# An ordinal approach to the processing of fuzzy queries with flexible quantifiers <br> Patrick Bosc, Ludovic Liétard, Henri Prade 

## To cite this version:

Patrick Bosc, Ludovic Liétard, Henri Prade. An ordinal approach to the processing of fuzzy queries with flexible quantifiers. Anthony Hunter; Simon Parsons. Applications of Uncertainty Formalisms, 1455, Springer, pp.58-87, 1998, Lecture Notes in Computer Science book series (LNCS), 978-3-540-65312-7. 10.1007/3-540-49426-X_4 . hal-04028802

## HAL Id: hal-04028802

## https://hal.science/hal-04028802

Submitted on 16 Mar 2023

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.

# An Ordinal Approach to the Processing of Fuzzy Queries with Flexible Quantifiers 

Patrick Bosc ${ }^{1}$, Ludovic Liétard ${ }^{1}$ and Henri Prade ${ }^{2}$<br>${ }^{1}$ IRISA/ENSSAT , Technopole Anticipa, BP 447, 22305 Lannion Cédex, France<br>e-mail: boscllietard@enssat.fr<br>${ }^{2}$ IRIT-CNRS, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cédex, France<br>e-mail: prade@irit.fr


#### Abstract

This paper studies queries to a database, involving expressions of the form ' Q A-x's are B's' where A and B are properties which may be fuzzy and with respect to which objects x's are evaluated, and where Q is a quantifier which may stand for 'all', or may leave room for exceptions ('at least $\mathrm{q} \%$ ', '(at least) most', etc.). An example of such a query is 'Find the departments where most young employees are well-paid'. Such queries are discussed from a modeling and evaluation point of view, taking also into consideration what the user intends to ask when (s)he addresses this type of queries to a database system. Clarifying what has to be evaluated is specially important in the case where A is fuzzy, since then the boundaries of A are ill-defined and A may be somewhat empty.


## 1 Introduction

One reason for introducing fuzzy sets in query models [2][3] is the representation of the preferences of the user which thus can be expressed in a simple way. This is clearly the case in a query asking, for instance, for 'an apartment not too expensive and not too far from downtown'. A benefit of the fuzzy set modeling is then to provide a framework for rank-ordering the answers according to their compatibility with the fuzzy request. However, this is not always the motivation underlying the use of fuzzy terms in a query. If for instance, we ask for 'the average of the salaries of the young people' about whom information is stored in the database, the use of a fuzzy label such as 'young' is rather a matter of convenience, which avoids to refer to a precise and somewhat arbitrary age threshold. In order that the query makes sense, it is necessary that the result does not vary too much with the slightly different possible interpretations of the word 'young' (in the given context), and if there is some brittleness of the result with respect to the different interpretations, we expect that the system will inform us about this state of fact, or will take it into account in its evaluation. A range of possible values for the average salary, according to the different possible interpretations of 'young', should then be returned by the system; see [10] for the treatment of such queries.

Queries involving fuzzy quantifiers, such as 'Find the departments where most young employees are well-paid', or 'What are the days where almost all early trains are overcrowded?', seem also to be motivated by some robustness issue. Indeed 'most'
(then understood as 'at least most') is often a way of expressing some implicit proviso for exceptions rather than a way of really specifying a proportion in a fuzzy way. It should be understood in a flexible way. Indeed, addressing queries involving the universal quantifier 'all' instead of 'most' to a database, such as asking for departments where all employees are well-paid, may very often lead to empty answers. As in the average salary example, the use of fuzzy categories like 'young' or 'well-paid' is also here a matter of convenience and implicitly presupposes that the result of the query does not vary too much with the possible interpretations of the words. In the following, a qualitative model is proposed for representing fuzzy expressions of the form 'Q A-x's are B's', and its use in the handling of queries is discussed.

Although this work could be related to different views of the cardinality of a fuzzy set and the modeling of fuzzily quantified statements $[8,12,14,16,20]$, the evaluation of 'Q A-x's are B's' with respect to an ordinary database (data are supposed to be precise and certain) is not envisaged here as the result of the matching of a count of the A's which are B's against $\mathrm{Q}^{1}$. Thus, the evaluation we are interested in, is rather viewed here as the extent to which all A's are B's up to some exceptions. Indeed we are primarily interested in rank-ordering cases (in our above example, 'the departments') according to the extent to which they satisfy a requirement of the form 'Q A-x's are B's'. Moreover, the approach only assumes the use of a purely ordinal scale (e.g., a finite, totally ordered, chain of levels) for defining the fuzzy sets, even if numbers in $[0,1]$ are used for convenience in the paper for encoding these levels.

Section 2 is devoted to a brief overview of basic notions related to fuzzy set theory involved in our evaluation problem. In Section 3, we assume that A is an ordinary subset of X, when evaluating expressions of the form 'Q A-x's are B's', before discussing the general case where A is a fuzzy set in Section 4. In each case, we first study the situation where Q is the universal quantifier ('for all'), before relaxing the evaluation with an exception-tolerant quantification.

In the following, $\mathrm{X}=\left\{\mathrm{x}_{1}, \ldots, \mathrm{x}_{\mathrm{n}}\right\}$ denotes a finite subset of objects, $a$ and $b$ are two attributes which apply to the elements of X, and A and B are two, fuzzy or not, subsets of the attribute domains of $a$ and $b$ respectively. For instance, X is a set of people, $a$ stands for 'age' and $b$ for 'salary', A for 'less than 30 years old', 'young', etc., and B for 'more than 10 k FF ', 'well-paid', etc.

## 2 Background

This section recalls different concepts associated to fuzzy set theory which will be used later on. Some fuzzy set notions are restated in Subsection 2.1, whereas Subsection 2.2 is devoted to fuzzy implications and fuzzy inclusions. The modeling of fuzzy quantifiers such as 'most' is discussed in Subsection 2.3.

[^0]
### 2.1 Fuzzy Sets

The concept of a fuzzy set, introduced by L.A. Zadeh [18], aims at extending the notion of a regular set in order to express classes with ill-defined boundaries (corresponding in particular to linguistic values, e.g., tall, young, well-paid, important, etc). This framework allows for a gradual transition between nonmembership and full membership. A degree of membership is associated to every element, $x$ and a fuzzy set $F$ over the referential $X$ (i.e., $F$ is a fuzzy subset of $X$ ) is defined by means of a membership function: $\mu_{F}$ from $X$ to $[0,1]$. For any $x$ in $X$, $\mu_{\mathrm{F}}(\mathrm{x})$ is the membership degree of x in F . It should be emphasized that the role of these degrees is first to rank-order the elements of the universe X , according to their compatibility with the fuzzy set F . Remember that we are using the interval $[0,1]$ here as an ordinal scale only; thus the ordering of the degrees is more important than their exact values.

Let $F$ be a fuzzy subset of $X$. The height of $F$ is denoted by $h(F)$ and is defined as the largest membership degree $\left(\mathrm{h}(\mathrm{F})=\max _{\mathrm{x} \in \mathrm{X}} \mu_{\mathrm{F}}(\mathrm{x})\right.$ ). When $\mathrm{h}(\mathrm{F})=1, \mathrm{~F}$ is said to be normalized. The $\alpha$-level cut of $F$ is the ordinary subset defined by $\left\{x, \mu_{F}(x) \geq \alpha\right\}$ and is denoted by $F_{\alpha}$. The support of $F$ is the ordinary subset defined by $\left\{x, \mu_{F}(x)>0\right\}$. The core of F is the ordinary subset $\left\{\mathrm{x}, \mu_{\mathrm{F}}(\mathrm{x})=1\right\}$ of elements which undisputedly belong to F .

Example 1. Let $X$ be a set of individuals $\{$ Angela, John, Mick, Peter, Mary\} and the fuzzy subset F of young people given by $\mu_{\mathrm{F}}($ Angela $)=\mu_{\mathrm{F}}($ Peter $)=1, \mu_{\mathrm{F}}(\mathrm{John})=$ $0.8, \mu_{\mathrm{F}}($ Mick $)=0.5$ and $\mu_{\mathrm{F}}($ Mary $)=0$. The intended meaning is that Peter and Angela are considered as young (completely) and form the core of the fuzzy set, while John is considered as rather young and Mick as somewhat young (whereas Mary is not at all young). F is normalized $(\mathrm{h}(\mathrm{F})=1$ ) and its support is the set \{Angela, John, Mick, Peter\}. The $\alpha$-level cut $\mathrm{F}_{0.8}$ is the set \{Angela, John, Peter\}

### 2.2 Fuzzy Implications and Inclusions

A fuzzy implication is an operator $(\rightarrow)$ defined from $[0,1] \times[0,1]$ to $[0,1]$ which must satisfy the characteristic properties [15]:

1) $\quad(a \rightarrow 1)=1$;
2) $\quad(0 \rightarrow a)=1$;
3) $(1 \rightarrow a)=a$;
4) if $\mathrm{b} \geq \mathrm{c},(\mathrm{a} \rightarrow \mathrm{b}) \geq(\mathrm{a} \rightarrow \mathrm{c})$ (increasing w.r.t. the second argument);
5) if $\mathrm{a} \leq \mathrm{c},(\mathrm{a} \rightarrow \mathrm{b}) \geq(\mathrm{c} \rightarrow \mathrm{b})$ (decreasing w.r.t. the first argument).

There are mainly two basic kinds of implication connectives: S-implications which are of the form 'not A or $\mathrm{B}^{\prime}$ (or equivalently 'not (A and not B)'), and R-
implications which are obtained by residuation, $\mathrm{a} \rightarrow \mathrm{b}=\sup \{\mathrm{t} \in[0,1], \mathrm{a} * \mathrm{t} \leq \mathrm{b}\}$ where $*$ is a conjunction operation. Since we restrict ourselves to ordinal scales in this paper, we should take $*=\mathrm{min}$, and then the S -implication is Dienes' implication $\mathrm{a} \rightarrow \mathrm{b}=\max (1-\mathrm{a}, \mathrm{b})$, while the R-implication is Gödel's implication $\mathrm{a} \rightarrow \mathrm{b}=1$ if a $\leq \mathrm{b}$ and $\mathrm{a} \rightarrow \mathrm{b}=\mathrm{b}$ if $\mathrm{a}>\mathrm{b}$. Note that $1-$ (.) denotes nothing more than the orderreversing function of the ordinal scale in case a non-numerical encoding would be used.

Given an implication connective $\rightarrow$, an inclusion index $d$ between fuzzy sets [1] is naturally defined under the form:

$$
\mathrm{d}(\mathrm{~A} \subseteq \mathrm{~B})=\min _{\mathrm{x}}\left(\mu_{\mathrm{A}}(\mathrm{x}) \rightarrow \mu_{\mathrm{B}}(\mathrm{x})\right)
$$

For Dienes' implication the above index gives:

$$
\mathrm{d}(\mathrm{~A} \subseteq \mathrm{~B})=\min _{\mathrm{x}} \max \left(1-\mu_{\mathrm{A}}(\mathrm{x}), \mu_{\mathrm{B}}(\mathrm{x})\right)
$$

and we have the characteristic property:

$$
d(A \subseteq B)=1 \Leftrightarrow \operatorname{support}(A) \subseteq \operatorname{core}(B)
$$

While using Gödel's implication, the inclusion index is such that:

$$
\begin{aligned}
\mathrm{d}(\mathrm{~A} \subseteq \mathrm{~B}) & =1 \Leftrightarrow \forall \mathrm{x}, \mu_{\mathrm{A}}(\mathrm{x}) \leq \mu_{\mathrm{B}}(\mathrm{x}) \\
& =\min _{\mathrm{x}: \mu_{\mathrm{A}}(\mathrm{x})>\mu_{\mathrm{B}}(\mathrm{x})} \mu_{\mathrm{B}}(\mathrm{x}) \text { otherwise } .
\end{aligned}
$$

### 2.3 The Quantifier 'Most'

Fuzzy quantifiers have been proposed by Zadeh [20] for modeling linguistic expressions such as 'most', 'a few', 'almost all', etc. Such linguistic quantifiers express an intermediate attitude between the classical, universal and existential, quantifiers. A relative linguistic quantifier Q , referring to an ill-known proportion (e.g., 'most'), is represented by a membership function $\mu_{\mathrm{Q}}$ from $[0,1]$ to $[0,1]$, such as $\mu_{\mathrm{Q}}(\mathrm{i})$ is the level of satisfaction of the statement ' Q conditions are satisfied' when i is the proportion of satisfied conditions. Figure 1 gives an interpretation for 'most'. This interpretation is, however, clearly context-dependent.

Example 2. In the case of Figure 1, the fuzzy statement 'most conditions are satisfied' is completely false if the proportion of satisfied conditions is less than $75 \%$ ( $\mu_{\text {most }}(\mathrm{i})=0$ if $\mathrm{i} \leq 0.75$ ). If the proportion equals $82.5 \%$, the truth degree of the statement is $0.5\left(\mu_{\text {most }}(0.825)=0.5\right)$. When $90 \%$ (or more) of the conditions are satisfied, 'most conditions are satisfied' is completely true. In addition, if 'most' applies to a number m of conditions, it is possible to define 'most' directly on the
number of satisfied conditions (i.e., on a subset $\{0,1,2, \ldots, m\}$ of integers). In this last case, 'most' is represented by a function $\mu_{\mathrm{Q}}$ defined by:

$$
\forall \mathrm{i} \in\{0,1,2, \ldots, \mathrm{~m}\}, \mu_{\mathrm{Q}}(\mathrm{i})=\mu_{\mathrm{most}}(\mathrm{i} / \mathrm{m}) .
$$

For example, if $m=12$, according to Figure 1 , we get $\mu_{Q}(0)=\mu_{Q}(1)=\ldots=\mu_{Q}(9)=0$, $\mu_{\mathrm{Q}}(10)=\mu_{\text {most }}(10 / 12)=5 / 9$ and $\mu_{\mathrm{Q}}(11)=\mu_{\mathrm{Q}}(12)=1$. The quantifier represented by $\mu_{\mathrm{Q}}$ can be then interpreted as 'at least about 11'


Fig. 1. A representation for the quantifier 'most'
In the following, we use a slightly different view for modeling 'most' in several examples. It is based on the idea of neglecting one or a few of the worst satisfied conditions (whereas the other conditions are fully taken into account). In that case, the quantifier is no longer genuinely gradual, and we are rather looking for the minimal relaxation of the requirement that 'all the conditions are satisfied' into 'all the conditions, except a few, are satisfied'.

## 3 Qualitative Modeling of 'Q A-x's are B's' with A Crisp

### 3.1 Case of the universal quantifier ( $\mathbf{Q}=$ 'all')

Let us first assume that A and B are not fuzzy, and Q is the universal quantifier 'all'. The index E which estimates to what extent it is true that ' $\mathrm{Q} A$-x's are B 's' is given by:

$$
\begin{align*}
\mathrm{E} & =\min _{\mathrm{x} \in \mathrm{~A}} \mu_{\mathrm{B}}(\mathrm{x}) \text { if } \mathrm{A} \neq \emptyset  \tag{1}\\
& =\left\{\begin{array}{l}
1 \text { if } \mathrm{A} \subseteq \mathrm{~B} \text { and } \mathrm{A} \neq \emptyset, \\
0 \text { otherwise. }
\end{array}\right.
\end{align*}
$$

Note that (1) still makes sense when B becomes fuzzy, since E is then all the greater as all the objects in A have a high degree of membership in B. In particular, we have:

$$
\begin{equation*}
\mathrm{E}=\max \left\{\beta \in(0,1] \text { s.t. } \mathrm{A} \neq \emptyset \text { and } \mathrm{A} \subseteq \mathrm{~B}_{\beta}\right\} \tag{2}
\end{equation*}
$$

$$
=0 \text { if } \mathrm{A} \not \subset\left\{\mathrm{x}, \mu_{\mathrm{B}}(\mathrm{x})>0\right\} \text { or } \mathrm{A}=\emptyset \text {, }
$$

where $B_{\beta}=\left\{\mathrm{x}, \mu_{\mathrm{B}}(\mathrm{x}) \geq \beta\right\}$ denotes the $\beta$-level cut of the fuzzy set B .
Example 3. Let $X$ be a set of employees described in Table 1, A (resp. B) be the predicate 'Age > 32' (resp. 'well-paid'). The evaluation of the query 'all employees over 32 are well-paid' yields:

$$
\begin{aligned}
\mathrm{E} & =\min _{\mathrm{e} \in \mathrm{EMP} \text { s.t. e.age }>32} \mu_{\text {well-paid }}(\mathrm{e}) \\
& =\min (1,0.2)=0.2
\end{aligned}
$$

EMP. | \#emp | name | age | salary | $\mu_{\text {well-paid }}$ |
| :---: | :---: | :---: | :---: | :---: |
| 189 | Smith | 56 | 60 | 1 |
| 57 | Jones | 29 | 80 | 1 |
| 876 | Kent | 43 | 38 | 0.2 |
| 217 | Allen | 24 | 45 | 0.5 |

Table 1

E as a scalar evaluation may appear to be too crude. It is possible to refine the ordering introduced by E by using a leximin ordering [5]. Namely we can rank-order the elements of A according to non-decreasing values of $\mu_{\mathrm{B}}$, i.e., $\mu_{\mathrm{B}}\left(\mathrm{x}_{\tau}(1)\right) \leq \ldots \leq$ $\mu_{\mathrm{B}}\left(\mathrm{x}_{\tau}(\mathrm{m})\right)$ where $\tau$ is a permutation and $|\mathrm{A}|=\mathrm{m}$. Then, if we have two evaluations E and $\mathrm{E}^{\prime}$ such that $\mathrm{E}=\mathrm{E}^{\prime}$, corresponding in one of the above examples to two different departments for instance, the two departments can be ordered according to the leximinordering between the two corresponding vectors $\left(\mu_{\mathrm{B}}\left(\mathrm{x}_{\tau}(\mathrm{i})\right)\right)_{\mathrm{i}}=1, \mathrm{~m}$ and $\left(\mu_{\mathrm{B}}\left(\mathrm{x}_{\tau^{\prime}}^{\prime}(\mathrm{i})\right)\right)_{\mathrm{i}=1, \mathrm{~m}}$; i.e., $\mathrm{E}>\mathrm{E}^{\prime}$ if and only if $\exists \mathrm{k}, \mu_{\mathrm{B}}\left(\mathrm{x}_{\tau}(\mathrm{t})\right)=\mu_{\mathrm{B}}\left(\mathrm{x}_{\tau^{\prime}}^{\prime}(\mathrm{t})\right)$ for $\mathrm{t}=1, \mathrm{k}$ and $\mu_{\mathrm{B}}\left(\mathrm{x}_{\tau}(\mathrm{k}+1)\right)>\mu_{\mathrm{B}}\left(\mathrm{x}^{\prime} \tau^{\prime}(\mathrm{k}+1)\right)$ for $\mathrm{k}+1 \leq \mathrm{m}$. For instance, if $\mathrm{E}=\min (0.2,1,1,1)$, $\mathrm{E}^{\prime}=\min (0.2,0.2,0.2,0.2)$, we will have $\mathrm{E}>\mathrm{E}^{\prime}$. This type of refinement could be also extended to the evaluation of E in the more complex cases considered below. However, this would not be discussed further for the sake of brevity.

### 3.2 Relaxing 'all' into 'most' $(\mathbf{Q}=$ 'most')

The evaluation of 'all A-x's are B's', as expressed by (1) is the min-aggregation (which, indeed, considers the universal quantifier a generalized conjunction) of the degrees of membership to $B$ of all the $x$ 's in $A$. Going from 'all' to 'most', the idea is to relax the requirement into 'in most cases, the x's which are A's are B's', or equivalently 'there are only a few cases of A's which are not B's, or are 'bad' B's'. This can be done by weighting the min-aggregation and by giving little importance to these few cases.

The quantifier 'most' is then represented by a subset of the set of integers $\{0,1$, $\ldots, m\}$ where $m$ is the cardinal of the set to which the quantifier refers (here A) rather than by a subset of proportions. More precisely, 'most' (understood as '(at least) most') has a non-decreasing membership function, i.e., $\mu_{\text {most }}(\mathrm{k}) \leq \mu_{\text {most }}(\mathrm{k}+1)$ for $\mathrm{k}=$ $0, \mathrm{~m}-1$, and such that $\mu_{\text {most }}(\mathrm{m})=1 . \mathrm{Q}=$ 'all' is represented by $\mu_{\mathrm{Q}}(\mathrm{k})=0, \forall \mathrm{k} \leq \mathrm{m}-$ 1 in this setting. In particular, 'at least $\mathrm{k}^{\prime}$ will be represented by $\mu_{\mathrm{Q}}(\mathrm{t})=0$ if $\mathrm{t}=0, \mathrm{k}-$ 1 and $\mu_{\mathrm{Q}}(\mathrm{t})=1$ if $\mathrm{t}=\mathrm{k}, \mathrm{m}$. Associated with 'most' is the (possibly fuzzy) subset I defined by $\mu_{\mathrm{I}}(\mathrm{k})=1-\mu_{\text {most }}(\mathrm{k}-1)$ for $\mathrm{k}=1, \mathrm{~m}$ and $\mu_{\mathrm{I}}(0)=1$. Thus if most $=\{\mathrm{k}, \ldots$, $m\}$, then $I=\{0, \ldots, k\}$. I represents the set of ranks of elements to be considered as important in the evaluation as we shall see in the following.

In this section, we consider the case where $A$ is not fuzzy and $A \neq \emptyset$ (if $A=\emptyset, E$ $=0$ ). Let $\mathrm{A}=\left\{\mathrm{x}_{1}, \ldots, \mathrm{x}_{\mathrm{m}}\right\}, \mathrm{m} \leq \mathrm{n}=|X|$. Assume that the elements of A are reordered according to the decreasing values of $\mu_{\mathrm{B}}$, i.e., $\mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma(1)}\right) \geq \mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma}(2)\right) \geq \ldots \geq \mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma}(\mathrm{m})\right)$ where $\sigma$ is a permutation of $\{1, \ldots, m\}$. Then the estimation $E$ is expressed by:

$$
\begin{equation*}
\mathrm{E}=\min _{\mathrm{i}=1, \mathrm{~m}} \max \left(\mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma}(\mathrm{i})\right), 1-\mu_{\mathrm{I}}(\mathrm{i})\right) \tag{3}
\end{equation*}
$$

(3) can be rewritten in order to let A explicitly appear. Namely:

$$
\begin{equation*}
\mathrm{E}=\min _{\mathrm{j}=1, \mathrm{n}} \max \left(\mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma^{\prime}}(\mathrm{j})\right), 1-\mu_{\mathrm{A}}\left(\mathrm{x}_{\mathrm{\sigma}^{\prime}(\mathrm{j})}\right), 1-\mu_{\mathrm{I}}(\mathrm{j})\right) \tag{4}
\end{equation*}
$$

where $\sigma^{\prime}(\mathrm{j})=\sigma(\mathrm{j})$ for $1 \leq \mathrm{j} \leq \mathrm{m}$ and $\sigma^{\prime}=$ identity for $\{\mathrm{m}+1, \ldots, \mathrm{n}\}$, since $\mu_{\mathrm{A}}\left(\mathrm{x}_{\sigma^{\prime}}(\mathrm{j})\right)$ $=0$, for $\mathrm{j}=\mathrm{m}+1, \mathrm{n}\left(\mu_{\mathrm{I}}(\mathrm{j})=0\right.$ for $\mathrm{j}=\mathrm{m}+1, \mathrm{n}$ for the sake of coherence $)$. It can be easily checked that (1) is a particular case of (3)-(4), letting $I=\{0,1, \ldots, m\}$, and $E$ defined by (3) is always greater or equal to E given by (1). (4) expresses the extent to which the k objects of A which are the more important with respect to their membership in B are indeed elements of $B$ with high membership. (3)-(4) is an example of an 'ordered weighted minimum' aggregation, and has been recently proposed in the context of the division of fuzzy relations [11].

As it can be seen, E has a high value as soon as the elements $\mathrm{x}_{\sigma(i)} \in \mathrm{A}$ which have low membership grades in B are neglected, i.e., are such that $\mu_{\mathrm{I}}(\mathrm{i})$ is close to 0 . In the particular case where $|\mathrm{A}|=\mathrm{m}=1$, we get $\mathrm{E}=0$ if $\mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma}(1)\right)=0$ since in this case $\mu_{\mathrm{I}}(1)=1-\mu_{\text {most }}(0)=1$ with $\mu_{\text {most }}(0)=0$ (if we want to keep at least one nonneglected element in $\mathrm{A}!$ ). When $\mathrm{A}=\mathrm{X}$, (3)-(4) yields the degree of membership in B of the $\mathrm{k}^{\text {th }}$ best element in X according to $\mu_{\mathrm{B}}$. In practice, quantifiers of the form 'at least $\mathrm{k}^{\prime}$ may be sufficient for our purpose, although (3)-(4) still make sense if 'most' and thus I becomes fuzzy (with monotonic, non-decreasing and non-increasing membership functions respectively, as already said).

Example 4. Let us consider 'most A- X 's are B's' applying to the set $\mathrm{X}=\left\{\mathrm{x}_{1}, \mathrm{x}_{2}, \mathrm{x}_{3}\right.$, $\left.x_{4}, x_{5}\right\}$ and $A=\left\{x_{1}, x_{2}, x_{3}, x_{4}\right\}$. The degrees of membership to $A$ and $B$ are given in Table 2. The quantifier 'most' is defined on $\{0,1,2,3,4\}$, since $m=4$, by:

$$
\mu_{\mathrm{most}}(0)=0, \mu_{\mathrm{most}}(1)=0, \mu_{\mathrm{most}}(2)=0.1, \mu_{\mathrm{most}}(3)=0.8, \mu_{\mathrm{most}}(4)=1
$$

Then, $\mu_{\mathrm{I}}(0)=1, \mu_{\mathrm{I}}(1)=1, \mu_{\mathrm{I}}(2)=1, \mu_{\mathrm{I}}(3)=0.9, \mu_{\mathrm{I}}(4)=0.2$. According to formula (3):

$$
\begin{array}{r}
\mathrm{E}=\min \left(\max \left(\mu_{\mathrm{B}}\left(\mathrm{x}_{1}\right), 1-\mu_{\mathrm{I}}(1)\right), \max \left(\mu_{\mathrm{B}}\left(\mathrm{x}_{2}\right), 1-\mu_{\mathrm{I}}(2)\right),\right. \\
\left.\max \left(\mu_{\mathrm{B}}\left(\mathrm{x}_{3}\right), 1-\mu_{\mathrm{I}}(3)\right), \max \left(\mu_{\mathrm{B}}\left(\mathrm{x}_{4}\right), 1-\mu_{\mathrm{I}}(4)\right)\right) .
\end{array}
$$

We finally get:

$$
\mathrm{E}=\min (\max (1,0), \max (0.8,0), \max (0.6,0.1), \max (0.2,0.8))=0.6
$$

This result is not surprising since $\mathrm{x}_{4}$, the worst B-element of A, (which would have given $E=0.2$ for $Q=$ 'all') is somewhat neglected since this last element is not completely important. Note that making $\mu_{\text {most }}(2)=0$ and $\mu_{\text {most }}(3)=1$, i.e., making Q $=$ 'at least 3 ' would not change the result of the evaluation $\downarrow$

| X | A | B |
| :---: | :---: | :---: |
| $\mathrm{x}_{1}$ | 1 | 1 |
| $\mathrm{x}_{2}$ | 1 | 0.8 |
| $\mathrm{x}_{3}$ | 1 | 0.6 |
| $\mathrm{x}_{4}$ | 1 | 0.2 |
| $\mathrm{x}_{5}$ | 0 | 0.7 |

Table 2
However, as already said, (3)-(4) evaluates the extent to which 'in most cases, the x's which are A's are also B's', and not to what extent 'most x's are A's and B's'. Indeed in Example 4, 'all x which are A's (except one) are B 's' yields $\mathrm{E}=0.6$, while 'all x (except one) are A's and B's' would yield an evaluation equal to 0.2 (neglecting $\mathrm{x}_{5}$ ).

## 4 Qualitative Modeling of 'Q A-x's are B's' with A Fuzzy

### 4.1 Case of the universal quantifier ( $\mathbf{Q}=$ 'all')

### 4.1.1 A is a normalized fuzzy set

Let us now consider the general situation where A is also a fuzzy set. We again start with the case $\mathrm{Q}=$ 'all', before relaxing it. Then it seems natural to require that E is all the greater as, whatever the crisp interpretations of A and B in terms of level-cuts $\mathrm{A}_{\alpha}$ and $\mathrm{B}_{\beta}$, the condition $\mathrm{A}_{\alpha} \subseteq \mathrm{B}_{\beta}$ holds. We first consider the case where A is normalized, i.e., $\forall \alpha, A_{\alpha} \neq \emptyset$. This leads to state that:

$$
\mathrm{E}=\left\{\begin{array}{l}
0 \text { if }\left\{\mathrm{x}, \mu_{\mathrm{A}}(\mathrm{x})=1\right\} \not \subset\left\{\mathrm{x}, \mu_{\mathrm{B}}(\mathrm{x})>0\right\}  \tag{5}\\
\max \left\{\min (1-\alpha, \beta),(\alpha, \beta) \in(0,1)^{2}\right. \\
\left\{1 \text { if }\left\{\mathrm{x}, \mu_{\mathrm{A}}(\mathrm{x})>0\right\} \subseteq\left\{\mathrm{x}, \mu_{\mathrm{B}}(\mathrm{x})=1\right\}\right.
\end{array}\right.
$$

Indeed, since the $\lambda$-cut $F_{\lambda}$ of a fuzzy set $F$ is all the larger as $\lambda$ is small, the statement 'all A-x's are $\mathrm{B}^{\prime} \mathrm{s}$ ' is all the more true as $\mathrm{A}_{\alpha} \subseteq \mathrm{B}_{\beta}$ holds for small $\alpha$ and large $\beta$, i.e., $\mathrm{A}_{\alpha}$ is large and $\mathrm{B}_{\beta}$ is small. Indeed when the support of $\mathrm{A},\left\{\mathrm{x}, \mu_{\mathrm{A}}(\mathrm{x})>0\right\}$, is included in the core of $\mathrm{B},\left\{\mathrm{x}, \mu_{\mathrm{B}}(\mathrm{x})=1\right\}$, we are completely certain that whatever the non fuzzy interpretations of A and B , we have $\mathrm{A}_{\alpha} \subseteq \operatorname{support}(\mathrm{A}) \subseteq \operatorname{core}(\mathrm{B}) \subseteq \mathrm{B}_{\beta}$, i.e., $A_{\alpha} \subseteq B_{\beta}$ holds for all $\alpha$ and $\beta$. In [13], it has been established that $E$ is nothing but the necessity ${ }^{2}$ of the fuzzy event B based on the possibility distribution $\mu_{\mathrm{A}}$, i.e., we have the equality:

$$
\begin{equation*}
\mathrm{E}=\min _{\mathrm{x} \in \mathrm{X}} \max \left(\mu_{\mathrm{B}}(\mathrm{x}), 1-\mu_{\mathrm{A}}(\mathrm{x})\right) . \tag{6}
\end{equation*}
$$

When A is non-fuzzy we recover (1). (6) can be also understood as a minaggregation weighted in terms of levels of importance [9]. Namely, it is all the more important to take into account x in the evaluation as $\mu_{\mathrm{A}}(\mathrm{x})$ is large, i.e., as x is indeed a typical element of A . In particular, if $\mu_{\mathrm{A}}(\mathrm{x})=0, \mathrm{E}$ does not depend on $\mu_{\mathrm{B}}(\mathrm{x})$; if $\mu_{\mathrm{A}}(\mathrm{x})=1$, we should have $\mathrm{E} \leq \mu_{\mathrm{B}}(\mathrm{x})$, while if $0<\mu_{\mathrm{A}}(\mathrm{x})<1$, E cannot be made equal to 0 just because we would have $\mu_{\mathrm{B}}(\mathrm{x})=0$ (in that case E will be just upper bounded by $1-\mu_{\mathrm{A}}(\mathrm{x})$ which reflects how much x is unimportant).

Note also that if $\mathrm{E}=\theta>0, \forall \mathrm{x} \in\left\{\mathrm{x}, \mu_{\mathrm{A}}(\mathrm{x})>1-\theta\right\}$ then $\mu_{\mathrm{B}}(\mathrm{x}) \geq \theta$, i.e., $\mathrm{A}_{\overline{1-\theta}} \subseteq$ $\mathrm{B}_{\theta}$, where $\mathrm{A}_{\alpha}^{-}$denotes the strict $\alpha$-cut of the fuzzy set $\mathrm{A}\left(\mathrm{A}_{\alpha}^{-}=\left\{\mathrm{x}, \mu_{\mathrm{A}}(\mathrm{x})>\alpha\right\}\right)$. Moreover, if $\mathrm{E}=\theta>0, \exists \mathrm{x} \in \mathrm{X},\left(\mu_{\mathrm{B}}(\mathrm{x})=\theta\right.$ and $\left.\mu_{\mathrm{A}}(\mathrm{x})>1-\theta\right)$ or $\left(\mu_{\mathrm{B}}(\mathrm{x}) \leq \theta\right.$ and $\mu_{\mathrm{A}}(\mathrm{x})=1-\theta$ ). In particular, $\mathrm{E}=0$ if and only if $\exists \mathrm{x} \in \mathrm{X}$ such that $\mu_{\mathrm{A}}(\mathrm{x})=1$ and $\mu_{\mathrm{B}}(\mathrm{x})=0$, i.e., if and only if there is an unchallenged element of A which does not belong to B at all, which is satisfying.

Generally speaking, E can be viewed as an inclusion index [1] of the form:

$$
\begin{equation*}
\mathrm{E}=\min _{\mathrm{x} \in \mathrm{X}} \mu_{\mathrm{A}}(\mathrm{x}) \rightarrow \mu_{\mathrm{B}}(\mathrm{x}) \tag{7}
\end{equation*}
$$

where $\mathrm{a} \rightarrow \mathrm{b}=\max (1-\mathrm{a}, \mathrm{b})$ is known as Dienes implication in the fuzzy set literature, as recalled in Section 2.2. One may then wonder about the usefulness of other implications and their meaning in terms of conditions over $\alpha$-cuts. If RescherGaines implication is chosen ( $\mathrm{a} \rightarrow \mathrm{b}=1$ if $\mathrm{a} \leq \mathrm{b}, \mathrm{a} \rightarrow \mathrm{b}=0$ otherwise), $\mathrm{E}=1$ iff $\mu_{\mathrm{B}}(\mathrm{x}) \geq \mu_{\mathrm{A}}(\mathrm{x})$ for all x , which means that for any $\alpha, \mathrm{A}_{\alpha} \subseteq \mathrm{B}_{\alpha}$. Thus, (7) with

2 Given a possibility distribution $\pi$ from $X$ to $[0,1]$ such that $\pi$ is normalized $(\exists x$, $\pi(x)=1)$, the necessity of a fuzzy event $B$ is defined by [6], $N(B)=\min _{x \in X}$ $\max \left(\mu_{\mathrm{B}}(\mathrm{x}), 1=\pi(\mathrm{x})\right)$ which is dually associated with a possibility measure [19], $\prod(\mathrm{B})=1-\mathrm{N}(\mathrm{B})=\max _{\mathrm{x}} \in \mathrm{X} \min \left(\mu_{\mathrm{B}}(\mathrm{x}), \pi(\mathrm{x})\right)$. Note that $\Pi(\mathrm{B} \cup \mathrm{C})=\max \left(\prod(\mathrm{B})\right.$, $\left.\prod(C)\right)$, while $N(B \cap C)=\min (N(B), N(C))$.

Rescher-Gaines implication estimates to what extent 'all x in X are at least as much B as they are $\mathrm{A}^{\prime}$; in particular if $\mu_{\mathrm{A}}(\mathrm{x})=0$, x has no influence on the evaluation.

Let us now consider Gödel implication: $\mathrm{a} \rightarrow \mathrm{b}=1$ if $\mathrm{a} \leq \mathrm{b}, \mathrm{a} \rightarrow \mathrm{b}=\mathrm{b}$ if $\mathrm{a}>\mathrm{b}$. In the context of (7), this implication leads to look for the smallest $\mu_{B}(x)$ such that $\mu_{\mathrm{B}}(\mathrm{x})<\mu_{\mathrm{A}}(\mathrm{x})$. Then it can be checked that

$$
\begin{align*}
\mathrm{E} & =\min \left\{\alpha \in[0,1), \mathrm{A}_{\alpha}^{-} \not \subset \mathrm{B}_{\alpha}^{-} \text {and } \mathrm{A}_{\alpha}^{-} \neq \mathrm{B}_{\alpha}^{-}\right\}  \tag{8}\\
& =1 \text { if } \forall \alpha, \mathrm{A}_{\alpha}^{-} \subseteq \mathrm{B}_{\alpha}^{-} .
\end{align*}
$$

Proof. $\left\{\alpha \in[0,1), \mathrm{A}_{\alpha}^{-} \not \subset \mathrm{B}_{\bar{\alpha}}^{-}\right.$and $\left.\mathrm{A}_{\alpha}^{-} \neq \mathrm{B}_{\bar{\alpha}}^{-}\right\}=\left\{\alpha \in[0,1), \mathrm{A}_{\alpha}^{-} \cap \mathrm{B}_{\bar{\alpha}}^{-} \neq \emptyset\right\}=\{\alpha \in$ $[0,1), \exists \mathrm{x}, \mu_{\mathrm{A}}(\mathrm{x})>\alpha$ and $\left.\mu_{\mathrm{B}}(\mathrm{x}) \leq \alpha\right\}$. Thus $\mathrm{E}=\min _{\mathrm{X}}\left\{\mu_{\mathrm{B}}(\mathrm{x}) \mid \mu_{\mathrm{B}}(\mathrm{x})<\mu_{\mathrm{A}}(\mathrm{x})\right\}$ and $\mathrm{E}=$ 1 if there is no x such that $\mu_{\mathrm{B}}(\mathrm{x})<\mu_{\mathrm{A}}(\mathrm{x})$

Thus E defined with Gödel implication refines the use of Rescher-Gaines implication, since in both cases $\mathrm{E}=1$ iff $\forall \alpha, \mathrm{A}_{\alpha} \subseteq \mathrm{B}_{\alpha}$, and E takes values intermediary between 0 and 1 if Gödel implication is used. These two implications express some simple conditions about the $\alpha$-cuts of A and B. This is particularly interesting, noticing that the inclusion of the $\alpha$-cuts of two fuzzy sets $A$ and $B$ is not monotonic with respect to $\alpha$. Indeed, we may have $\mathrm{A}_{\alpha_{1}} \subseteq \mathrm{~B}_{\alpha_{1}}, \mathrm{~A}_{\alpha_{2}} \not \subset \mathrm{~B}_{\alpha_{2}}$ and $\mathrm{A}_{\alpha_{3}} \subseteq$ $\mathrm{B}_{\alpha_{3}}$ with $\alpha_{1} \geq \alpha_{2} \geq \alpha_{3}$, as shown by the example $\mathrm{A}=\left\{1 / \mathrm{x}_{1}, 0.9 / \mathrm{x}_{2}, 0.5 / \mathrm{x}_{3}\right\}$, $\mathrm{B}=$ $\left\{0.8 / \mathrm{x}_{1}, 0.6 / \mathrm{x}_{2}, 0.2 / \mathrm{x}_{3}\right\}$ where $\mathrm{A}_{1} \not \subset \mathrm{~B}_{1}, \mathrm{~A}_{0.6} \subseteq \mathrm{~B}_{0.6}, \mathrm{~A}_{0.5} \not \subset \mathrm{~B}_{0.5}$, but $\mathrm{A}_{0.2} \subseteq \mathrm{~B}_{0.2}$. In this example $\mathrm{E}=0.2$ for Gödel implication since $\mathrm{A}_{\overline{0.2}} \not \subset \mathrm{~B}_{\overline{0.2}}$ and $\mathrm{A}_{\overline{0.2}} \neq \mathrm{B}_{\overline{0.2}}$.

### 4.1.2 A is an unnormalized fuzzy set

Let us now consider the case where $A$ is not normalized $(\mathrm{h}(\mathrm{A})<1$, where $\mathrm{h}(\mathrm{A})=$ $\max _{\mathrm{x} \in \mathrm{X}} \mu_{\mathrm{A}}(\mathrm{x})$ denotes the height of the fuzzy set A$)$. If we continue to use (6) in such a case, it can be easily checked that we would have $E \geq 1-h(A)$. In particular if $A=$ $\emptyset$, we get $E=1$, which is not satisfying, since $E$ should also reflect to what extent it makes sense to speak of the elements of A. Then E can be defined as:

$$
\begin{equation*}
\mathrm{E}=\min \left(\mathrm{h}(\mathrm{~A}), \min _{\mathrm{x} \in \mathrm{X}} \max \left(\mu_{\mathrm{B}}(\mathrm{x}), 1-\mu_{\mathrm{A}} *(\mathrm{x})\right)\right. \tag{9}
\end{equation*}
$$

where $\mu_{\mathrm{A}} *(\mathrm{x})=1$ if $\mu_{\mathrm{A}}(\mathrm{x})=\mathrm{h}(\mathrm{A})$ and $\mu_{\mathrm{A}} *(\mathrm{x})=\mu_{\mathrm{A}}(\mathrm{x})$ otherwise. Note that we recover (6) when $\mathrm{h}(\mathrm{A})=1$. E is a graded version of the condition, $\mathrm{A} \neq \varnothing$ and $\mathrm{A} \subseteq \mathrm{B}$, in the non-fuzzy case. A is renormalized in $A^{*}$ in (9) in order to have a meaningful degree of inclusion in (9). Indeed, if we keep A instead of $A^{*}$, we would have $E \geq \min (h(A), 1$ $-h(A)$ ) which is a lower bound which does not depend on $B$, which is not satisfying. The method proposed for normalizing $A$ is in agreement with the idea of a finite scale where the quotient $\mu_{\mathrm{A}}(\mathrm{x}) / \mathrm{h}(\mathrm{A})$ does not make sense. Moreover, it leaves a maximum number of membership grades untouched. The renormalization involved in expression (9) is illustrated hereafter.

| X 1 | A | B |
| :---: | :---: | :---: |
| $\mathrm{x}_{1}$ | 0.7 | 0.6 |
| $\mathrm{x}_{2}$ | 0.8 | 0.5 |
| $\mathrm{x}_{3}$ | 0.6 | 0 |

Table 3a

| X 2 | A | B |
| :---: | :---: | :---: |
| $\mathrm{x}_{1}$ | 0.2 | 0.8 |
| $\mathrm{x}_{2}$ | 0.5 | 0 |
| $\mathrm{x}_{3}$ | 0.3 | 0.2 |

Table 3b

Example 5. Let us consider two sets X 1 and X 2 as described in Tables 3a and 3b. The evaluation of the query 'all A-x's are B's' yields:

$$
\begin{aligned}
& \mathrm{E}(\mathrm{X} 1)=\min (0.8, \max (0.6,0.3), \max (0.5,0), \max (0,0.4))=0.4 \\
& \mathrm{E}(\mathrm{X} 2)=\min (0.5, \max (0.8,0.8), \max (0,0), \max (0.2,0.7))=0 .
\end{aligned}
$$

All the elements of X 1 with the higher degrees of membership in A have a non-zero degree of membership in $B$ and $E$ is thus non-zero. On the contrary, the element $x_{2}$ of X 2 with the highest degree $\mu_{\mathrm{A}}\left(\mathrm{x}_{2}\right)=\mathrm{h}(\mathrm{A})$ is such that $\mu_{\mathrm{B}}\left(\mathrm{x}_{2}\right)=0$, and $\mathrm{E}=0$, which is natural since $\mathrm{x}_{2}$ is considered to be in A as soon as we sufficiently relax our idea of A (into $\mathrm{A}_{\alpha}$ with $\alpha \leq 0.5$ ) in order to have $\mathrm{A} \neq \emptyset$ (and $\mathrm{x}_{2}$ is anyway the best representative of A in X 2$)$

### 4.2 Relaxing 'all' into 'most' $(\mathbf{Q}=$ 'most' $)$

More generally, we want to evaluate statements of the form Q A-x's are B's where 'all' is relaxed into a quantifier allowing for some exceptions. Three different types of evaluations can be distinguished (this is already the case when A is non-fuzzy). The first case corresponds to a relaxation of 'all A-x's are B's' into 'in most cases, if $x$ is an A it is a B also'. Such an evaluation is clearly different of the statement 'in most cases, the x's are A's and B's', which correspond to a second type. Indeed, in the first case there is no restriction on the x 's which are not A , while there are at most a few x's of this kind in the second case. In these two evaluations, the referential remains X . This was not the case in the evaluation (3), proposed in Section 3, where we were focusing on A only, ignoring what was happening in $\mathrm{X}-\mathrm{A}$. When A becomes fuzzy, this third type of evaluation becomes more tricky, since there are several crisp representatives of A depending on the $\alpha$-level cuts we consider thus leading to different evaluations in general. We now briefly discuss the three evaluations, starting with the two first ones, which were not developed for A crisp (although they apply as well to this case), in order to differentiate them from the third type of evaluation.

### 4.2.1 Interpretation 1: 'For most $x$ 's, those which are $A$ 's are $B$ 's'

The idea is then to give no importance to the worst counterexamples to the statement A's are B 's, i.e., the x 's which maximize $\min \left(\mu_{\mathrm{A}}(\mathrm{x}), 1-\mu_{\mathrm{B}}(\mathrm{x})\right)$. Thus, we rank-order the x 's in X according to the decreasing values of $1-\min \left(\mu_{\mathrm{A}}(\mathrm{x}), 1-\mu_{\mathrm{B}}(\mathrm{x})\right)=$
$\max \left(\mu_{\mathrm{B}}(\mathrm{x}), 1-\mu_{\mathrm{A}}(\mathrm{x})\right)$ and we give no importance to the last $\mathrm{n}-\mathrm{k}$ x's (if $\mathrm{Q}=$ 'at least $\mathrm{k}^{\prime}$ with $\mathrm{n}-\mathrm{k}$ much smaller than n in practice). The evaluation E is obtained by applying a formula close to (3) where $\mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma}(\mathrm{i})\right)$ is changed into $\mu_{\mathrm{A} \cup \mathrm{B}}\left(\mathrm{x}_{\sigma}(\mathrm{i})\right.$ ), and m is changed into n which refers to the cardinality of X :

$$
\begin{equation*}
\mathrm{E}=\min _{\mathrm{i}=1, \mathrm{n}} \max \left(\mu_{\mathrm{B}}\left(\mathrm{x}_{\sigma(\mathrm{i})}\right), 1-\mu_{\mathrm{A}}\left(\mathrm{x}_{\sigma(\mathrm{i})}\right), 1-\mu_{\mathrm{I}}(\mathrm{i})\right) \tag{10}
\end{equation*}
$$

where the values $\mu_{\overline{\mathrm{A}} \cup \mathrm{B}}\left(\mathrm{x}_{\sigma(\mathrm{i}}\right)$ are such that: $\mu_{\overline{\mathrm{A}} \cup \mathrm{B}}\left(\mathrm{x}_{\sigma(1)}\right) \geq \ldots \geq \mu_{\overline{\mathrm{A}} \cup \mathrm{B}}\left(\mathrm{x}_{\sigma}(\mathrm{n})\right)$ and I is the set of ranks of elements somewhat important $\left(\mu_{\mathrm{I}}(\mathrm{k})=1-\mu_{\mathrm{Q}}(\mathrm{k}-1)\right.$ and $\left.\mu_{\mathrm{I}}(0)=1\right)$. Note that the x's outside the support of A have no direct influence in the evaluation (since they contribute a term equal to 1 in the min-aggregation). However, it is the whole cardinality of X which is taken into account in case I refers to a relative quantifier expressing a proportion.

Example 6. Let us consider the case described in Table 4. If the worst case $\mathrm{x}_{10}$ is completely neglected (i.e., $\mu_{\mathrm{I}}(10)=0$ and $\mu_{\mathrm{I}}(9)=\ldots=\mu_{\mathrm{I}}(1)=1$ ) we get $\mathrm{E}=0.4$. In the same spirit, if the worst two elements $\mathrm{x}_{9}$ and $\mathrm{x}_{10}$ are completely neglected (i.e., $\mu_{\mathrm{I}}(10)=\mu_{\mathrm{I}}(9)=0$ and $\left.\mu_{\mathrm{I}}(8)=\ldots=\mu_{\mathrm{I}}(1)=1\right)$ we get $\mathrm{E}=0.5$. Clearly, the greater the number of neglected elements, the greater the evaluation. Note that E may remain high although there are only a few x's which are A's

| X | A | B | $\max \left(\mu_{\mathrm{B}}, 1-\mu_{\mathrm{A}}\right)$ |
| :---: | :---: | :---: | :---: |
| $\mathrm{x}_{1}$ | 0 | 1 | 1 |
| $\mathrm{x}_{2}$ | 0 | 0.9 | 1 |
| $\mathrm{x}_{3}$ | 0.9 | 1 | 1 |
| $\mathrm{x}_{4}$ | 0.8 | 0.9 | 0.9 |
| $\mathrm{x}_{5}$ | 0.9 | 0.8 | 0.8 |
| $\mathrm{x}_{6}$ | 0.3 | 0.2 | 0.7 |
| $\mathrm{x}_{7}$ | 1 | 0.5 | 0.5 |
| $\mathrm{x}_{8}$ | 0.5 | 0.5 | 0.5 |
| $\mathrm{x}_{9}$ | 1 | 0.4 | 0.4 |
| $\mathrm{x}_{10}$ | 0.8 | 0.1 | 0.2 |

Table 4

### 4.2.2 Interpretation 2: 'Most $x$ 's are A's and B's'

In this case, we rank-order the x 's according to the decreasing values of $\min \left(\mu_{\mathrm{A}}(\mathrm{x})\right.$, $\mu_{\mathrm{B}}(\mathrm{x})$ ). In other words, if Q means 'at least k , we are looking for a subset C of X such that:

$$
\mathrm{ICl}=\mathrm{k}, \mathrm{C} \subseteq \mathrm{~A}_{\alpha}, \mathrm{C} \subseteq \mathrm{~B}_{\beta} \text { with } \alpha \text { and } \beta \text { as large as possible, }
$$

since:

$$
\mathrm{C} \subseteq \mathrm{~A}_{\alpha} \cap \mathrm{B}_{\beta} \Leftrightarrow \mathrm{C} \subseteq(\mathrm{~A} \cap \mathrm{~B})_{\min (\alpha, \beta)}
$$

This leads to rank-order the k best elements of X according to their decreasing values of $\mu_{\mathrm{A} \cap \mathrm{B}}$ and to assign, via $\mu_{\mathrm{I}}$, a degree of importance equal to 1 to the k best rated elements, the others having a level of importance equal to 0 . Then E is given again by a formula close to (3) where $\mu_{B}\left(x_{\sigma}(i)\right)$ becomes $\mu_{A \cap B}\left(\mathrm{x}_{\sigma}(\mathrm{i})\right)$ and n is the cardinality of X:

$$
\begin{equation*}
\mathrm{E}=\min _{\mathrm{i}=1, \mathrm{n}} \max \left(\mu_{\mathrm{A} \cap \mathrm{~B}}\left(\mathrm{x}_{\sigma(\mathrm{i})}\right), 1-\mu_{\mathrm{I}}(\mathrm{i})\right) \tag{11}
\end{equation*}
$$

where the values $\mu_{\mathrm{A} \cap \mathrm{B}}\left(\mathrm{x}_{\sigma}(\mathrm{i})\right)$ are such that: $\mu_{\mathrm{A} \cap \mathrm{B}}\left(\mathrm{x}_{\sigma}(1)\right) \geq \ldots \geq \mu_{\mathrm{A} \cap \mathrm{B}}\left(\mathrm{x}_{\sigma}(\mathrm{n})\right)$ and I is the set of ranks of elements somewhat important $\left(\mu_{\mathrm{I}}(\mathrm{k})=1-\mu_{\mathrm{Q}}(\mathrm{k}-1)\right.$ for $\mathrm{k}>0$ and $\mu_{\mathrm{I}}(0)$ $=1$ ).

Example 7. Let us consider the condition 'At least about 4 x's are A's and B's' where the (fuzzy) quantifier is defined by $\mu_{\mathrm{Q}}(0)=0, \mu_{\mathrm{Q}}(1)=0, \mu_{\mathrm{Q}}(2)=0.2, \mu_{\mathrm{Q}}(3)=.5$, $\mu_{\mathrm{Q}}(4)=1, \mu_{\mathrm{Q}}(5)=1, \mu_{\mathrm{Q}}(6)=1$ and the set X is given in Table 5.

Then $\mu_{\mathrm{I}}(0)=1, \mu_{\mathrm{I}}(1)=1, \mu_{\mathrm{I}}(2)=1, \mu_{\mathrm{I}}(3)=0.8, \mu_{\mathrm{I}}(4)=0.5, \mu_{\mathrm{I}}(5)=0=\mu_{\mathrm{I}}(6)$.

$$
\begin{aligned}
\mathrm{E} & \left.=\min \left(\max \left(\mu_{\mathrm{A} \cap \mathrm{~B}}\left(\mathrm{x}_{4}\right)\right), 1-\mu_{\mathrm{I}}(1)\right), \max \left(\mu_{\mathrm{A} \cap \mathrm{~B}}\left(\mathrm{x}_{1}\right)\right), 1-\mu_{\mathrm{I}}(2)\right), \\
& \left.\left.\left.\max \left(\mu_{\mathrm{A} \cap \mathrm{~B}}\left(\mathrm{x}_{3}\right)\right), 1-\mu_{\mathrm{I}}(3)\right), \max \left(\mu_{\mathrm{A} \cap \mathrm{~B}}\left(\mathrm{x}_{2}\right)\right), 1-\mu_{\mathrm{I}}(4)\right)\right) \\
& =\min (\max (0.9,0), \max (0.8,0), \max (0.5,0.2), \max (0.2,0.5)) \\
& 0.5
\end{aligned}
$$

| X | A | B |
| :---: | :---: | :---: |
| $\mathrm{x}_{1}$ | 1 | 0.8 |
| $\mathrm{x}_{2}$ | 1 | 0.2 |
| $\mathrm{x}_{3}$ | 0.5 | 0.5 |
| $\mathrm{x}_{4}$ | 0.9 | 1 |
| $\mathrm{x}_{5}$ | 0 | .9 |
| $\mathrm{x}_{6}$ | 0.5 | 0 |

Table 5

Note that this view is in accordance with the usual result when $A$ is not fuzzy stating that checking if 'at least $k$ elements of A are B 's' is equivalent to estimating if 'at least $k$ elements of $X$ are (A and B)'s'.

### 4.2.3 Interpretation 3: 'Most $A$ 's are $B$ 's' where $A$ is fuzzy

As already said, in such an evaluation we restrict ourselves to the x 's which are somewhat $A$. Clearly, depending on the $\alpha$-level cut $A_{\alpha}$ we consider, we would get different evaluations $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)$ by applying (3). Note that $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)$ does not vary in a monotonic way with $\alpha$ in general. This should not be surprizing. Indeed, it is wellknown that the proportion of A's which are B's varies in a nonmonotonic way when A is replaced either by a superclass, or a subclass of $A$. Then the information provided by the different $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)$ 's can be summarized in different ways.

A first approach is to propose the following evaluation:

$$
\begin{equation*}
\mathrm{E}(\mathrm{~A})=\max _{\alpha} \min \left(\alpha, \mathrm{E}\left(\mathrm{~A}_{\alpha}\right)\right) \tag{12}
\end{equation*}
$$

which corresponds to a weighted disjunction of the possible results. This is clearly an optimistic evaluation. Since we may have $\mathrm{E}=1$ because $\mathrm{E}\left(\mathrm{A}_{1}\right)=1$ while for instance $\mathrm{E}\left(\mathrm{A}_{0.9}\right)=0$ !

Example 8. Let us consider the query 'most A-x's are B's' where the (fuzzy) quantifier 'most' means that the worst B element in A is neglected. The set X is given in Table 6. The different $\alpha$-level cuts are: $\mathrm{A}_{1}=\left\{\mathrm{x}_{1}\right\}$ and $\mathrm{A}_{0.9}=\left\{\mathrm{x}_{1}, \mathrm{X}_{2}, \mathrm{X}_{3}\right\}$. We have for $\mathrm{A}_{1}: \mu_{\mathrm{I}}(1)=1$ (because this set has only one element) and thus: $\mathrm{E}\left(\mathrm{A}_{1}\right)=\max$ (1$\left.\mu_{\mathrm{I}}(1), \mu_{\mathrm{B}}\left(\mathrm{x}_{1}\right)\right)=1$. Consequently the result given by (12) is $\mathrm{E}=1$ (because $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)=1$ for $\alpha=1$ ). This result is obviously optimistic because, as it can be seen in Table 7, we would have to neglect the worst two elements in $\mathrm{A}_{0.9}$ in order to keep this evaluation E •

| X | A | B |
| :---: | :---: | :---: |
| $\mathrm{x}_{1}$ | 1 | 1 |
| $\mathrm{x}_{2}$ | 0.9 | 0.1 |
| $\mathrm{x}_{3}$ | 0.9 | 0.1 |

Table 6
Consequently, another natural evaluation which may be considered is:

$$
\begin{equation*}
\mathrm{E}(\mathrm{~A})=\min _{\alpha} \mathrm{E}\left(\mathrm{~A}_{\alpha}\right) \tag{13}
\end{equation*}
$$

which guarantees that, for any $\alpha, \mathrm{E}\left(\mathrm{A}_{\alpha}\right)$ is larger than or equal to E . This result corresponds to a conjunction of the possible results and is clearly a pessimistic
evaluation (since we may have $\mathrm{E}=0$ because $\mathrm{E}\left(\mathrm{A}_{0.1}\right)=0$ while for instance $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)=$ 1 for any $\alpha \neq 0.1$ !

Expression (13) is a conjunction which can be weighted in order to modulate its pessimistic behavior:

$$
\begin{equation*}
\mathrm{E}(\mathrm{~A})=\min _{\alpha} \max \left(1-\mu_{\mathrm{I}}(\alpha), \mathrm{E}\left(\mathrm{~A}_{\alpha}\right)\right) \tag{14}
\end{equation*}
$$

where $\mu_{\mathrm{I}}(\alpha)$ is the importance of the $\alpha$-level cut $\mathrm{A}_{\alpha}$. The idea underlying (14) is that the $\alpha$-level cuts with large $\alpha$ are the most important ones, while $\alpha$-level cuts with low $\alpha$ can be neglected (since they involve elements which are not strongly in A). In the particular case where $\forall \alpha \geq \alpha^{\prime}, \mu_{\mathrm{I}}(\alpha)=1$ and $\mu_{\mathrm{I}}(\alpha)=0$ otherwise, we obtain:

$$
\begin{equation*}
\mathrm{E}(\mathrm{~A})=\min _{\alpha \geq \alpha^{\prime}} \mathrm{E}\left(\mathrm{~A}_{\alpha}\right) \tag{15}
\end{equation*}
$$

where $\alpha^{\prime}$ can be viewed as a membership threshold for A-elements. It means that if $\mu_{\mathrm{A}}(\mathrm{x})<\alpha^{\prime}$ then x does not sufficiently belong to A and thus can be neglected.

In the special case where $\mu_{\mathrm{I}}(\alpha)=\alpha$ in (14), which expresses that 'the higher $\alpha$, the more important $\mathrm{A}_{\alpha}$ and $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)^{\prime}$, we get:

$$
\begin{equation*}
\mathrm{E}(\mathrm{~A})=\min _{\alpha} \max \left(1-\alpha, \mathrm{E}\left(\mathrm{~A}_{\alpha}\right)\right) \tag{16}
\end{equation*}
$$

The expressions (12) and (16) can be viewed respectively as the possibility and the necessity (see note 2), of a fuzzy event corresponding to the $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)$ 's based on the possibility distribution $\pi(\alpha)=\alpha$ for $\alpha \in[0,1]$ ( $\alpha$ being the possibility that $\mathrm{A}_{\alpha}$ represents the fuzzy set A ; if $\mathrm{A}_{\alpha}=\varnothing$ then $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)=0$ ). The estimates (12) and (16) can be viewed as scalar summaries of the fuzzy-valued estimate $\mathrm{E}^{*}$ where $\mu_{\mathrm{E}}^{*}\left(\mathrm{E}_{\alpha}\right)=$ $\alpha$; see Dubois and Prade [7]. In the particular case where Q is 'for all' (then $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)$ is the minimum of $\mu_{\mathrm{B}}(\mathrm{x})$ for x belonging to $\mathrm{A}_{\alpha}$ ) it has been pointed out [4] that (6) and (16) lead to the same result. (It can be seen as a consequence of (5), by introducing the degree of inclusion of $A_{\alpha}$ into the fuzzy set $B$ defined by (1)-(2), in the expression (5).)

Example 9. Let us consider the query 'most A-x's are B's' where the (fuzzy) quantifier 'most' means that the worst B element in A can be neglected. The set X is given in Table 7. The different $\alpha$-level cuts are: $\mathrm{A}_{1}=\left\{\mathrm{x}_{1}\right\}, \mathrm{A}_{0.9}=\left\{\mathrm{x}_{1}, \mathrm{x}_{2}\right\}, \mathrm{A}_{0.5}=$ $\left\{\mathrm{x}_{1}, \mathrm{x}_{2}, \mathrm{x}_{3}, \mathrm{x}_{4}\right\}, \mathrm{A}_{0.4}=\left\{\mathrm{x}_{1}, \mathrm{x}_{2}, \mathrm{x}_{3}, \mathrm{x}_{4}, \mathrm{x}_{5}, \mathrm{x}_{6}\right\}$. We have for $\mathrm{A}_{1} \mu_{\mathrm{I}}(1)=1$ (because this set has only one element) and thus (3) gives: $\mathrm{E}\left(\mathrm{A}_{1}\right)=\max \left(1-\mu_{\mathrm{I}}(1), \mu_{\mathrm{B}}\left(\mathrm{x}_{1}\right)\right)=1$. Considering $\mathrm{A}_{0.9}$ we have $\mu_{\mathrm{I}}(1)=1$ and $\mu_{\mathrm{I}}(2)=0$ thus $\mathrm{E}\left(\mathrm{A}_{0.9}\right)=1$ (the same result would be obtained even without neglecting $\mathrm{x}_{2}$ ). Considering $\mathrm{A}_{0.5}$ we have $\mu_{\mathrm{I}}(1)=$ $\mu_{\mathrm{I}}(2)=\mu_{\mathrm{I}}(3)=1$ and $\mu_{\mathrm{I}}(4)=0$ and thus $\mathrm{E}\left(\mathrm{A}_{0.5}\right)=0.8$ (neglecting $\mathrm{x}_{4}$ ). Considering $\mathrm{A}_{0.4}$ we have $\mu_{\mathrm{I}}(1)=\mu_{\mathrm{I}}(2)=\mu_{\mathrm{I}}(3)=\mu_{\mathrm{I}}(4)=\mu_{\mathrm{I}}(5)=1$ and $\mu_{\mathrm{I}}(6)=0$ thus $\mathrm{E}\left(\mathrm{A}_{0.4}\right)=0$ (since we are neglecting at most one element). If definition (13) is chosen for $E$ we have:

$$
\mathrm{E}(\mathrm{~A})=\min _{\alpha} \mathrm{E}\left(\mathrm{~A}_{\alpha}\right)=0,
$$

which is a pessimistic evaluation (since this result is induced by a poor member of A (0.4)). If definition (15) is taken for $E$, we face the problem of choosing $\alpha^{\prime}$. In this example, let $\alpha^{\prime}$ be 0.5 , we get:

$$
\mathrm{E}(\mathrm{~A})=\min _{\alpha \geq 0.5} \mathrm{E}\left(\mathrm{~A}_{\alpha}\right)=0.8
$$

However, one may argue that $\alpha^{\prime}$ is a precise boundary that cannot be always clearly justified. Furthermore, if we choose $\alpha^{\prime}=0.4$ in this example we get $E(A)=0$ which shows that two different but close thresholds ( 0.5 and 0.4 ) could lead to two extremely different results ( 0.8 and 0 ). Is 0.4 or 0.5 more appropriate to give a significant result? That is why the evaluation given by expression (16) may be preferred:

$$
\mathrm{E}(\mathrm{~A})=\min _{\alpha} \max \left(1-\alpha, \mathrm{E}\left(\mathrm{~A}_{\alpha}\right)\right)=0.6
$$

In this last case, each value $\mathrm{E}\left(\mathrm{A}_{\alpha}\right)$ is weighted by the importance $\alpha$ of the considered $\alpha$-level cut and the contribution of each $\mathrm{A}_{\alpha}$ to the overall result is more significant when $\mathrm{A}_{\alpha}$ only gathers elements which are strong members of A

| X | A | B |
| :---: | :---: | :---: |
| $\mathrm{x}_{1}$ | 1 | 1 |
| $\mathrm{x}_{2}$ | 0.9 | 1 |
| $\mathrm{x}_{3}$ | 0.5 | 0.8 |
| $\mathrm{x}_{4}$ | 0.5 | 0.2 |
| $\mathrm{x}_{5}$ | 0.4 | 0 |
| $\mathrm{x}_{6}$ | 0.4 | 0 |

Table 7

## 5 Conclusion

This paper has mainly intended to provide a preliminary discussion of the qualitative handling of evaluations of the form 'Q A-x's are B's'. In practice, considering a query like 'Find the departments where most young people are well-paid', we would first rank-order the departments according to the extent to which all young people are wellpaid'. If no (or too few) departments are retrieved with a positive evaluation, we would restart the evaluation process changing 'all' into 'all except one', and then relaxing the requirement still more, if necessary.It is also clear that when $h(A)$ is less than 1, i.e., A is somewhat empty, this fact should be notified to the user.

Besides, we have insisted on the qualitative nature of the evaluation to provide, since it is not always clear in practice that the membership grades can receive a genuine interpretation in terms of a real number. It is then important to keep the evaluation and the ranking process as robust as possible. However, it would be interesting to clarify the differences between the approach proposed here and an OWA operation-based aggregation of the elementary evaluations for each item $x$ [17], since OWA aggregation can be also nicely interpreted in terms of quantifiers. The expressions used in this paper are Ordered Weighted Minimums [11], rather than Ordered Weighted Averages (OWA's) which cannot be defined on purely ordinal scales as the Ordered Weighted Minimums. More generally, it would be useful to undertake some practical experiments in order to assess the approach and its situation with respect to those based on cardinalities (e.g., [17, 20]).

More generally, it could also be of interest to see the problem of the evaluation of expressions of the form 'Q-A's $x$ are B's' as the detection of the (in)stability of the evaluations corresponding to various crisp approximations of $A$ or $B$ in terms of $\alpha$ cuts. For instance, we may look for what (high) values of $\alpha$ the evaluation of 'Q $\mathrm{A}_{\alpha}$ elements are B's' remains constant (once the interpretation of Q is chosen). This can be related to data summarization issues since the $\alpha$-level cuts with high $\alpha$ may be viewed as different levels of approximation regarding the evaluation of the condition 'Q A-x's are B'. Going back to our example, once a department has been identified as satisfying the condition 'most young people are well-paid' to some extent (for some interpretation of 'most' of the type 'all except a few'), it might be interesting to find out if, for instance, the condition is more, or is less, satisfied if we neglect the people who are not really young. With such queries, it is also important to make clear that the evaluation may be quite different for apparently rather similar expressions, for example, if we look for 'the departments where all people are rather well-paid', or for 'the departments where almost all people are very-well-paid'.

## References

[1] W. Bandler, L.J. Kohout, Fuzzy power sets and fuzzy implication operators. Fuzzy Sets and Systems, 4, 1980, 13-30.
[2] P. Bosc, O. Pivert, SQLf: A relational database language for fuzzy querying. IEEE Trans. on Fuzzy Systems, 3, 1995, 1-17.
[3] P. Bosc, H. Prade, An introduction to the fuzzy set and possibility theory-based treatment of soft queries and uncertain or imprecise databases. In: Uncertainty Management in Information Systems: From Needs to Solutions (Ph. Smets, A. Motro, eds.), Kluwer Academic Publ., 1997, Chapter 10, 285-324
[4] P. Bosc, L. Liétard, Une interprétation pour 'Q B X sont A'. BUSEFAL (IRIT, Univ. P. Sabatier, Toulouse, France), 68, 1996, 9-19.
[5] D. Dubois, H. Fargier, H. Prade, Refinements of the maximin approach to decisionmaking in fuzzy environment. Fuzzy Sets and Systems, 81, 1996, 103-122.
[6] D. Dubois, H. Prade, Fuzzy Set and Systems: Theory and Applications. Academic Press, New York, 1980.
[7] D. Dubois, H. Prade, Evidence measures based on fuzzy information. Automatica, 21, 1985, 547-562.
[8] D. Dubois, H. Prade, Fuzzy cardinality and the modeling of imprecise quantification. Fuzzy Sets and Systems, 16, 1985, 199-230.
[9] D. Dubois, H. Prade, Weighted minimum and maximum operations. Information Sciences, 39, 1986, 205-210.
[10] D. Dubois, H. Prade, Measuring properties of fuzzy sets: A general technique and its use in fuzzy query evaluation. Fuzzy Sets and Systems, 38(2), 1990, 137-152.
[11] D. Dubois, H. Prade, Semantics of quotient operators in fuzzy relational databases. Fuzzy Sets and Systems, 78, 1996, 89-93.
[12] L. Liétard, Contribution à l'interrogation flexible de bases de données: Etude des propositions quantifiées floues. Thesis, Université de Rennes I, France, 1995.
[13] H. Prade, Modal semantics and fuzzy set theory. In: Fuzzy Set and Possibility Theory - Recent Developments (R.R. Yager, ed.), Pergamon Press, New York, 1982, 232-246.
[14] M. Wygralak, Vaguely Defined Objects. Kluwer Academic Publ., Dordrecht, 1996.
[15] R.R. Yager, An approach to inference in approximate reasoning. Int. J. of ManMachine Studies, 13, 1980, 323-328.
[16] R.R. Yager, General multiple-objective decision functions and linguistically quantified statements. Int. J. of Man-Machine Studies, 21, 1984, 389-400.
[17] R.R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making. Trans. on Systems, Man and Cybernetics, 18, 1988, 183-190.
[18] L.A. Zadeh, Fuzzy sets, Information and Control, 8, 1965, 338-353.
[19] L.A Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems, 1, 1978, 3-28.
[20] L.A. Zadeh, A computational approach to fuzzy quantifiers in natural languages. Computer Mathematics with Applications, 9, 1983, 149-183.


[^0]:    ${ }^{1}$ Indeed we are not interested here in knowing to what extent it is true that, for instance, "there are approximately eighty per cent of the A's which are B's", but rather to find out under what acceptable weakening of "for all" into "most", it is completely true that "most A's are B's" (or at least, to know if such a weakening exists).

