Extending Microaggregation Procedures using Defuzzification Methods for Categorical Variables

Josep Domingo-Ferrer, Vicenç Torra

Abstract— **Defuzzification is one of the fundamental steps in the development of fuzzy knowledge based systems. Given a fuzzy set μ over the reference set X, defuzzification applied to μ returns an element of X. While a large number of methods exists for the case of X being a numerical scale, only few methods are applicable when X corresponds to a categorical scale.**

Aggregation procedures have been extensively used in defuzzification in numerical scales. This is so because defuzzification has been studied as equivalent to the computation of an expected value. In this work we present the reversal approach, we study defuzzification procedures for their application to aggregation. We focus on the development of defuzzification methods for the case of X being an ordinal scale. This is, X is a set of finite values in which a total order is defined. Our ultimate goal is to apply these methods to microaggregation (a Statistical Disclosure Risk).

Index Terms— **Defuzzification, ordinal scales, aggregation procedures, selection procedures.**

I. Introduction

Fuzzy knowledge based systems (see [1] and [2] for details) are one of the most successful applications of fuzzy sets [3]. These systems are rule based systems in which the predicates of the antecedents and consequents are defined in terms of fuzzy sets. This is, given a fuzzy rule of the form

$$\textbf{if } X \textit{ is } A \textbf{ and } \ldots \textbf{and } Y \textit{ is } B \textbf{ then } Z \textit{ is } C$$

X, Y and Z are variables and A, B and C are fuzzy sets on the reference sets of X, Y and Z. We will denote the reference set of a variable X by D_X. For example, if X corresponds to temperature, A can correspond to a fuzzy set describing values near zero degrees, in this case, the reference set D_X corresponds to real numbers or a subset of them ($D_X = \mathbb{R}$).

Given a piece of information on the variables of the antecedent (their value or a fuzzy set describing possible values), the information is propagated into the conclusion to get information about the possible values of the variable in the conclusion. In the case above, given values for variables $X, \cdots, Y$, a system would obtain possible values for Z. These values are usually described by means of a fuzzy set on the reference set of Z (i.e., D_Z). If no information can be inferred, the set of possible values for Z is usually set equal to the empty set. In fact, this corresponds to disjunctive systems (conjunctive systems make, in this case, the conclusion equal to the whole set [1]). In this way, the conclusion of each rule takes into account available information (the values of the variables) and the relevance of this information to the rule at hand (in which degree the antecedent is fulfilled).

Most fuzzy knowledge based systems correspond to one stage fuzzy systems. These systems are defined by means of a flat set of rules, all having the same set of variables in the antecedent and the same variables in the consequence. Then, in a given instant of time, all rules are applied and the conclusions of all rules are combined to build the conclusion of the system. This combined conclusion corresponds to a fuzzy set (we will denote this combined fuzzy set by $\mu_{\mathbb{C}}$). A final step of the system is defuzzification. This is to transform the fuzzy set $\mu_{\mathbb{C}}$ into a value in the reference set of the conclusion variable. For the example above, this is: $defuzzification(\mu_{\mathbb{C}}) \in D_Z$

When studying defuzzification methods an important aspect to be taken into account is the scale of the reference set. Three main types can be distinguished:

Numerical: The conclusion is either a real number or a natural number.

Ordinal scale: The conclusion is categorical (e.g., linguistic labels) and values are ordered (there is a total relation over the categories).

Nominal scale: The conclusion is categorical but there is no ordering relation between categories.

While most control systems fall into the first type, most configuration systems fall in the second and third type (e.g., the one in [4]).

At present there exists a large number of defuzzification procedures for numerical scales, while existing methods for ordinal scales are more limited. In this work we study defuzzification methods for ordinal scales. Our interest in these methods is for their application to the aggregation of values (fusion of information) in these scales.

In fact, usually the relationship between aggregation (fusion) and defuzzification is observed in the other direction. This is so because, two main approaches are considered in defuzzification (this is explained in more detail in the rest of the paper):

Element selection: From the reference set of the conclusion variable, one element has to be selected. This selection has to be based on the combined fuzzy set. In the example above, this means to select a value from the set $\mu_{\mathbb{C}}$ for the variable Z in the reference set D_Z.

Aggregation: The defuzzified value corresponds to the aggregation of the available information. This can be seen as

J. Domingo-Ferrer is with the Department of Computer Engineering and Maths - ETSE, Universitat Rovira i Virgili, Avga Països Catalans 26, E-43007 Tarragona, Catalonia. E-mail: jdomingo@etse.urv.es

V. Torra is with the Institut d'Investigació en Intel·ligència Artificial - CSIC, Campus UAB s/n, E-08193 Bellaterra, Catalonia. E-mail: vtorra@iiia.csic.es

equivalent to returning an expected value of the combined fuzzy set.

According to the second approach, almost all aggregation procedures can be used for defuzzification. This is particularly interesting in the case that the variables are defined in numerical scales due to the large number of aggregation procedures. Also, the selection of an appropriate procedure makes possible the customization of an application. In the case of variables on ordinal or numerical scales, this approach is not as appropriate because there are only a few applicable aggregation procedures. Nevertheless, the relationship can be studied in the reversal direction to develop new aggregation procedures, and at the same time extending defuzzification methods.

II. Aggregation Procedures in Ordinal Scales

The development of aggregation procedures for ordinal scales is briefly reviewed in [5]. A more detailed review that also analyzes the properties of the procedures in the view of prototype selection in clustering is given in [6]. It has to be noted that, in some sense, the properties of a defuzzification method are similar to the ones of the methods for selecting the prototype of a class and, thus, the analysis in that work is relevant here.

In the rest of the paper we assume that the values to be aggregated (or the domain X in the case of defuzzification) are the following ones: $L = \{l_0, \cdots, l_R\}$ where $l_0 \leq_L l_1 \leq_L \cdots \leq_L l_R$.

Aggregation procedures in ordinal scales can be classified into three main classes according to the semantics underlying the categories (or linguistic labels) in L:

Explicit quantitative or fuzzy scales: There is a function that assigns to each value l_i a value in a numerical (or fuzzy scale). Aggregation operators can be defined on L according to this translation function. An operator of this class is described in [5].

Implicit numerical scale: In this class of operators the translation is not explicitly defined but it is implicitly assumed. Two operators of this class are the Linguistic OWA [7] and the Linguistic WOWA [8].

Operating directly on qualitative scales: The operators of this class do not assume the existence of any translation function either implicitly or explicitly. Operators are only based on operators that can be directly defined in the ordinal scale. Example of operators that can be defined in the ordinal scale are the *minimum* and the *maximum* (that rely on the relation $\leq_L$) and also t-norms and t-conorms (because these operators can be defined axiomatically – see [9]).

When we restrict to the third class, the number of existing operators is small. In fact, four main families can be distinguished:

Plurality rule: This function selects the set of most frequent elements. Weighted plurality rule can also be defined. In this case, the *most frequent* elements are the ones that accumulate *larger weights.*

Median and order statistics: This function selects the element that occupies the central position when all the elements to be aggregated are ordered. A weighted median can also be defined. We classify *order statistics* procedures in this class, because they correspond to the selection of those elements that occupy other positions (e.g., first, second or last) than the central one.

Sugeno integral: This integral can be interpreted as an aggregation operator. In this case, values are aggregated taking into account a fuzzy measure. This fuzzy measure is used to measure the importance of the information sources (the sources that supply the values to be aggregated). It is well known that the Sugeno integral generalizes several other aggregation operators as the weighted minimum and the weighted maximum.

Ordinal weighted mean: This operator is the ordinal counterpart of the weighted mean. The definition of this operator tries to mimic the definition of the numerical weighted mean. Instead of addition and product, it uses t-conorms and t-norms. Ordinal Choquet integral has also been defined using the same approach.

Some of the difficulties underlined in [6] for the application of these operators to prototype selection are the following ones:

1. *Difficulty for defining the parameters by non experienced users:* Operators like the Sugeno integral and the Ordinal weighted mean that need that someone defines fuzzy measures or t-norms/t-conorms are not appropriate because these parameters are not easy to define.
2. *Inconvenience of non-compensative operators:* For most reviewed operators, the result of the aggregation should be one of the values to be aggregated. This means that the average of a small and a large value cannot be a value somewhere in between. Instead, compensation is allowed in the numerical case. E.g., the aggregation of 0 and 1 is 0.5 when the aggregation operator is the arithmetic mean.
3. *Difficulty of defining parametric operators:* Parameterization of existing operators is difficult. Some of the operators require weights in ordinal scales. Fuzzy measures are even more difficult to define because they are defined over parts of the sources and, here, the only information is the membership value for each element.

Taking all these aspects into account it seems that the most appropriate aggregation procedure for prototype selection is the median. However, it does not allow for compensation and does not include any parameterization. To overcome these two difficulties, the CWOW-Median was defined in [6]. Its definition is as follows:

Definition 1: Let $\mathbf{p} : \mathbf{X} \rightarrow \mathbf{D} \subset \mathbf{R}$ be a weighting vector, let Q be a non-decreasing fuzzy quantifier, then a mapping $CWOW - Median_{\mathbf{p}} : L^N \rightarrow L$ is a *Convex WOW-Median* of dimension N if:

$$CWM_{\mathbf{w}}(a_1, \cdots, a_N) = a \text{ iff } acc'''(a) > 0.5 \geq acc'''(b)$$

where acc''' is the WOW-weighting vector of (L, acc'') and Q, $acc''(a) = acc'(a)/\sum_{b \in L} acc'(b)$, $acc'(a) = \min(\max_{b \leq a} acc(b), \max_{b \geq a} acc(b))$, $acc(a) = \sum_{f(x_j)=a} p(x_j)$ and where b is the element next to b in L. This is, $b = \max\{x | x \in L, x < a\}$.

Where the WOW-weighting vector is computed according to the following definition:

Definition 2: Let $(a_i, p_i)_{i=1,N}$ be a pair defined by a value a_i and the importance of this value (the value p_i) expressed in a given domain $D \subset \mathbb{R}^+$, and let Q be a fuzzy non-decreasing fuzzy quantifier. Then, the WOW-weighting vector $\omega = (\omega_1, \cdots, \omega_N)$ for $(a, \mathbf{p}) = ((a_1, \cdots, a_N), (p_1, \cdots, p_N))$ and Q is defined as follows:

$$\omega_i = Q\left(\frac{\sum_{j \leq i} p_{\sigma(j)}}{\sum_{j \leq N} p_{\sigma(j)}}\right) - Q\left(\frac{\sum_{j < i} p_{\sigma(j)}}{\sum_{j \leq N} p_{\sigma(j)}}\right)$$

where $\{\sigma(1), ..., \sigma(N)\}$ is a permutation of $\{1, ..., N\}$ such that $a_{\sigma(i-1)} \geq a_{\sigma(i)}$ for all $i = \{2, ..., N\}$ (i.e. $a_{\sigma(i)}$ is the i-th largest element in the collection $a_1, ..., a_N$), and Q is a non-decreasing fuzzy quantifier. This is, Q is a monotonic function (i.e., $Q(a) \geq Q(b)$ for all $a > b$) such that $Q(0) = 0$ and $Q(1) = 1$.

III. Defuzzification Procedures

In this section we review some of the defuzzification procedures. For more details see [12] and [10].

Defuzzification methods have been studied from different perspectives. Yager [10] views defuzzification in the more general framework of a *selection problem*. This is, selecting an element using the information represented in the fuzzy set. Additional knowledge can also be taken into account in this selection (e.g. using constraints about values of variables in [4]). This latter case corresponds, according to Yager [11], to the so-called *Knowledge based defuzzification*.

To study the defuzzification process, Yager proposed a general architecture. This is shown in Figure 1. Following this figure, the output of the fuzzy rule based system is a fuzzy set F (we will use μ_F to denote its membership function). The defuzzification of this fuzzy set is achieved via a two stage process. First, the fuzzy set F is transformed into a probability distribution P, and then one element is selected from the probability distribution (two general methods $S1$ and $S2$ – see below – are considered for selection). In [11], Yager introduces an additional stage for transforming the original fuzzy set into a *transformed* one F' (with membership function $\mu_{F'}$) using additional knowledge.

Here we consider an hybrid approach consisting on only two stages as in Figure 1 but in which the first stage (this is called *fuzzy set transformation*) considers all the additional knowledge, and its output F' can be another fuzzy set. The second stage is named *Element selection*. This is shown in Figure 2.

We now review some of the existing alternative approaches for the two stages of fuzzy set transformation and

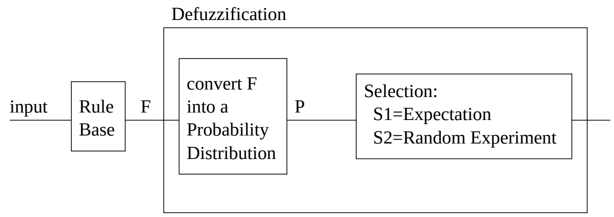

Fig. 1. Defuzzification process following Yager

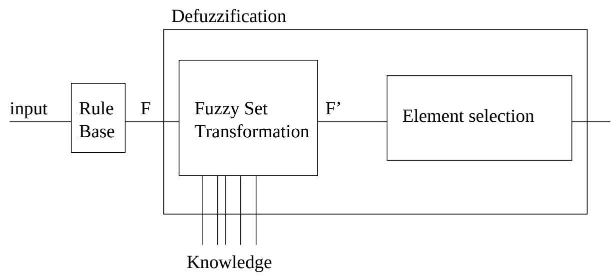

Fig. 2. Our approach to the defuzzification process

element selection. The following two subsections are devoted to these aspects.

A. Fuzzy Set Transformation

Below we list some of the appropriate fuzzy set transformation processes. An alternative method based on clustering is described in [11].

Normalization: Membership values are scaled so that they add to one. This is:

$$\mu_{F'}(x) = \frac{\mu_F(x)}{\sum_{x_i \in X} \mu_F(x_i)}$$

This process can also be understood as transforming the fuzzy set into a probability distribution. The well-known Center of Area method (see e.g. [12], [10] for its description) requires the application of this transformation.

Selection of most possible values: Objects x_i in X with a larger membership value are selected. Let $F_{max} = max \mu_F(x)$, then:

$$\mu_{F'}(x) = \begin{cases} 1 & \text{if } \mu_F(x) = F_{max} \\ 0 & \text{otherwise} \end{cases}$$

The Mean of Maxima method requires the application of this transformation.

α-cut of the membership function: All values with a membership value less than a given α are disregarded.

$$\mu_{F'}(x) = \begin{cases} \mu_F(x) & \text{if } \mu_F(x) \geq \alpha \\ 0 & \text{otherwise} \end{cases}$$

This transformation is applied to avoid the inclusion of values that have a possibility below a given threshold.

As these definitions are functionally defined, composition of transformations are possible in a single defuzzification method. For example, we can apply a given α-cut and then normalize the resulting membership function to obtain a probability distribution that disregards some values with a low possibility.

B. Element Selection

For element selection, Yager [11] made the classification given below. The two classes considered roughly correspond to the two approaches for defuzzification informally reviewed in the introduction (and to $S1$ and $S2$ in Figure 1): element selection and aggregation.

Blending methods: The defuzzified value is obtained as the combination of available solutions. These methods usually use some kind of average to combine the solutions. Differences between methods correspond to different ways of averaging the values. For example, the following aggregation operators have been considered in the literature:

1. **Arithmetic mean of the values:** The Mean of Maxima can be computed using this procedure for element selection.
2. **Weighted mean of the values:** Usually weights are linearly proportional to the membership values in μ_F. This would be the case of the Center of Area. The Mean of Maxima can also be computed using this approach (in this case, all weights are equal).

Celibate methods: They do not combine solutions but take one of the elements of X as its solution. Some particular examples of celibate methods are:

1. Random selection of one of the elements (with random numbers following e.g. a Normal distribution).
2. Random selection with the probability of selecting an element being proportional to its fuzzy membership value. This is the RAGE (RAndom GEneration defuzzification) family of methods [11].

IV. Aggregation and Defuzzification Procedures

As seen in Section III, aggregation is one of the blending methods for element selection. In particular, given a membership function μ_F on the (discrete) reference set X, the defuzzification of μ_F can be seen as the aggregation of the elements $x_i \in X$ with respect to the *weighting vector* μ_F. Thus, $\mu_F(x_i)$ is interpreted as the weight of x_i.

According to this, for a given measure scale (e.g., numerical, ordinal, nominal), all aggregation operators in that kind of scale that use as additional information a numerical weighting vector can be used for defuzzification.

The other relationship between aggregation and defuzzification procedures is that the latter can be seen as an aggregation procedure when values are weighted. This is, given a set of values $a_1, \cdots, a_N$ (with $a_i \in X$) to be aggregated with weights $p_1, \cdots p_N$ (p_i is the weight attached to a_i), the following *"fuzzy set"* on X can be defined: $\mu_{A,p}(x_i) = \sum_{a_j = x_i} p_j$. Then, we can define the aggregation of $a_1, \cdots, a_N$ with respect to $p_1, \cdots, p_N$ as the defuzzified value of $\mu_{A,p}$.

V. Extending Defuzzification Methods in Ordinal Scales

The architecture described above for defuzzification is suitable for any kind of scale. However, while most membership transformations (as the ones described in this work) are independent of the type of scale, this is not as clear for selection methods. Note that blending methods are difficult to apply due to the difficulty of defining aggregation functions in ordinal scales. In particular, the arithmetic mean and the weighted mean (enumerated above) are not applicable because the values to be aggregated belong to an ordinal scale.

To replace these numerical aggregation operators, the operators described in Section II can be used.

Without considering their properties, the (weighted) plurality rule, the (weighted) median, the weighted minimum and maximum and (after some adaptation) the ordinal weighted mean can be applied. Instead, the Sugeno integral cannot. The main difficulty for the Sugeno integral is that when applied in ordinal scales it needs a fuzzy measure in the same ordinal scale than the values. This is difficult to be defined and, moreover, the only available information is numerical (the weights p_i) instead of ordinal. The second inconvenient us that the role of the fuzzy measure is similar to the role of the weights, and both elements are difficult to be combined.

Of these available and applicable operators, the most appropriate one is the Median. The ordinal weighted mean could be used, but it requires a definition of the t-norms and t-conorms involved in the process. The inconvenience of the plurality rule and the median is that they do not allow for compensation. However, the plurality rule has an additional inconvenient because a small variation on the input data can provoke a large variation in the output. Therefore, the result is not much stable. This is a drawback specially relevant when considering the application of procedures to defuzzification.

The CWOW-Median can also be applied and has the advantages of the Median and, moreover, it allows for some compensation (recall that the median lacks this property) and it also includes a parameterization by means of the fuzzy non-decreasing quantifier Q. As shown in [6], this parameterization allows a smooth transition between the smallest value being aggregated and the largest one. This is achieved with $Q(x) = x^\alpha$ and, respectively, with $\alpha = 0$ and $\alpha = \infty$

The CWOW-Median can be decomposed into two stages to adapt it to the architecture for defuzzification described in Section III. The decomposition of the operator into three components (convex transformation, WOW transformation and median) suggests new defuzzification methods. They are built through the combination of these transformations with alternative (other than Median) selection procedures.

In the next section, we introduce some new fuzzy set transformation functions. Besides of the new transformations based on CWOW-Median, we introduce aggregation based ones and we present an example using the Choquet integral.

A. Fuzzy Set Transformation

We start with the two transformations corresponding to the components of the CWOW-Median operators.

Convex membership function: This is to permit the selection of a value with a null membership function if it is located between elements with non-null membership functions. This is solved making the fuzzy set a convex fuzzy set. Given the fuzzy set μ_F, the convex fuzzy set $\mu_{F'}$ is defined as:

$$\mu_{F'}(x) = \min(\max_{b \leq x} \mu_F(b), \max_{b \geq x} \mu_F(b))$$

WOW transformation: Given a non-decreasing fuzzy quantifier Q, the new membership function is defined as:

$$\mu_{F'}(l_i) = Q\left(\frac{\sum_{j \leq i} \mu_F(l_j)}{\sum_{l_j \in L} \mu_F(l_j)}\right) - Q\left(\frac{\sum_{j < i} \mu_F(l_j)}{\sum_{l_j \in L} \mu_F(l_j)}\right)$$

Aggregation transformation: Given a membership function, the new membership function is defined using an aggregation operator. This is, the value for a given category is the aggregation of the membership of nearby categories. This is, given an aggregation operator $\mathbb{C}$, we compute $\mu_{F'}(l_i)$ by:

$$\mu_{F'}(l_i) = \mathbb{C}(\mu_F(l_1), \cdots, \mu_F(l_R))$$

The aggregation operator needs to be customized (via a parameterization) for all categories l_i. This can be better expressed by:

$$\mu_{F'}(l_i) = \mathbb{C}_i(\mu_F(l_1), \cdots, \mu_F(l_R))$$

where $\mathbb{C}_i$ means that the operator (or some *internal* parameters) depends on the label l_i.

Note that if the aggregation operator is not customized for each category, the result would be the same for all categories, because in our definition the function always receives the same parameters.

The transformation based on aggregation can be used to make the membership function smoother. This transformation is similar to the smoothing of data in signal processing. An example of using aggregation for fuzzy set transformation is described in subsection V-A.1.

A.1 Example: Choquet Integral for Defuzzification

As said above, the approach for using aggregation operators for fuzzy set transformation is to consider the new value $\mu_{F'}(l_i)$ for a given category l_i as the aggregation of the value $\mu_F(l_i)$ with the values $\mu_F(l_j)$ for $j \neq i$. To compute this value, we use the Choquet integral. The Choquet integral is a numerical aggregation operator that aggregates some values with respect to a fuzzy measure.

Definition 3: A fuzzy measure μ on a set X is a set function $\mu : \wp(X) \to [0,1]$ satisfying the following axioms:
(i) $\mu(\emptyset) = 0$, $\mu(X) = 1$ (boundary conditions)
(ii) $A \subseteq B$ implies $\mu(A) \leq \mu(B)$ (monotonicity)

Fuzzy measures replace the axiom of additivity in probability measures ($\mu(A \cup B) = \mu(A) + \mu(B)$ when $A \cap B = \emptyset$) by a more general one: monotonicity. Thus, probability measures are also fuzzy measures. Fuzzy measures are used in Choquet integrals to express the importance of a set of information sources and their redundance and complementariness. When additivity is not satisfied, it means that the importance of a set is not the addition of the importance of the elements by themselves. In our case, as we aggregate the membership values of the categories, the measure corresponds to the importance of the category and their relationship with other categories. This can express whether a category can be aggregated with another one or not.

The definition of the Choquet integral follows:

Definition 4: Let μ be a fuzzy measure on X. The *Choquet integral* of a function $f : X \to \mathbb{R}$ with respect to μ is defined by:

$$(C)\int f d\mu = \sum_{i=1}^{n} (f(x_{s(i)}) - f(x_{s(i-1)}))\mu(A_{s(i)})$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that $0 \leq f(x_{s(1)}) \leq ... \leq f(x_{s(N)}) \leq 1$, $A_{s(i)} = \{x_{x(i)}, ..., x_{s(N)}\}$ and $f(x_{s(0)}) = 0$.

In our case, and according to the remark in the previous section that the aggregation operator has to be parameterized in a convenient way for each category, we adapt the Choquet integral to compute the membership of the category l_i (this is, to compute the value $\mu_{F'}(l_i)$) as follows:

Definition 5: Let μ_{l_i} be a fuzzy measure on X for category l_i. The *Choquet integral* of the membership function $\mu_F : X \to \mathbb{R}$ with respect to μ_{l_i} is defined by:

$$\mu_{F'}(l_i) = \sum_{j=1}^{n} (\mu_F(l_{s(j)}) - \mu_F(l_{s(j-1)}))\mu_{l_i}(A_{s(j)})$$

where $f(x_{s(j)})$ indicates that the indices have been permuted so that $0 \leq f(x_{s(1)}) \leq ... \leq f(x_{s(N)}) \leq 1$, $A_{s(j)} = \{x_{s(j)}, ..., x_{s(N)}\}$ and $f(x_{\sigma(0)}) = \emptyset$.

To make the integral applicable, we need a fuzzy measure. This measure depends on our prior knowledge about the categories, and, as said, is different for each category l_i (we denote this measure by μ_{l_i} as above). In our case, we propose the following fuzzy measure (the measure is defined in terms of its Möbius transformation m_{l_i}):

Given a category $l_i \in X$ and a constant $K \in [0,1]$, we define the Möbius transform m_{l_i} of the fuzzy measure μ_{l_i} as follows:

$$m_{l_i}(A) = \begin{cases} K & \text{if } A = l_i \\ K/|A| & \text{if } l_i \in A \text{ and } A \text{ is convex} \\ 0 & \text{otherwise} \end{cases}$$

with a convex set A, we mean a set such that for all a, b in A, all the values c such that $a \leq c \leq b$ also belong to A.

The rational of this measure is that the only sets A to be taken into account when computing the membership value of the category l_i are the ones that are *connected* with l_i. This is, if a set contains l_j with $l_j > l_i$ but does not contains a category in between, the set is not considered. The relative importance of a set is larger for smaller sets, and the maximum importance is given to the set with the category l_i.

To avoid that $\mu_{F'}(l_i) > 1$, we need that the fuzzy measure defined above is normalized in the sense that $\mu_{l_i}(X) = 1$. This can be achieved in two ways. The simplest way is to define an alternative Möbius transform as follows:

$$m'_{l_i}(B) = m_{l_i}(B) / \sum_{A \subseteq X} m_{l_i}(A)$$

and use this function to compute the measure μ_{l_i}. Recall (see [13]) that this is achieved as follows:

$$\mu_{l_i}(B) = \sum_{A \subseteq B} m'_{l_i}(A)$$

An alternative way is to use the same normalization factor for all fuzzy measures μ_{l_i}. This means to define the normalization factor (NF) as follows:

$$NF = \max_{l_i \in X} \sum_{A \subseteq X} m'_{l_i}(A)$$

and define

$$m''_{l_i}(B) = m_{l_i}(B)/NF$$

The first normalization approach means that when $\mu_F(x) = \alpha$ for all $x \in X$, then $\mu_{F'}(x) = \alpha$ for all $x \in X$. Thus, all elements are equally possible after transformation.

The second approach is more appropriate when in a constant membership function (as above, $\mu_F(x) = \alpha$ for all $x \in X$) we are more interested in central elements of X. In other words, this means that a given x_i not only receives support from the x_i itself but also from contiguous values x_{i+1} and x_{i-1}. In this sense, extreme values have less support because they have only one neighbor. The same comment applies to extreme values of convex membership functions.

VI. Extending Aggregation Operators in Ordinal Scales

The new defuzzification processes introduced in Section V define, when combined with the selection processes in Section III, some new defuzzification procedures. These procedures are suitable, according to the transformation found in Section IV, for aggregation. In particular, the use of RAGE selection introduces several new aggregation operators.

Some examples of these new operators are the following:

- The combination of convex membership function and the RAGE selection to define the Convex-RAGE defuzzification process.
- The combination of the convex membership function, the WOW transformation and the convex RAGE leads to the CWOW-RAGE defuzzification method.
- The combination of convex membership function, the WOW transformation and the median leads to the CWOW-Median.
- The Choquet transformation (aggregation with fuzzy measure) and the median would lead to the Choquet median.

VII. Conclusions

In this work we have considered defuzzification methods based on the architecture described by Yager in [11]. We have reviewed transformation functions and methods for element selection. The combination of one element of each stage leads to several defuzzification methods. All these methods can also be seen as aggregation procedures. In particular, the use of random selection procedures in combination with membership function transformation procedures lead to new aggregation operators.

VIII. Acknowledgments

Partial support of the European Community under the contract "CASC" IST-2000-25069 and of the CICYT under the project "STREAMOBILE" (TIC2001-0633-C03-01/02) is acknowledged.

References

[1] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, U.K., 1995.
[2] V. Torra, A review on the construction of Hierarchical fuzzy systems, Int. J. of Intel. Systems, vol. 17 pp. 531-543, 2002.
[3] L. A. Zadeh, Fuzzy sets, Information and control, vol. 8 pp. 338-353, 1965.
[4] V. Torra, A. Sodan, A Multi-Stage System in Compilation Environments, Fuzzy Sets and Systems vol. 105 pp. 49-61, 1999.
[5] V. Torra, Aggregation of linguistic labels when semantics is based on antonyms, Int. J. of Intel. Syst., vol. 16 pp. 513-524, 2001.
[6] J. Domingo-Ferrer, V. Torra, Median based aggregation operators for prototype construction in ordinal scales, submitted.
[7] F. Herrera, E. Herrera-Viedma, J.L. Verdegay, A Sequential Selection process in Group Decision Making with a Linguistic Assessment Approach, Information Science, vol. 85 pp. 223-239, 1995.
[8] V. Torra, The Weighted OWA operator, International Journal of Intelligent Systems, vol. 12 pp. 153-166, 1997.
[9] G. Mayor, J. Torrens, On a Class of Operators for Expert Systems, Int. J. of Intelligent Systems, vol. 8, pp. 771-778, 1993.
[10] R. R. Yager, D. P. Filev, Essentials of fuzzy modeling and control, John Wiley & Sons, New York, 1994.
[11] R. R. Yager, Knowledge-based defuzzification, Fuzzy Sets and Systems, vol. 80 pp. 177-185, 1996.
[12] D. Driankov, H. Hellendoorn, M. Reinfrank, An introduction to Fuzzy Control, Springer-Verlag, USA, 1996.
[13] J.-L. Marichal, Aggregation operators for multicriteria decision aid, PhD Dissertation, Faculte des Sciences, Universite de Liege, Liege, Belgium, 1999.