boolean) properties that a bird may have, and ii) the *possibility* that Tweety does *not* have the property y is all the greater as y is less important for birds,

$$\forall y \in Y^{bird}, \pi_{y(Tweety)}(yes) = 1 \text{ and } \pi_{y(Tweety)}(no) = 1 - Y^{bird}_{imp}(y)$$

where Y^{bird}_{imp} is the fuzzy set of *important properties* for birds. Equivalently, the *certainty* that Tweety has property y is all the greater as y is more important for birds, and is equal to $Y^{bird}_{imp}(y)$.

One might think of another representation of "Tweety is a bird" that would perfectly parallel the one used for restricting the possible values of a single-valued attribute (e.g., representing "Bob is young" by $\pi_{age(Bob)}(u) = \mu_{young}(u)$). Namely, one may also represent "Tweety is a bird" as a possibility distribution over a set of mutually exclusive species of birds

$$\forall x \in X^{bird},\ \pi_{type(Tweety)}(x) = X^{bird}_{typ}(x).$$

This would also enable us to conclude, using the subcontext R in Table 1, that indeed the certainty that Tweety has property y is equal to $Y^{bird}_{imp}(y)$. But, the representation in terms of more or less certain properties is generally more compact in practice than the one in terms of objects (e.g., in our example, we have considered only four types of birds for simplicity!).

As typicality has allowed us in this section to restrict the extension of a concept, the next section addresses the converse concern, namely *enlarging* the meaning of a term to *close* terms having *similar* meanings. For instance, the term "married" may be understood strictly, or may sometimes be interpreted in a flexible way by replacing it by a (possibly weighted) disjunction such as "married or in a civil union".

4 Flexible understanding of linguistic terms for maintaining consistency

When a statement involving a vague term is asserted, its precise meaning may differ depending on the context in which it is used and on the person who is asserting it. As a result, upon receiving such a statement, we are often uncertain about its precise meaning. In some scenarios, this uncertainty may be intended by the speaker. Indeed, vagueness may among others be used to help the listener focus on what is relevant in the given context, it may soften complaints or criticisms, or it may be used to deliberately put the listener on the wrong tracks without the risk of being accused of lying [22]. This latter aspect of communication is also stressed in [27], where a bipolar view on assertability is put forward, distinguishing between situations in which a statement is definitively assertable, situations in which it is merely acceptable to assert it, and situations in which it cannot be asserted without condemnation; see [26] for a Bayesian treatment of this issue. A similar view on vagueness in dialogues is also suggested by Wahlster [45], giving the example of a dialogue between a hotel owner and a customer, who wants to know whether the room he is considering to rent is 'large'. Clearly, when the owner subsequently claims that the room is 'quite large', it is not at all certain that the customer would agree when seeing the room. Note that in the latter case, a vague term is used even though the hotel owner may know the exact surface of the room. In addition to the problem of finding the most appropriate linguistic label to describe a precisely known state of affairs, the goal of using vague language may also be to cope with a lack of precise knowledge, as already discussed in Section 2.

In general, successful communication depends on the ability of the participants to establish a common ground which is sufficiently specific. The required alignment between speaker and listener may happen explicitly, e.g. through clarification requests and reformulations, but also implicitly. In the latter case, the listener may for instance revise earlier assumptions on the meaning of assertions when they turn out to be inconsistent with subsequent assertions. It is important to note that these considerations do not only apply to vague terms: natural language terms with a precise and unambiguous meaning are often used in a flexible way, encompassing situations that are not normally associated with the term. For instance, suppose we initially believe that John is married, while later we notice that he answers negatively to the question *Are you married?* on a web form. We may then revise our earlier belief by assuming a more liberal understanding of the term *married*, viz. we take our earlier beliefs to mean that John is either married or in a civil union, and thus, given our new observation, that he is in a civil union. Note that there is a certain duality between this idea of flexibly using a precise term and the idea of typicality which was discussed in the previous section. Indeed, when we reinterpret statements about birds as statements about typical birds, we tighten our understanding of some well-defined concept, while in the marriage example, we enlarge our understanding of a well-defined concept.

In a computational setting, where knowledge is encoded in some logic and is provided to us by a number of different sources, assumptions need to be made about what the considered terms, e.g. associated with atomic propositions, mean. As sources may have used some terms in a slightly unexpected way, when integrating knowledge from different sources, we may end up with a knowledge base that is logically inconsistent. This observation suggests a view on maintaining consistency in a knowledge representation setting based on a flexible reading of propositions. Indeed, when a knowledge base turns out to be inconsistent with an observation, we may restore consistency by getting rid of the *less entrenched* formulas, but also by weakening some formulas, assuming a more tolerant reading of the terms underlying it. For the ease of presentation, in the following we assume a scenario involving two agents: the *listener*, whose initial knowledge includes all relevant integrity constraints (e.g. indicating which properties are mutually exclusive), and the *speaker*, which corresponds to an external source that provides us with new information. We also assume that the listener has background knowledge about how the terms used by the speaker should be understood, i.e. some form of alignment is assumed between the listener and the speaker.

To deal with inconsistencies that are due to flexible language usage, it is useful to notice that the exact meaning of an atomic property p as it is understood by the speaker may not necessarily be expressible in the language of the listener. For instance, what the speaker calls 'cold weather' may have a more narrow meaning than how the listener understands this term, and there may be no other term whose understanding by the listener exactly corresponds to the speaker's understanding of 'cold weather'. Let us write p_{speak} and p_{list} to denote the understanding of the atomic property p by the speaker and listener respectively. We may then consider the weakest formula α^-_{list} which is expressible in the language of the listener and which entails p_{speak}, and the strongest formula α^+_{list} which is expressible in the language of the listener and which is entailed by p_{speak}. Conversely, p_{list} may not be expressible in the language of the speaker and we may need to consider entailing and entailed statements α^-_{speak} and α^+_{speak}.

The two agents are perfectly aligned when for each atomic property p, the four corresponding formulas α^-_{list}, α^+_{list}, α^-_{speak} and α^+_{speak} are known. Note in particular that modeling the (mis)alignment between speaker and listener due to flexible language does not, as such, require the use of degrees. However, in most situations, we will only have incomplete knowledge about how the understanding of a given term by the speaker is related to its understanding by the listener. It may thus be useful to introduce degrees to discriminate between more and less plausible alignments. In practice, e.g. our information about α^-_{list} may be encoded in a graded way, differentiating between formulas α which definitely entail p_{speak} and formulas that plausibly entail it. As a form of plausible reasoning, we may initially make some rather strong assumptions on the alignment between both sources (e.g. assuming α^{-}_{list} to be a rather weak formula), and then revise these assumptions (e.g. assuming α^{-}_{list} to be a stronger formula) when inconsistencies arise. For example, if the speaker tells us that it is warm outside, we may initially take this to mean that the temperature is above $30°C$, and later revise this to "above $25°C$". As another example, suppose that the speaker tells us that John is married, then we may model the resulting knowledge, in the language of the listener, using the following possibilistic logic knowledge base:

$$K = \{(married_{list}, 0.5), (married_{list} \vee civil\text{-}union_{list}, 0.75), \\ (married_{list} \vee civil\text{-}union_{list} \vee cohabitation_{list}, 1)\}$$

where the certainty weights are interpreted in a *purely ordinal* fashion, as lower bounds on the necessity that the corresponding proposition is satisfied [13]. The knowledge base K models the fact that we are certain that $married_{speak}$ entails $married_{list} \vee civil\text{-}union_{list} \vee cohabitation_{list}$, and somewhat certain that it entails $married_{list} \vee civil\text{-}union_{list}$; with even more limited certainty we believe that it entails $married_{list}$. In other words, the idea is to progressively weaken what is claimed by the speaker, and believe the resulting propositions with increasing certainty. This way of modeling the alignment between speaker and listener can be generalized to a setting where information is provided by multiple sources. It can be used as a basis for merging propositional information coming from different sources, by treating conflicts among these sources as evidence for misalignment [38, 39]. The degrees that are used can also be given a probabilistic flavor, or we may avoid the use of numerical degrees altogether and rely on symbolic, partially ordered certainty scores [39]. Moreover, it should be stressed that, as for degrees of typicality, the degrees we consider here are not directly related to categorical membership. Indeed, degrees are used to encode what we know about the alignment; their purpose is to allow us to find the most likely alignment between speaker and listener which does not lead to logical inconsistency.

Essentially, the approach outlined above deals with inconsistencies by assuming that when somebody asserts some proposition, we may (only) deduce (with complete certainty) that something similar to that proposition is the case, where we do not only consider the logical models of that proposition as possible worlds, but also those worlds that can be related to models by assuming a flexible understanding of the underlying terms. Such a similarity-based view on logical consequence may also be studied in an abstract way by considering similarity-based consequence operators of the form $\models^{\lambda}$, where $p \models^{\lambda} q$ means that all models of p are similar to some model of q and λ is a tolerance parameter on the required strength of similarity [8]. This idea stands in contrast to non-monotonic consequence relations where p entails q when the most typical models of p are also models of q [25]. Finally, note that a similarity-based view in this spirit was also put forward in [35], in the context of belief revision.

5 Conclusion

In this paper, we have contrasted the notion of borderline cases as they relate to linguistic vagueness, with scenarios in information processing where borderline cases need to be explicitly handled. We have, in particular, emphasized that the introduction of degrees may be useful in the latter case—or even required—to appropriately deal with borderline cases. To illustrate this point, we have discussed three different scenarios, in which degrees refer to fuzziness, typicality, and similarity respectively.

Specifically, we have first considered the use of fuzzy labels, focusing on cognitive uses such as expressing flexible requests, describing known states of fact, or summarizing or classifying situations. By interpreting the meaning of linguistic terms as fuzzy sets, the vagueness of these terms becomes a useful feature, as it e.g. provides an intuitive way to convey that a given piece of information should be understood with some tolerance. Note that such uses of fuzzy sets remain in general distinct from their use to compactly encode (crisp) information, which was very early recognized in the case of multiple-valued truth-functional logics by de Finetti [6]. Second, within the framework of formal concept analysis, we have discussed how the degree to which an object is a typical instance of some class can be related to the importance of properties in the definition of this class. Finally, we have addressed the importance of similarity, as the basis for a flexible understanding of properties, for inconsistency management.

While borderline cases play a central role in each of the scenarios we discussed, it is important to notice that the associated degrees are not degrees of truth. We have left aside classical issues related to vagueness such as approximate truth (e.g., [46, 32]) which would deserve longer developments, and should not be confused with uncertainty about binary truth [12].

BIBLIOGRAPHY

[1] Bělohlávek, R. Fuzzy Galois connections. *Math. Logic Quart.*, 45, 497–504, 1999.

[2] Benferhat, S., Dubois, D., Prade, H. Practical handling of exception-tainted rules and independence information in possibilistic logic. *Applied Intelligence*, 9, 101–127, 1998.

[3] Bennett, B. Spatial vagueness. In: *Methods for Handling Imperfect Spatial Information*, (R. Jeansoulin, O. Papini, H. Prade, S. Schockaert, eds.), vol. 256 in Studies in Fuzziness and Soft Computing series, Springer-Verlag, 15–47, 2010.

[4] Bezdek, J.C., Keller, J., Krisnapuram, R., Pal, N.R. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. Kluwer Acad. Publ., 1999.

[5] Czyzowicz, J., Mundici, D., Pelc, A. Ulam's searching game with lies. *J. of Combinatorial Theory*, Series A, 52, 62–76, 1989.

[6] De Finetti, B. La logique de la probabilité, *Actes Congrès Int. de Philos. Scient.*, Paris, 1935, Hermann et Cie Editions, Paris, IV1–IV9, 1936.

[7] Djouadi, Y., Dubois D., L., Prade, H. Graduality, uncertainty and typicality in formal concept analysis. In: *35 years of Fuzzy Set Theory: Celebratory Volume Dedicated to the Retirement of Etienne E. Kerre*, (C. Cornelis, G. Deschrijver, M. Nachtegael, S. Schockaert, Y. Shi, eds.), vol. 261 in Studies in Fuzziness and Soft Computing series, 127–147, Springer Verlag, 2010.

[8] Dubois, D., Esteva, F., Garcia, P., Godo, L., Prade, H. A logical approach to interpolation based on similarity relations, *International Journal of Approximate Reasoning*, 17, 1–36, 1997.

[9] Dubois, D., Esteva, F., Godo, L., Prade, H. An information-based discussion of vagueness: six scenarios leading to vagueness. In: *Handbook of Categorization in Cognitive Science*, (H. Cohen, C. Lefebvre, eds.), Chapter 40, 891–909, Elsevier, 2005.

[10] Dubois, D., Prade, H. Measuring properties of fuzzy sets: a general technique and its use in fuzzy query evaluation. *Fuzzy Sets and Systems*, 38, 137–152,1990.

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

[12] Dubois, D., Prade, H. Possibility theory, probability theory and multiple-valued logics: A clarification. *Annals of Mathematics and Artificial Intelligence*, 32, 35–66, 2001.

[13] Dubois, D., Prade, H. Possibilistic logic: a retrospective and prospective view. *Fuzzy Sets and Systems*, 144, 3–23 2004.

[14] Dubois, D., Prade, H., Sudkamp, T. On the representation, measurement, and discovery of fuzzy associations. *IEEE Trans. on Fuzzy Systems*, 13, 250–262, 2005.

[15] Farreny, H., Prade, H. On the best way of designating objects in sentence generation. *Kybernetes*, 13 (1), 43–46, 1984.

[16] Freund, M., Desclés J.-P., Pascu, A., Cardot, J. Typicality, contextual inferences and object determination logic. *Proc. of the 17th Inter. Florida Artificial Intelligence Research Society Conf.* (FLAIRS), (V. Barr, Z. Markov, eds.), Miami Beach, Fl, USA, 5 p., 2004.

[17] Fine, K., Vagueness, truth and logic. *Synthese*, 54, 235–259, 1975. Reprinted in *Vagueness: A Reader*, (R. Keefe and P. Smith, eds.), MIT Press, Cambridge,119–150, 1996.

[18] Freund, M. On the notion of concepts. I. Artificial Intelligence, 172, 570–590, 2008.

[19] Freund, M. On the notion of concepts. II. Artificial Intelligence, 173, 167–179, 2009.

[20] Gaines, B.R., Foundations of fuzzy reasoning. *Int. J. Man-Machine Studies*, 8, 623–668, 1976.

[21] Hähnle, R. Proof theory of many-valued logic – linear optimization – logic design: connections and interactions. *Soft Computing*, 1, 107–119, 1997.

[22] Jucker, A. H., Smith, S.W., Lüdge, T. Interactive aspects of vagueness in conversation, *Journal of Pragmatics*, 35 (12), 1737–1769, 2003.

[23] Keefe R. *Theories of Vagueness*. Cambridge University Press, 2000.

[24] Keefe R., Smith, P. (eds.) *Vagueness: A Reader*. MIT Press, Cambridge, 1996.

[25] Kraus, S., Lehmann, D., Magidor, M. Nonmonotonic reasoning, preferential models and cumulative logics, *Artificial Intelligence*, 44, 167–207, 1990.

[26] Kyburg, A.: When vague sentences inform: A model of assertability. Synthese 124,175–192, 2000.

[27] Lawry, J. Imprecise bipolar belief measures based on partial knowledge from agent dialogues. *Proc. of the 4th Inter. Conf. on Scalable Uncertainty Management*, LNAI 6379, Springer, 205–218, 2010.

[28] Margolis, E., Laurence, S. Concepts. In: *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), E.N. Zalta (ed.), `http://plato.stanford.edu/archives/fall2008/entries/concepts`.

[29] Mauris, G., Benoit, E., Foulloy, L. Fuzzy symbolic sensors – From concept to applications. *Measurement*, 12, 357–384, 1994.

[30] Medin, D. L., Rips, L. J. Concepts and categories: Memory, meaning and metaphysics. In: The Cambridge Handbook of Thinking and Reasoning, Chapter 3, (K.J. Holyoak, R.G. Morrison (Jr.), eds.), 37–72, 2005.

[31] Meghini, C., Sebastiani, F., Straccia, U. A model of multimedia information retrieval. *J. of the ACM*, 48, 909–970, 2001.

[32] Niiniluoto, I. *Truthlikeness*. D. Reidel, Dordrecht, 1987.

[33] Osherson, D. N., Smith, E. E. On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9, 59–72, 1981.

[34] Pasquier, N., Bastide, Y., Taouil, R., Lakhal, L. Efficient mining of association rules using closed itemset lattices. *Inf. Syst.,* 24(1): 25–46, 1999.

[35] Rodriguez, R., Garcia, P., Godo, L., Relating similarity-based models, counterfactuals and theory change. *Proc. of the 3rd Europ. Congress on Intelligent Techniques and Soft Computing* (EUFIT'95), Aachen, Aug. 28–31, 230–234, 1995.

[36] Rosch, E., Mervis, C. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605, 1975.

[37] Schockaert, S., Janssen, J., Vermeir, D., De Cock, M. Answer sets in a fuzzy equilibrium logic. *Proc. of the 3rd Inter. Conf. on Web Reasoning and Rule Systems*, 135–149, 2009.

[38] Schockaert, S. and Prade, H. An inconsistency-tolerant approach to information merging based on proposition relaxation, In:*Proc. of the 24th AAAI Conf. on Artificial Intelligence*, 363–368, 2010.

[39] Schockaert, S., Prade, H. Solving conflicts in information merging by a flexible interpretation of atomic propositions. Artificial Intelligence, to appear.

[40] Shramko, Y., Wansing, H., Truth Values, *The Stanford Encyclopedia of Philosophy* (Summer 2010 Edition), (E.N. Zalta, ed.), `http://plato.stanford.edu/archives/sum2010/entries/truth-values`.

[41] Smith, N. J. J. Vagueness as closeness. *Australasian J. of Philosophy*, 83, 157–83, 2005.

[42] Sorensen, R. Vagueness. *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), (E.N. Zalta, ed.), `http://plato.stanford.edu/archives/fall2008/entries/vagueness/`.

[43] Sorensen, A. *Vagueness and Contradiction*. Oxford University Press, 2001.

[44] van Fraassen, B. C. Presuppositions, supervaluations and free logic. In: *The Logical Way of Doing Things*, (K. Lambert, ed.), Yale University Press, 67–91, 1969.

[45] Wahlster, W. Implementing fuzziness in dialogue systems. In: *Empirical Semantics*, (B. Rieger, ed.), Brockmeyer, Bochum, 1980.

[46] Weston, T. Approximate truth. *J. of Philosophical Logic*, 16, 203–227, 1987.

[47] Wille, R. Restructuring lattice theory: an approach based on hierarchies of con- cepts. In: *Ordered Sets*, (I. Rival, ed.), Reidel, Dordrecht, 445–470, 1982.
[48] Williamson, T. *Vagueness*. Routledge, London, 1994.
[49] Zadeh, L. A. PRUF: A meaning representation language for natural languages. *Int. J. of Man-Machine Studies*, 10, 395–460, 1978.
[50] Zadeh, L. A. A note on prototype theory and fuzzy sets. *Cognition*, 12, 291–297, 1982.

Henri Prade
IRIT – Institut de Recherche en Informatique de Toulouse
University of Toulouse
118 route de Narbonne
31062 Toulouse Cedex 09, France
Email: prade@irit.fr

Steven Schockaert
Ghent University
Department of Applied Mathematics and Computer Science
Krijgslaan 281, 9000 Gent, Belgium
Email: steven.schockaert@ugent.be