



HAL
open science

Vagueness, typicality, and uncertainty in class hierarchies

Didier Dubois, Henri Prade, Jean-Paul Rossazza

► **To cite this version:**

Didier Dubois, Henri Prade, Jean-Paul Rossazza. Vagueness, typicality, and uncertainty in class hierarchies. *International Journal of Intelligent Systems*, 1991, *Fuzzy Methods for Knowledge-Based Systems*, 6 (2), pp.167–183. 10.1002/int.4550060205 . hal-04074781v2

HAL Id: hal-04074781

<https://hal.science/hal-04074781v2>

Submitted on 13 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vagueness, Typicality and Uncertainty in Class Hierarchies

Didier Dubois, Henri Prade, Jean-Paul Rossazza

Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex (France)

The paper presents an object-centered representation, where both a range of allowed values and a range of typical values can be specified for the attributes describing a class. These ranges may be fuzzy. Then various kinds of (graded) inclusion relations can be defined between classes. Inheritance mechanisms are discussed in this framework, as well as other kinds of reasoning tasks such as classification. The architecture of a software system implementing these ideas is outlined.

1. INTRODUCTION

The object-centered representations (O.C.R.) stem from the notion of frames expressed by Minsky¹ and the concept of object defined in Smalltalk². They model knowledge as explicit entities named objects (or frames) rather than by means of relations bearing on entities, as in predicative and relational formalisms. An object may be an abstract concept (for example the concept "bird") or a real entity (the living bird named Tweety). So in an O.C.R. the objects have an explicit existence inside the knowledge base: all the information concerning an object is gathered at the same place and forms this object. The pieces of information are not scattered in the base as they are in predicative and relational formalisms. The objects having the same properties are gathered into classes which are organized into hierarchies. This organization allows an object to inherit properties from its class (if the object is an instance) or from its superclasses (if the object is a class) thus simplifying the specification of the knowledge base. Moreover a system using such an organization can easily access its knowledge and so explain its reasoning. An O.C.R. may be used in association with production rules or as one among different representation methods the user can choose.

From a cognitive point of view an O.C.R. is interesting because the human being seems to use both the notions of hierarchy and class in his reasoning. By using such hierarchies we can hope on the one hand that the system will be able to reason

in a way closer to the human mind and on the other hand that the knowledge base will be more understandable and easier to model. But the reasoning process of a human being is much more complex than the simple use of class hierarchies: a human being can plausibly answer questions even if he has not all the necessary information for a sure answer. He can do that because he uses different kinds of reasoning patterns (see e.g., Collins and Michalski³): generalization, specialization, analogy and default reasoning, and because he can handle vagueness, uncertainty and typicality.

So, in order to accommodate various forms of plausible reasoning we need an O.C.R. allowing to cope with typicality, uncertainty and vagueness. Although object-centered representations and possibility theory can get along quite easily, there have been only a few attempts to mix them: let us mention Rundensteiner and Bandler⁴ who rather study fuzzy semantic networks, Vignard⁵ who introduces fuzzy-valued attributes in classes, Graham and Jones^{6,7} who propose a generalization of the notion of frame allowing for partial inheritance, Granger⁸ who solves classification problems, using a class hierarchy with fuzzy thresholds and weighted attributes and Torasso and Console⁹ who define a frame-based representation with three kinds (necessary, sufficient and supplementary) of weighted attributes.

In the approach proposed in this paper, classes are intensionally described in terms of attributes for which we distinguish between the range of *allowed* values and the range of *typical* values, as explained in the following section. These ranges are described by fuzzy sets¹⁰ in the most general case. The existence of two kinds of range for each attribute induces four basic types of possible inclusion between two classes. In section 2 suitable degrees of inclusion are discussed in order to take into account the possible fuzziness of ranges in the evaluation. Section 3 of the paper briefly presents the architecture of a system based on such classes, and section 4 deals with inference issues in this framework.

II. FUZZY CLASSES OF OBJECTS AND VALUES OF ATTRIBUTES

A class can theoretically be considered from two different view points: i) an extensional one where the class is defined by the list of its members; ii) an intensional one where the class is defined by a set of attributes and their admissible values (or equivalently in terms of a compound predicate). Practically we can only use the intensional definition. In the following a class is defined by means of a set of *attributes*; each attribute has a *domain*, a *range* and a *typical range*. Only single-valued attributes are considered in this paper. The domain of an attribute *a* is the set of all the values the attribute may take for all objects; the attribute domain does not depend on the class. The range and the typical range will allow us to define a class intensionally and to capture the notion of typicality. When speaking of a class, we will use the word 'range' for expressing what is usually called an attribute value in a classical O.C.R. and distinguishing it from a value in a range.

A. The Range and the Typical Range of an Attribute

The range $R(a,C)$ of an attribute a of a class C is the set of allowed values that members of C can take for the attribute a . The range of an attribute may be fuzzy because some of its values are regarded as *atypical* (i.e. less possible or feasible than other values): then an atypical value of the range will be assigned a membership degree strictly less than 1. However a nonatypical value will not necessarily be a somewhat typical value. We want to point out that a class is fuzzy not because of a lack of information but in order to allow for flexibility and encompass borderline cases.

The typical range $T(a,C)$ of an attribute a of the class C is the set of more or less typical values that a member of C can take for a . The typical range can naturally be represented by a fuzzy set because the notion of typicality is intrinsically gradual. Moreover, before being somewhat typical a value must be not atypical, that is if we write $Su(T(a,C))$ for the support of $T(a,C)$ and $Co(R(a,C))$ for the core of $R(a,C)$ we should have

$$Co(R(a,C)) \supseteq Su(T(a,C)) \quad (1)$$

where the support of a fuzzy set is the set of elements whose membership is strictly positive, and the core is the set of elements whose membership is equal to 1. In practice, an equality rather than an inclusion will often be assumed since it might be difficult to imagine a nontypical value (not belonging whatsoever to $T(a,C)$) that is at the same time not atypical (belonging to $Co(R(a,C))$). When a typical range is not explicitly stated, we will implicitly consider, in order to be coherent with the above requirement, that the core of the range is the typical range. The typical ranges may of course be used as (fuzzy) default values¹¹ that apply to an instance of C taken at random; they may also be viewed as defining class prototypes^{12,13}.

Let us consider a simple example:

class	Birds
attribute	way_of_locomotion
domain	{fly, walk, swim, crawl, ...}
range	{1/fly, 0.6/walk, 0.2/swim}
typical range	{1/fly}
...	...

Here, as it has been said, the attribute `way_of_locomotion` is considered as being single-valued (in fact we can consider that we are representing the *main* way of locomotion of birds). Some values of the range have received a membership degree (indicated before the '/') different from 1 in order to point out their atypicality (here the numbers mainly reflect an ordering). In the example the nonatypical value is the typical value.

When none of the ranges of the attributes of a class is fuzzy, we have an ordinary class. Clearly we are free to allow or not to allow the ranges or the typical ranges of attributes to be fuzzy. This a matter of option; in practice it is perhaps not necessary

to allow both for fuzzy ranges and fuzzy typical ranges.

B. Subclasses

In classical O.C.R., a subclass $C2$ is defined from its superclass $C1$ either by specialization of its ranges (i.e. the subclass ranges must be included in the corresponding ranges of $C1$) or by addition of new attributes. In our representation the subclasses are going to be defined in a similar way once we have specified what does 'included' mean for fuzzy ranges. However the distinction between range and typical range induces the existence of four types of inclusion depending if we are comparing the ranges or the typical ranges of $C1$ with the ranges or the typical ranges of $C2$. Obviously this comparison may be limited to the context corresponding to a prescribed subpart of the attributes involved in the descriptions of $C1$ and $C2$.

1. Inclusion of Two Fuzzy Ranges

The degree of inclusion $N(B \mid A)$ of a fuzzy set A into a fuzzy set B can be defined by the expression¹⁴

$$N(B \mid A) = \text{Inf}_u \{ \mu_A(u) * \rightarrow \mu_B(u) \} \quad (2)$$

where $* \rightarrow$ denotes a fuzzy implication. We have four main inclusion degrees exemplified by considering the following fuzzy implications: Gödel's, Dienes', Lukasiewicz' and the reciprocal of Gödel's implication (see the annex for their different properties). The minimal requirement we demand for the fuzzy ranges $A = R(\mathbf{a}, C2)$ and $B = R(\mathbf{a}, C1)$ in order to say that the inclusion of A into B holds to some extent is the following

$$\text{Su}(R(\mathbf{a}, C1)) \supseteq \text{Su}(R(\mathbf{a}, C2)) \quad (3)$$

i.e. there are no values of \mathbf{a} which somewhat belong to A (i.e. which are somewhat allowed in $C2$) and which do not at all belong to B (i.e. which are forbidden in $C1$). This property is ensured by the degree of inclusion based on Gödel implication (see the annex). This is why this degree of inclusion, denoted N_G in the following, is chosen for defining the inclusions between classes. Note that

$$N_G(B \mid A) = 1 \Leftrightarrow \mu_A \leq \mu_B$$

that is, if B contains A , any object belongs to B at least as much as it belongs to A .

Moreover

$$N_G(B \mid A) > 0 \Rightarrow \forall u, \mu_A(u) > 0 \text{ and } \mu_B(u) = 0;$$

i.e. (3) holds, as expected.

2. Inclusion Between Classes

Since a class is defined through the conjunction of its attributes (and of their ranges), we define the inclusion degree between two classes as a conjunctive aggregation of the inclusion degrees of their ranges. The attributes are supposed to be noninteractive¹⁵ (i.e. any n-tuple of values on the Cartesian product of the ranges is (at least somewhat) allowed for an instance of the class). To perform the conjunction we want to choose two associative conjunctive operations \times and $*$ verifying

$$N_G(B_1 \times \dots \times B_n \mid A_1 \times \dots \times A_n) = * \{N_G(B_i \mid A_i) \mid i \in \langle 1, n \rangle\} \quad (4)$$

where \times stands for the Cartesian product and $*$ for the conjunction of the degrees; $* \{a_i \mid i \in I\}$ is the result of combining the a_i 's by operation $*$. It can be proved (see annex) that if the Cartesian product is defined by means of min operation (in order to be in agreement with the noninteractivity of the attributes¹⁵) we should have $*$ = min in (4), provided that A_1, \dots, A_n are normalized (i.e. $\forall i, \exists ui \mu_{A_i}(ui) = 1$), so we can define the inclusion degree $N_G(C_1 \mid C_2)$ of C_2 in C_1 with respect to the ranges as

$$N_G(C_1 \mid C_2) = \min \{N_G(R(a, C_1) \mid R(a, C_2)) \mid a \text{ is an attribute of } C_2\} \quad (5)$$

where $\min\{f(a) \mid a \in \mathcal{A}\}$ denotes the value of the smallest of the $f(a)$'s when a ranges in \mathcal{A} . Here we consider that all the attributes of C_2 which are not explicitly mentioned as attributes of C_1 can always be implicitly considered as attributes of C_1 with a range equal to their domain (for these ranges Gödel's inclusion degree is equal to 1) so the minimum in (5) can be taken on the attributes of C_1 . Note that

$$N_G(C_1 \mid C_2) > 0 \Leftrightarrow \forall a, N_G(R(a, C_1) \mid R(a, C_2)) > 0.$$

Similarly we define the three other inclusion degrees $N_G(C_1 \mid T(C_2))$, $N_G(T(C_1) \mid C_2)$ and $N_G(T(C_1) \mid T(C_2))$ where $T(C_i)$ means that we substitute $T(a, C_i)$ for $R(a, C_i)$ in (5). It is clear that the following inequalities hold due to (1)

$$N_G(C_1 \mid T(C_2)) \geq N_G(C_1 \mid C_2) \geq N_G(T(C_1) \mid C_2) \quad (6)$$

$$N_G(C_1 \mid T(C_2)) \geq N_G(T(C_1) \mid T(C_2)) \geq N_G(T(C_1) \mid C_2) \quad (7)$$

Given a collection of classes described in terms of ranges and typical ranges of attributes, it is then possible to build a valued-graph expressing the different inclusion relationships between classes and their degrees. It can be interpreted as a weighted semantic network. The definitions of the inclusion degrees can be restricted to a context expressed by selecting a subpart of the attributes involved in the description of classes. It can be checked that N_G satisfies the following transitivity property

$$N_G(C \mid A) \geq \min(N_G(C \mid B), N_G(B \mid A)). \quad (8)$$

Thus by transitivity lower bounds of degrees of inclusion can be easily obtained. The

use of this property in classification processes is emphasized in section 4.2.

C. The Possible Range and the Credible Range of an Instance

We have called instances, the terminal objects of the hierarchy which are not able to generate other objects. Like classes, they are described in terms of attributes. The available knowledge, possibly incomplete about an instance x is represented for an attribute a by a *possible* range $P(a,x)$ and a *credible* range $\text{Cred}(a,x)$. $P(a,x)$ is the (fuzzy) set of all the more or less possible values of the attribute a for the instance x . Its membership function is viewed as a possibility distribution acting as an elastic constraint on the value of the attribute a for x . It is supposed to be a sure information (since given by the operator or by inheritance). $\text{Cred}(a,x)$ denotes a range obtained in an *uncertain* way (i.e. by a plausible reasoning technique whose validity is not guaranteed). It will in particular allow for default reasoning. In order to be coherent with the possible range, the credible range has to satisfy the following relation (which is weaker than (1)):

$$\text{Su}(P(a,x)) \supseteq \text{Co}(\text{Cred}(a,x)) \quad (9)$$

that is to say each completely credible value is at least a little possible. Unlike a class range, $P(a,x)$ is a set of mutually exclusive values: the real value of the (single-valued) attribute a for the instance x is in $\text{Su}(P(a,x))$ but we do not know precisely where.

Note that (9) is a very weak relation: as we know that no value outside $P(a,x)$ is allowed, a stronger relation could be $\text{Su}(P(a,x)) \supseteq \text{Su}(\text{Cred}(a,x))$. However a plausible reasoning may produce uncertain credible ranges which would not satisfy this relation (indeed an uncertain fuzzy value has a membership function which is not zero anywhere: since nothing is certain, everything remains somewhat possible, may be with a low degree). Their uncertainty would then have to be restricted to $\text{Su}(P(a,x))$.

The uncertainty of a credible range arises from the fact that it has been produced as a default range, whereas the uncertainty of a possible range denotes the uncertainty of the piece of information itself. Thus in the following example

Tweety	
Way_of_locomotion:	Possible range {1/fly, 0.6/walk, 0.2/swim}
	Credible range {1/fly}

the possible range comes from the knowledge that Tweety is a bird, and the credible range is obtained by default reasoning. If the source of information only asserts that it is 0.8 certain that Tweety is a bird, the possible range would be {fly, 0.6/walk, 0.2/swim 0.2/crawl, 0.2/...} since any value of the domain remains in this case possible at the degree $1 - 0.8 = 0.2$. Hence the approach distinguishes between plausible values whose uncertainty is due to default reasoning (the credible ranges) and plausible values whose uncertainty is due to the lack of reliability of the source that provides the information (uncertain possible ranges supplied by the user).

D. Certainty of Membership of an Object in a Class

When the value $P(\mathbf{a},x)$ of an attribute \mathbf{a} for an object x is precisely known (i.e. corresponds to a singleton of the domain of \mathbf{a} , its membership degree to the range $R(\mathbf{a},C)$ of the class C is given by $\mu_{R(\mathbf{a},C)}(P(\mathbf{a},x))$; more generally when $P(\mathbf{a},x)$ is imprecisely or fuzzily known, i.e. $P(\mathbf{a},x)$ is represented by a possibility distribution which does not restrict a unique possible value, we only have a bracketing of this degree; see Cayrol et al¹⁶ on this point. Therefore we want to be sure that the real value of the attribute for the considered object is in the range of the class even if the corresponding instance value is fuzzily known. This leads to use Dienes' inclusion degree (see the annex) since

$$N(R(\mathbf{a},C) \mid P(\mathbf{a},x)) = 1 \Leftrightarrow \text{Co}(R(\mathbf{a},C)) \supseteq \text{Su}(P(\mathbf{a},x)) \quad (10)$$

i.e. we conclude in favor of completely certain (and a full) membership of x to C if and only if all the more or less possible values of $P(\mathbf{a},x)$ are among the values which undisputedly belong to the range. The membership degree evaluation of an object x to a class C is defined as the conjunction of the membership degree evaluations with respect to ranges defining C . Dienes' inclusion degree can be factorized¹⁷ as Gödel's inclusion degree, provided that each $P(\mathbf{a},x)$ is normalized, i.e.

$$N(C \mid x) = \min \{N(R(\mathbf{a},C) \mid P(\mathbf{a},x)) \mid \mathbf{a} \text{ is an attribute of } C \} \quad (11)$$

$N(C \mid x)$ represents the certainty for x to be in C , in spite of the ill-known description of x . In the same manner we may want to know the extent to which x is a typical instance of the class C . It is estimated by $\min\{N(T(\mathbf{a},C) \mid P(\mathbf{a},x)) \mid \mathbf{a} \text{ is an attribute of } C\}$. It is also possible to compute credible degrees by using the credible ranges of x . Finally, it can be checked that we have a transitivity property analogous to (8):

$$N(C1 \mid x) \geq \min(N(C2 \mid x), N_G(C1 \mid C2)). \quad (12)$$

Note that $N(C \mid x)$ is generally not a degree of membership (except if x is precisely described and/or $N(C \mid x) = 1$). It is a degree of certainty of membership. Particularly $N(C \mid x) = 0$ does not mean that x does not belong to C , but only that there is total ignorance about it. In order to conclude that x does not belong to C , we must ensure that the degree of possibility of membership $\prod(C \mid x) = 1 - N(\neg C \mid x)$ is zero, where $\neg C$ is the complement of C (defined by taking the complement to 1 of the membership functions of the ranges). However it is not usual to compute $\prod(C \mid x)$ in usual object-centered representations.

E. Handling Typicality

The way the classes are described, as well as the various kinds of inclusion relations which can be computed from the descriptions or declared by the user, enables us to handle typicality. A statement like "typically, students are adults" can be understood as $T(\text{Stu}) \subseteq \text{Adu}$, while a statement like "typically adults are

employees" may rather mean $T(\text{Adu}) \subseteq T(\text{Emp})$, where Stu , Adu and Emp denote the classes of students, adults and employees respectively, and $T(C)$ is the subclass of typical elements of class C . These statements are for instance compatible with the following descriptions

Stu		Adu	
Age	$R: \geq 15$ $T: [18,25]$	Age	$R: \geq 18$ $T: [25,60]$
Income	$R: \{\text{scholarship, parents' money, salary}\}$ $T: \{\text{scholarship, parents' money}\}$	Income	$R: \langle \text{any kind} \rangle$ $T: \{\text{salary}\}$
Emp			
Age	$R: [16,65]$ $T: [20,60]$		
Income	$R: \{\text{salary}\}$ $T: \{\text{salary}\}$		

Note that the description blocks undesirable transitivity effects since $T(\text{Stu}) \subseteq \text{Adu}$ \mathbf{Z} $T(\text{Adu}) \subseteq T(\text{Emp})$; thus we cannot deduce that a typical student is an employee.

III. SKETCH OF THE ORGANIZATION OF THE PROPOSED OBJECT-CENTERED REPRESENTATION SYSTEM

The heart of the system¹⁸ implementing the above-presented approach contains a knowledge base and a supervisor. The knowledge base is divided into a hierarchical component and a descriptive component. The hierarchical component can be viewed as a simplified semantic network, based on weighted "is-a" links (between an instance and its classe(s)) and "a-kind-of" links (between classes). It allows us to draw quick inferences when no description is required. The descriptive component provides a semantic justification to the weighted hierarchy: every hierarchical link has to refer to an inclusion or membership relation in terms of intensional description. The existence of two modes of representation within the system recalls but contrasts with Coupey's proposal in the RS¹⁹ system, where a NETL-like²⁰ semantic network is coupled together with a syntactic declarative description in a variant of Reiter's default logic²¹.

The supervisor monitors the execution of complex requests, possibly by calling specific procedures of the base manager for reasoning steps, computing inclusion degrees... It has to determine which requests of the base manager have to be called, and how they must be called. Further research is needed about it.

Any access to the base requires the base manager. It gathers the different functions manipulating the knowledge base objects. These functions may be called by the user interface for consultation or updating requests, or by the supervisor when the processing of a complex request requires knowledge about some objects of the base.

The objects description requires particular data structures in order to represent fuzzy ranges. The symbolic and numerical processor is specially designed to cope

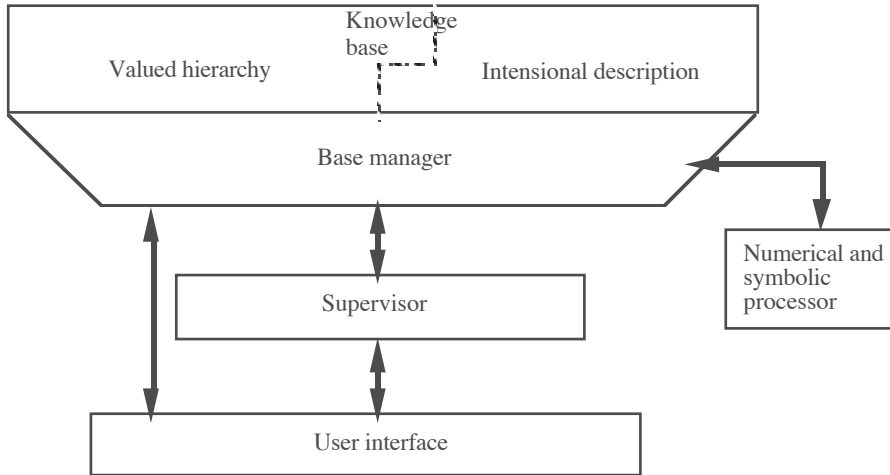


Figure 1. General organization of the object-centered representation system

with these structures. Particularly, it can create fuzzy ranges, perform intersection, compute inclusion degrees between ranges, or evaluate membership degrees of instances.

IV. VARIOUS REASONING MODES

Inheritance and classification are two opposite ways of reasoning: inheritance tries to deduce the properties of an object according to its position in the hierarchy, whereas classification tries to locate an object in a hierarchy according to its properties. The expressive power of our representation offers not only these two ways of reasoning but also variations thereof. Finally, we present some other kinds of reasoning as well. In the following, when a class $C2$ is declared to be a subclass of $C1$, it implicitly means that the corresponding degree of inclusion (based on N_G) is equal to 1.

A. Inheritance

Three kinds of inheritance are available: typical inheritance, 'normal' inheritance and atypical inheritance. Describing $C2$ as a typical subclass of $C1$ implies that $C2$'s possible ranges inherit $C1$'s typical ranges. Describing $C2$ as a 'normal' subclass of $C1$ implies that $C2$'s possible ranges inherit $C1$'s possible ranges and $C2$'s typical ranges inherit $C1$'s typical ranges. The atypical inheritance is a bit more complex: describing $C2$ as an atypical subclass of $C1$ means that at least one attribute of $C2$ is atypical, but not necessarily all of them. So atypical inheritance requires to specify which attributes are atypical. For each of these attributes, $C2$'s possible range inherits the intersection of $C1$'s possible range and of the complement of $C1$'s typical range (the complement $\neg F$ of a fuzzy set F is defined by $\mu_{\neg F} = 1 - \mu_F$ as usual). Other attributes of $C2$ inherit in a 'normal' way.

Consider, for example, the "Mammal" class below

Mammal		
way_of_birth	Range	{1/viviparous, 1/oviparous}
	Typical range	{1/viviparous}
way_of_locomotion	Range	{0.2/fly, 1/walk, 0.3/jump, 0.5/swim}
	Typical range	{1/walk}
skeleton	Range	{1/yes}
	Typical range	-
suckle	Range	{1/yes, 0.5/no}
	Typical range	{1/yes}

Saying "Dogs are typical mammals" will cause

Dogs		
way_of_birth	Range	{1/viviparous}
	Typical range	-
way_of_locomotion	Range	{1/walk}
	Typical range	-
skeleton	Range	{1/yes}
	Typical range	-
suckle	Range	{1/yes}
	Typical range	-

whereas "African mammals are mammals" will cause

African Mammals		
way_of_birth	Range	{1/viviparous, 1/oviparous}
	Typical range	{1/viviparous}
way_of_locomotion	Range	{0.2/fly, 1/walk, 0.3/jump, 0.5/swim}
	Typical range	{1/walk}
skeleton	Range	{1/yes}
	Typical range	-
suckle	Range	{1/yes, 0.5/no}
	Typical range	{1/yes}

and "cetaceans are atypical mammal due to way_of_locomotion" will imply

Cetacean		
way_of_birth	Range	{1/viviparous, 1/oviparous}
	Typical range	{1/viviparous}
way_of_locomotion	Range	{0.2/fly, 0.3/jump, 0.5/swim}
	Typical range	-
skeleton	Range	{1/yes}
	Typical range	-
suckle	Range	{1/yes, 0.5/no}
	Typical range	{1/yes}.

In the last example, the system will renormalize the fuzzy range "way_of_locomotion". Here the result will be intuitively good but in the case of atypical inheritance the obtained possible ranges may not be adequate as long as nothing is said about the kind of atypicality involved in the new subclass. For instance the atypical inheritance would give strange results for bats in the above example. We can see that the user must refine this description by specifying the actual kind of atypicality of the new subclass. The consistency of any such additional information must obviously be checked.

The locality property of atypical inheritance sometimes applies to typical inheritance, if the latter is understood as related to a "typical" attribute of a class: in other words, if $C2$ is declared to be a typical subclass of $C1$, it may mean that typical inheritance applies for the typical attributes of $C1$ only and normal inheritance for other attributes. For instance the typical attributes for birds would include the presence of feathers and the flying capability, but not the colour.

The above three kinds of inheritance can be refined by using only a subset of $C1$'s attributes. Such a subset may be viewed as the context in which $C2$ is a subclass of $C1$. Mind that an atypical inheritance in a restricted context is different from an atypical inheritance where the atypical attributes are specified.

The multiple inheritance problem is solved in a simple way: the subclass ranges inherit the intersection of the ranges of the super-classes, each super-class range being inherited according to the nature of inheritance. According to the types of inheritance different conditions must be satisfied for the amount of overlapping of the ranges or typical ranges of the involved superclasses. For instance if C inherits from $C1$ and $C2$ in a normal way, then $\forall a, R(a,C1) \cap R(a,C2)$ must be normalized. Otherwise it would mean that only atypical elements of $C1$ or $C2$ belong to C . Similarly $T(a,C1) \cap T(a,C2)$ should not be empty. When C is a typical subclass of both $C1$ and $C2$, then $T(a,C1) \cap T(a,C2)$ should be normalized. In case of double atypical inheritance, however, $T(a,C1) \cap T(a,C2)$ may be empty, and $R(a,C1) \cap R(a,C2)$ must be not empty, although not necessarily normalized. All these conditions are coherence requirements whose violation should be notified to the user. Lastly note that multiple inheritance can be dissymmetric; C can be an a typical subclass of $C1$ and atypical subclass of $C2$, which leads to still another kind of coherence condition.

In the same way as we have three kinds of inheritance for a class, we have three kinds of inheritance for an instance. Describing x as a typical C implies that the possible ranges of x inherit typical ranges of C . Describing x as a 'normal' C implies that the possible ranges of x inherit possible ranges of C and that the credible ranges of x inherit typical ranges of C . We also have the same process of atypical inheritance with an instance as with a class. Again, the inheritance may be restricted to a context, again we perform the intersection of the involved ranges in case of multiple inheritance, and again the consistency of any additional information has to be checked.

When x is simply described as a C , its typical ranges may be viewed as more precise but more uncertain ranges than x 's possible ranges: that's why they are inherited as x 's credible ranges. Note that by doing this we perform a kind of default reasoning.

B. Classification

Classification consists in best locating a (possibly ill-described) instance or class in the hierarchy. The natural way is to look for membership or inclusion of the considered item in the most general classes first, and then to go down along the branches of the hierarchy as far as possible.

The great advantage of our representation with regard to the classification comes from the fact that it is useless to carry on the classification process on subclasses of a class $C1$ if the classification process has already failed on $C1$. This is a consequence of the transitivity relations (8) and (12) since given $N_G(C1 \mid C2) > 0$ and $N_G(C1 \mid C) = 0$ (or $N(C1 \mid x) = 0$), (8) and (12) gives

$$\begin{aligned} 0 = N_G(C1 \mid C) &\geq \min(N_G(C1 \mid C2), N_G(C2 \mid C)) \\ \text{and} \quad 0 = N(C1 \mid x) &\geq \min(N_G(C1 \mid C2), N(C2 \mid x)) \end{aligned}$$

which entails $N_G(C2 \mid C) = 0$ and $N(C2 \mid x) = 0$. So, the semantic network underlying our representation helps the classification process and is useful for fast (bottom-up) inferences. In a general way, representations allowing such cuts in a classification process are not able to deal with exceptions, whereas representations dealing with exceptions have difficulties to perform classification because they have to examine all the objects of the base. Our formalism deals with both of them.

An object (even if its description is imprecise) may belong to several classes with various degrees of membership. The results of a classification process is given as a list of all the somewhat interesting classes.

We are also able to perform a kind of "credible" classification. Its main idea is to refine the description of the unknown object by using the typical ranges of the classes for which the classification process succeeds. For example if the process succeeds in comparing Tweety and Bird, we can continue the classification on Bird's subclasses taking for granted the fact that Tweety flies.

C. Other Reasoning Modes

1. Other Indices

The relations between two objects can be refined by using degrees other than Dienes and Gödel inclusion indices. We have:

- the possibility degree that an instance x belongs to a class C : $\prod(C \mid x) = \min\{\prod(R(a,C) \mid P(a,x)) \mid a \text{ is an attribute of } C\}$; where $\prod(R(a,C) \mid P(a,x)) = \sup_u \min(\mu_{R(a,C)}(u), \mu_{P(a,x)}(u))$;
- the intersection degree of two classes $C1$ and $C2$, given by $\prod(C1 \mid C2) = \min_a \{\prod(R(a,C1) \mid R(a,C2))\}$;
- the resemblance between two classes $C1$ and $C2$. It can be computed by symmetric indices of the form

$$\min\{\min_{i=1..n} N(R(a_i, C1) \circ T \mid R(a_i, C2)), \min_{i=1..n} N(R(a_i, C2) \circ T \mid R(a_i, C1))\}$$

where N defines a degree of inclusion in the sense of (2) and T is a tolerance (fuzzy) relation such that $A \circ T \supseteq A$. The composition $A \circ T$ being defined in terms of membership function by $\mu_{A \circ T}(u) = \sup_{u'} \min(\mu_A(u'), \mu_T(u, u'))$.

2. Search for a context

Rather than computing indices between two objects, we may prefer to know on which attributes these objects satisfy a given relation. For example, we can ask "in which context cetaceans are fish" (the response would be "by their place of living", "by their shape", ...), or in which context Tweety is a typical bird", or "in which context John looks like Fred"....

3. Bottom-Up Inheritance

A class $C1$ can be defined or its description refined by pointing out a prototypical subclass $C2$, as suggested by Kayser²², assuming $C2$ is already present in the system or intensionally described by the user. Then by pointing out $C2$ as a typical subclass of $C1$, the description of $C1$ is partially refined from that of $C2$, namely the attribute ranges of $C2$ become the typical ranges of its superclass $C1$. For instance defining "birds" by means of "blackbirds" as being typical birds, we may conclude that typical birds fly if it were not known yet.

E. CONCLUSION

The proposed approach presents an object-centered representation completely based on possibility theory. The distinction between class and typical class allows both to deal with exceptions in inference processes and to perform classification easily. Moreover we handle three kinds of inheritance (typical, atypical and normal) for the definition of new objects. Various degrees can also be used to refine our knowledge about the relationships that objects can satisfy. Our paper has not considered the estimation of membership grades in the various fuzzy classes. When domains are infinite and are intervals in the real line, membership grades are often obtained by interpolation between extreme membership values, as with trapezoidal fuzzy intervals. When domains are finite and without any a priori order, membership grades could be derived by means of data analysis methods applied to samples of objects in a given context. Yager²³ also uses possibility theory in nonmonotonic inheritance systems; however his view of typicality is based on a representation of possibility-qualification and is close to the spirit of default logics, while our approach handles classes of typical objects in an explicit manner, as particular subclasses.

The object-centered representation we presented may be used in association with (fuzzy) production rules expressing relations between attribute values; it may also include information about cardinalities of classes. This may be of interest for

- 22 D. Kayser, "What kind of thing is a concept?," *Computational Intelligence*, **4**, 158–165 (1988).
- 23 R.R. Yager "Nonmonotonic inheritance systems," Tech. Report #MII-804, Iona College, Machine Intelligence Institute, New Rochelle, N.Y., 1988.

APPENDIX

Here we present four different inclusion degrees of the form $\text{Inf}_u \{ \mu_A(u) * \rightarrow \mu_C(u) \}$, and their meaning when they take the values 0 or 1. $\text{Co}(A)$ stands for the core of a fuzzy set A and $\text{Su}(A)$ for its support. The inclusion relationships that appear in Table 1 are characteristic properties for extreme values of the inclusion indices. For continuous membership functions and Gödel implication, the condition $\text{Su}(A) \supseteq \text{Su}(C)$ should be replaced by $\forall \alpha, A_\alpha \supseteq C_\alpha$ where $A_\alpha = \{u \mid \mu_A(u) \geq \alpha\}$.

Table A.I. Characteristic properties of 4 indices of inclusion

	$\mu_A * \rightarrow \mu_C$	$N(C \mid A) > 0$	$N(C \mid A) = 1$
Dienes	$\text{Max}(\mu_C(u), 1 - \mu_A(u))$	$\text{Su}(C) \supseteq \text{Co}(A)$	$\text{Co}(C) \supseteq \text{Su}(A)$
Lukasiewicz	$\text{Min}(1, 1 - \mu_A(u) + \mu_C(u))$	$\text{Su}(C) \supseteq \text{Co}(A)$	$C \supseteq A$
Gödel	1 if $\mu_A(u) \leq \mu_C(u)$ $\mu_C(u)$ if $\mu_A(u) > \mu_C(u)$	$\text{Su}(C) \supseteq \text{Su}(A)$	$C \supseteq A$
Gödel's reciprocal	1 if $\mu_A(u) \leq \mu_C(u)$ $1 - \mu_A(u)$ if $\mu_A(u) > \mu_C(u)$	$\text{Co}(C) \supseteq \text{Co}(A)$	$C \supseteq A$

Proof of the decomposability result.

$$N_G(C \times D \mid A \times B) = \text{Min}(N_G(C \mid A), N_G(D \mid B)) \quad \text{when } \exists(s,t) / \mu_{A \times B}(s,t) = 1$$

$$\text{with } \mu_{A \times B}(s,t) = \text{min}(\mu_A(s), \mu_B(t)) \quad \mu_{C \times D}(s,t) = \text{min}(\mu_C(s), \mu_D(t))$$

$$N_G(C \mid A) = \text{inf}_s (\mu_A(s) * \rightarrow \mu_C(s)),$$

$$\text{and } a * \rightarrow c = \begin{cases} 1 & \text{if } c \geq a \\ c & \text{if } c < a \end{cases}$$

Let $\Phi(s,t)$ and $\Psi(s,t)$ and $A - C$ be defined by

$$\Phi(s,t) = \text{min}(\mu_A(s), \mu_B(t)) * \rightarrow \text{min}(\mu_C(s), \mu_D(t))$$

supporting or discounting plausible inferences. These are topics for further research.

Lastly we may contemplate weighting the attributes describing a class, in order to distinguish between typical and less typical attributes, as done by Torasso and Console⁹.

References

- 1 M. Minsky, "A framework for representing Knowledge," In *The Psychology of Computer Vision*, P.H. Winston (Ed.), McGraw-Hill, New York, 1975, pp. 211–277.
- 2 A. Goldberg and D. Robson, *Smalltalk-80: the Language and its Implementation*, Addison Wesley, Reading, MA, 1983.
- 3 A. Collins and R. Michalski, "The logic of plausible reasoning: a core theory," *Cognitive Science*, **13**, 1–49 (1989).
- 4 E. A. Rundensteiner and W. Bandler, "The equivalence of knowledge representation schemata: 'semantic networks' and 'fuzzy relational products'," *Proc. of NAFIPS (North-Amer. Fuzzy Inform. Proces. Soc.) Conf.*, New Orleans, 1986, pp. 477–501.
- 5 P. Vignard, *Un mécanisme d'exploitation à base de filtrage flou pour une représentation des connaissances centrées objets*, Thèse 3ème cycle, Institut National Polytechnique de Grenoble, France, 1985.
- 6 I. Graham and P. L. Jones, "A theory of fuzzy frames," *BUSEFAL* (LSI-IRIT, Université Paul Sabatier, Toulouse, France), n° 31, 109–132 and n° 32, 120–135, 1987.
- 7 I. Graham and P. L. Jones, *Expert systems: Knowledge, Uncertainty and Decision*, Chapman and Hall, 1988.
- 8 C. Granger, "An application of possibility theory to object recognition," *Fuzzy Sets and Systems*, **2** 8(3), 351–362 (1988).
- 9 P. Torasso and L. Console, "Approximate Reasoning and Prototypical knowledge," *Int. J. of Approximate Reasoning*, **3**(2), 157–177 (1989).
- 10 L.A. Zadeh, "Fuzzy sets," *Information and Control*, **8**, 338–353 (1965).
- 11 R.R. Yager, "Linguistic representations of default values in frames," *IEEE Trans. on Systems, Man and Cybernetics*, **14**(4), 630–633 (1984).
- 12 B. Cohen and G. Murphy, "Models of concepts," *Cognitive Science*, **8**, 27–58 (1984).
- 13 L. A. Zadeh, "A note on prototype theory," *Cognition*, **12**, 291–297 (1982).
- 14 W. Bandler and L. Kohout, "Fuzzy power sets and fuzzy implication operators," *Fuzzy Sets and Systems*, **4**, 13–30 (1980).
- 15 L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," *Information Sciences*, Part 1: **8**, 199–249; Part 2: **8**, 301–357; Part 3: **9**, 43–80 (1975).
- 16 M. Cayrol, H. Farreny and H. Prade, "Fuzzy pattern matching," *Kybernetes*, **11**, 103–116 (1982).
- 17 D. Dubois and H. Prade (with the collaboration of H. Farreny, R. Martin-Clouaire and C. Testemale) *Possibility Theory — An Approach to the Computerized Processing of Information*, Plenum Press, New-York, 1988.
- 18 J.-P. Rossazza, *Utilisation de hiérarchies de classes floues pour la représentation de connaissances imprécises et sujettes à exceptions: le système "SORCIER"*, Thèse 3ème cycle, Université Paul Sabatier, Toulouse, France, 1990.
- 19 P. Coupey, *Etude d'un réseau sémantique avec gestion des exceptions*, Thèse 3ème cycle, Université de Paris Nord, Villetaneuse, France, 1989.
- 20 S.E. Fahlman, *NETL: a system for representing and using real-world knowledge*, MIT Press, Cambridge, Ma., 1979
- 21 R. Reiter, "A Logic for Default Reasoning," *Artificial Intelligence*, **13**(1–2), 81–132 (1980).

$$\begin{aligned} \Psi(s,t) &= \min(\mu_A(s) * \rightarrow \mu_C(s), \mu_B(t) * \rightarrow \mu_D(t)) \\ A - C &= \{s / \mu_A(s) > \mu_C(s)\} \end{aligned}$$

Let us notice that

$$\min(N_G(C \mid A), N_G(D \mid B)) = \inf_{s,t} \{\Psi(s,t)\}.$$

Moreover

$$\begin{aligned} \Phi(s,t) &= \begin{cases} 1 & \text{if } (s,t) \notin (A \times B) - (C \times D) \\ \min(\mu_C(s), \mu_D(t)) & \text{if } (s,t) \in (A \times B) - (C \times D) \end{cases} \\ \Psi(s,t) &= \begin{cases} 1 & \text{if } s \notin A - C, t \notin B - D \\ \mu_C(s) & \text{if } s \in A - C, t \notin B - D \\ \mu_D(t) & \text{if } s \notin A - C, t \in B - D \\ \min(\mu_C(s), \mu_D(t)) & \text{if } s \in A - C, t \in B - D \end{cases} \end{aligned}$$

Clearly

$$\begin{aligned} \Phi(s,t) = \Psi(s,t) & \quad \text{if } s \in A - C, t \in B - D \\ \text{or } & \quad \text{if } s \notin A - C, t \notin B - D \end{aligned}$$

since $\mu_A(s) > \mu_C(s)$ and $\mu_B(t) > \mu_D(t)$ entails $\min(\mu_A(s), \mu_B(t)) > \min(\mu_C(s), \mu_D(t))$ and similarly changing ' $>$ ' into ' \leq '.

It is then simple to check that $\Psi(s,t) \leq \Phi(s,t)$ in the general case. Indeed

$$\begin{aligned} \text{if } & \mu_A(s) > \mu_C(s), \mu_B(t) \leq \mu_D(t) \quad \text{and} \quad \min(\mu_A(s), \mu_B(t)) > \min(\mu_C(s), \mu_D(t)) \\ \text{then } & \mu_C(s) \leq \mu_D(t) \quad \text{and so } \Phi(s,t) = \Psi(s,t) = \mu_C(s). \end{aligned}$$

The case when $\mu_A(s) \leq \mu_C(s)$ and $\mu_B(t) > \mu_D(t)$ is dealt with similarly.

In other words

$$\Phi(s,t) = \Psi(s,t) \text{ or } \Phi(s,t) = 1 \text{ (since } (A \times B) - (C \times D) \supseteq (A - C) \times (B - D)\text{)}.$$

To prove the result, it is enough to check that the minimum of Ψ is attained for at least one value where $\Phi = \Psi$.

The minimum of Ψ is attained on $(A - C) \times (B - D)$; indeed for instance

$$\begin{aligned} \inf_{s \in A - C, t \notin B - D} \Psi(s,t) &= N_G(C \mid A) \geq \inf_{(s,t) \in (A - C) \times (B - D)} \Psi(s,t) \\ &= \min(N_G(C \mid A), N_G(D \mid B)) \end{aligned}$$

This is true of course if $(A - C) \times (B - D) \neq \emptyset$ i.e. neither $\mu_A \leq \mu_C$ nor $\mu_B \leq \mu_D$ hold. In that case the decomposability result is obtained since $\Phi = \Psi$ on $(A - C) \times (B - D)$.

If $(A - C) \times (B - D) = \emptyset$ then $\mu_A \leq \mu_C$ or $\mu_B \leq \mu_D$. Let us assume $\mu_A \leq \mu_C$

only. The definition of $\Psi(s,t)$ simplifies since it is equal to $\mu_B(t) * \rightarrow \mu_D(t)$. The expression of $\Phi(s,t)$ boils down to

$$\Phi(s, t) = \begin{cases} 1 & \text{if } \min(\mu_A(s), \mu_B(t)) \leq \min(\mu_C(s), \mu_D(t)) \\ \mu_D(t) & \text{otherwise} \end{cases}$$

indeed $\min(\mu_A(s), \mu_B(t)) > \min(\mu_C(s), \mu_D(t))$ implies $\mu_B(t) > \mu_D(t)$ when $\mu_A(s) \leq \mu_C(s)$; moreover we also get

$$\mu_C(s) \geq \min(\mu_A(s), \mu_B(t)) > \min(\mu_C(s), \mu_D(t)) \text{ so that } \mu_C(s) > \mu_D(t).$$

To get the decomposability result, it is enough to find a minimum (s,t) of Ψ with $\Phi(s,t) = \Psi(s,t)$.

Let t' such that $\mu_B(t') * \rightarrow \mu_D(t') = \inf_t \{\mu_B(t) * \rightarrow \mu_D(t)\} = \mu_D(t')$.

Clearly $\mu_D(t') < 1$ since $\mu_B \leq \mu_D$ does not hold, and any (s,t') is a minimum of Ψ . Hence we are left to find (s,t') such that $\min(\mu_A(s), \mu_B(t')) > \mu_D(t')$ i.e. (due to $\mu_B(t') > \mu_D(t')$ find s such that $\mu_A(s) > \mu_D(t')$).

Now $\mu_D(t')$ can take any value smaller than 1. Hence only the normalization of A ensures that $\exists s', \mu_A(s') > \mu_D(t')$. In that case $\Phi(s',t') = \Psi(s',t') = \mu_D(t')$ and the decomposability holds.

Clearly if $\mu_B \leq \mu_D$ is assumed, we need the normalization of B .

The last case is when $\mu_B \leq \mu_D$ and $\mu_A \leq \mu_C$; then $\Phi = \Psi = 1$.

The requirement that A and B be normalized, needed for decomposability, is the extension of a nonemptiness condition in the nonfuzzy case, namely

$$C \times D \supseteq A \times B \Leftrightarrow C \supseteq A \text{ and } D \supseteq B$$

provided that $A \neq \emptyset$ and $B \neq \emptyset$ (or $A = B = \emptyset$).