About vagueness, typicality and similarity

Henri Prade

IRIT - CNRS
University of Toulouse
France
e-mail: prade@irit.fr

Introduction

- Vagueness is related to the *informal* idea that

  the contours or limits of the scope of words lack *precision* or *clarity*

- *borderline* cases

- truth values *gap*

a vague concept partitions the universe into more than two parts ?

many **distinct** informational scenarios in *categorization*-like tasks

that may lead to such a situation:

- gradualness / graduality (*fuzzy* set, partial pre-orderings)
- lack of knowledge in the precise extension of a property, *sub-definite* sets
- precisely-defined properties with *closeness* relations
- co-existing distinct views by *multiple agents*
- incomplete information, ill-known attribute values (*twofold* fuzzy sets)
- lack of *discernibility* due to the limited descriptive power of a finite vocabulary (*rough* sets)

importance to distinguish between the scenarios

Dubois, D., Esteva, F., Lluis 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.), ch. 40, 891–909, Elsevier, 2005.

Uncertainty in meaning ?

Here, 'vagueness' is used in a more restricted way ...

Vagueness = 'uncertainty in meaning' ? (Schefe, 1980)

this may be understood in different ways !!

A statement of the form $X\ is\ A$

where X refers to a *single*-valued attribute

conveys a piece of information

- that is *imprecise* if the *extension* of the set representing A is *not a singleton*

  e.g. *'Peter is* **25 or 26***'*, *'John is* **between 20 and 30** *years old'*

- that is *fuzzy* if A is a fuzzy set

  e.g. *'Bob is* **young***'*

Possibilistic representation of imprecise and fuzzy information

In both cases, we are *uncertain* about what is the precise age:

- Peter may be 25 or may be 26 $age(Peter) \in \{25, 26\}$

  $age(John) \in [20, 30]$

- the *possibility* that $age(Bob)$ is 25 is taken as

  the *degree of membership* of 25 to the fuzzy set representing '*young*'

$$\pi_{age(Bob)}(u) = \mu_{young}(u) \ \forall u$$ (Zadeh, 1978)

But ...

- this does not mean that ‘*young*’ is vague, or involve any uncertainty in itself,

  at least as long as there is some agreement on the membership function representing it in the considered context

- the gradualness of a predicate s.t. ‘*young*’ is essentially due to the fact that

  the numerical attribute ‘age’ has an attribute domain which is a *continuum*,

  interfaced with a *discrete set of linguistic labels*

  (here such as ‘*young*’, ‘*old*’, and so on)

- other issue: choice of the *scale* for the membership degrees

  what kind of degrees are meaningful?

Introducing typicality in formal concept analysis

Formal Concept Analysis (FCA) (Wille, 1982)

defines concepts as (**extension**, **intension**)-pairs (X, Y)

s. t. $X = \{x \in Obj | R(x) \supseteq Y\}$ and $Y = \{y \in Prop | R^{-1}(y) \supseteq X\}$

from a relation R, called '**context**', which states for each object $x \in Obj$ the complete set of its properties $y \in Prop$.

- *Gradualness* in properties can be taken into account by allowing **R** to be **fuzzy** (Belohlavek).
- *Typicality* can be introduced in FCA *by keeping* **R crisp**, and introducing degrees among objects and among properties.

two principles

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

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

Djouadi, Y., Dubois D., L., Prade, H. On the possible meanings of degrees when making formal concept analysis fuzzy. Proc. Eurofuse'2009*, Pamplona, Sept.16-18.*

Bird example

Table 1:

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

Example: What is a bird?

birds: $X = \{albatross, parrot, penguin, kiwi\}$

bird properties: Y=$\{$'$laying\ eggs$', '$having\ two\ legs$', '$flying$', '$having\ feathers$'$\}$)

typicality $X^{\mathbf{t}}$ $X^t(albatross) = X^t(parrot) = 1$, $X^t(penguin) = \alpha$, $X^t(kiwi) = \beta$

with $1 > \alpha > \beta$ (kiwis do not fly and have no feathers).

- fuzzy set of *important properties*, according to *(B)*

  $Y^i(y) = min_x X^t(x) \rightarrow R(x, y)$, with **a** $\rightarrow$ **1 = 1** and **a** $\rightarrow$ **0 = 1 - a**

  It expresses that an object not having property y makes a property all the less important in the definition of the concept *bird* as this bird is considered more typical.

- Let $Y^i(y)$ define the degree of importance of property y, in the definition of *bird*, $\forall y$.

  fuzzy set of *typical objects*, according to *(A)*

  $\mu(x) = min_y Y^i(y) \rightarrow R(x, y)$, using$(1 - a) \rightarrow 0 = a$

  We get $\mu(albatross) = Y^i(parrot) = 1$, $\mu(penguin) = \alpha$; $\mu(kiwi) = \beta$

- We have $\forall x$ $\mu(x)$ = $X^t(x)$ a (fuzzy) *Galois connection*

Representing 'Tweety is a bird'

'Tweety is a bird'

Y^i_{bird} the fuzzy set of *important properties* for birds

$\forall y \in Prop$

$$\pi_{Tweety}(y^c) = 1 - Y^i_{bird}(y)$$

$$\pi_{y(Tweety)}(no) = 1 - Y^i_{bird}(y)$$

where y^c is the negation of y

- the *possibility* that Tweety has *not* property y is all the greater as y is less important for birds
- the *certainty* that Tweety has property y is all the greater as y is more important for birds
- to parallel with 'Bob is young' $\pi_{age(Bob)}(u) = \mu_{young}(u)$

Zadeh about vagueness

vagueness should not be confused with *fuzziness*

"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." (Zadeh, 1978)

- We propose to understand 'vague' in a slightly different way

  The statement $X\ is\ A$ is regarded as *vague*, whatever the specificity of A,

  - only if the membership function of A is **open to variability**

  - or in other words if the *precise location* of the border of A is *not known*

- A is more like a Gentilhomme *flou set*

  where some elements are **certainly** in A, and others are only **possibly** in A.

Taking advantage of vagueness in dialogues

HAM-RPM dialogue system in natural language (German) (for hotel reservation)

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

- acknowledges the role of *fuzziness* in dialogues
- *approximate* matching of a requirement with a state of fact
- also the idea that 2 agents may use the *same word* with slightly *different meaning*: e. g. "*a room quite large*"

Vagueness pervades social interaction, e. g. bargaining

Need for different representations for a vague term

according to the situation

- choosing a vague label in agreement with a **precisely known** state of facts
  and maybe with some goal in mind, as in the example of the hotel manager:
  "*our rooms are large*"
  different labels are *more or less possible* for describing the same state of facts
- a vague piece of information is received
  this is **all what is known**
  e.g. for the customer, the room is said to be *large*
  $\pi_{size(room)}(u) = \mu_{large^*}(u)$ where u ranges between 10 and 30 m^2
  $large^* =$
  modified/discounted view of customer's understanding of *large* (for a ** hotel room)
- finding a *linguistic* (compound) label for an *ill-known situation*
  "*linguistic approximation*" (Zadeh, 1978)

Approximate truth and similarity-based revision

while p is false, p^* may be true, when p entails p^*

if *all* the models of p^* are **close** to a model of p

p "*not far to be true*"

Revision:

- getting rid of the *less entrenched* propositions
- or weakening propositions

  by *enlarging* their set of models by *similarity* (if possible)

  for trying to restore *consistency*

Applies as well for *fusion*

Vague understanding for maintaining consistency

2 reliable sources: K_1, K_2

K_1 says 'tomorrow the sky will be overcast'(oc)

K_2: says 'tomorrow the sky will be open' (os)

Classical merging operators Δ lead to $\Delta(K_1, K_2) = \{oc \vee os\}$

Let us introduce the literal pc = 'tomorrow the sky will be partially cloudy'

which is *close* both to oc and oc

and have a *vague* reading of sources as K_1 =$\{oc \vee pc\}$ and K_2 =$\{pc \vee os\}$.

This enlargement of propositions by proximity enables us to get $\Delta(K_1, K_2)$ =$\{pc\}$

- can be generalized by allowing *progressive enlargements* of the meaning of propositions using *possibilistic logic*

  S. Schockaert, H. Prade. Merging conflicting propositional knowledge by similarity. Proc. 21st IEEE Int. Conf. on Tools with Artif. Intell. (ICTAI-09), Nov. 2-4, 2009, Newark, US

Example 2

More *jointly exhaustive* and *pairwise disjoint* literals between oc and os say $pc_{-1} \vee pc_0 \vee pc_1$ with easy-to-guess meanings

Then, K_1 and K_2 may now be weakened into

$$K_1 = \{oc \vee pc_{-1} \vee pc_0\} \text{ and } K_2 = \{os \vee pc_0 \vee pc_1\}$$

and the result will be $\Delta(K_1, K_2) = \{pc_0\}$

A more liberal interpretation:

$$K_1 = \{oc \vee pc_{-1} \vee pc_0 \vee pc_1\} \text{ and}$$

$$K_2 = \{os \vee pc_{-1} \vee pc_0 \vee pc_1\}$$

we get $\Delta(K_1, K_2) = \{pc_{-1} \vee pc_0 \vee pc_1\}$

Possibilistic logic encoding

Example

$$K = \{\neg p \vee q, r\}$$

is changed into

$$K_{pos} = \{(\neg p \vee q, \beta), (r, \beta),$$
$$(\neg p_* \vee q^*, \alpha), (r^*, \alpha),$$
$$(\neg p_{**} \vee q^{**}, 1), (r^{**}, 1),$$
$$(\neg p \vee p^*, 1), (\neg p^* \vee p^{**}, 1), (\neg p_* \vee p, 1), (\neg p_{**} \vee p_*, 1)\}$$

with $1 > \alpha > \beta > 0$

Level of inconsistency $K = K_1 \cup K_2$

$Inc(K_{pos}) = \beta$ then $\Delta(K_1, K_2) = (K_{pos})_\alpha$

the set of formulas with level at least equal to α

Conclusion

- vagueness $\neq$ gradualness
- vagueness $\neq$ (non) typicality
- vagueness $\neq$ lack of specificity, generality
- vagueness $\neq$ ambiguity

  (to be related to a lack of **convexity**:

  comes from Latin ‘anceps’ = which has two heads)

- vagueness = extension of the concept open to variability

  vagueness is a matter of *flexibility in meaning*

  vagueness may be useful !